

UC Davis

UC Davis Previously Published Works

Title

Solving Transfer Puzzles: Multi-Factor Explanations

Permalink

<https://escholarship.org/uc/item/8nf1n0z9>

Journal

International Journal of Bilingualism

ISSN

1367-0069

Authors

Filipović, Luna

Hawkins, John A

Publication Date

2026-07-14

DOI

10.1177/13670069261464843

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

Solving Transfer Puzzles: Multi-Factor Explanations

International Journal of Bilingualism

1–20

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13670069261464843

journals.sagepub.com/home/ijbLuna Filipović¹  and John A. Hawkins^{1,2} 

Abstract

Aims and Objectives/Purpose/Research Questions: This article proposes a multi-factor model of second language learning and bilingualism, CASP (an acronym for Complex Adaptive System Principles), which can help solve some puzzles in language transfer, involving when it occurs and when it does not. The key goal is to try to move the field beyond discussions that are focused on just one or two of the many factors of relevance (e.g., typological similarity vs. distance and levels of proficiency).

Design/Methodology/Approach: This study gives a critical and comparative discussion of previous research on word order variation and null versus overt subjects in grammars, as carried out within two different theoretical frameworks, namely the Competition Model and Processability Theory. We explain why conflicting results and explanations have come out of these camps.

Data and Analysis: We use examples from CASP-based research and from recent experimental studies involving different types of bilingual speakers, as well as findings from previous studies that used a variety of empirical methodologies (e.g., fieldwork and classroom research) and theoretical approaches.

Findings/Conclusions: The CASP model accounts for when and why transfer occurs, and crucially, it can explain when and why it does not occur, in many of the bilingual outputs discussed in the field, which has often been elusive to previous research.

Originality: We offer both theoretical and empirical insights that help us solve some of the perennial puzzles in bilingualism and pave the way towards capturing different bilingual outputs of different types of bilinguals who learn and use their languages under different circumstances.

Significance/Implications: We explain why any model of bilingualism must be a multi-factor one, incorporating typological factors, internal factors (such as age, proficiency and dominance) and external factors (i.e., specific interactional circumstances; e.g., single vs. dual language condition; formal vs. informal) and their varying collaborative and competing relationships.

Keywords

bilingualism, CASP, efficiency, pro-drop, transfer, word order

¹University of California, Davis, USA²University of Cambridge, UK

Corresponding Author:

Luna Filipović, Department of Linguistics, University of California, Davis, 270 Kerr Hall, I Shields Avenue, Davis, CA 95616, USA.

Email: lfilipovic@ucdavis.edu

Introduction

Acquisition of another language, especially after the Critical Period (whenever that may be precisely) involves the prior existence of a first language, which can be a help or a hindrance depending on the typological profile of each language. Linguistic theory and applied linguistic research have been divided over the role, importance and impact of transfer (Odlin, 1989) – the mechanism by which the first language (the main/stronger language – L1; usually acquired early and with full proficiency) impacts the acquisition of the weaker second language (L2, with lower proficiency, as in the case of late L2 or heritage acquisition). It seems that some linguistic properties are very productively transferred (e.g., phonetic properties), which is understandable considering that the door for fully native acquisition and consistently native-like production shuts for phonetics early in childhood (most likely by the age of 8, or by the time official schooling begins; see Austin et al., 2019, for a discussion and review). For other linguistic properties transfer effects have not been observed as consistently in all L1-L2 pairs at all times. There is also an ongoing debate with regard whether to full versus piecemeal feature-based transfer is at play in bilingualism and multilingualism (see Guo & Yuan, 2024, for a succinct recent overview), but such discussions tend to focus on one or two of the many factors that do indeed play a role (e.g., typological similarity vs. distance and levels of proficiency). What the field still seems to need is a predictive theory that can account for the different pathways of bilingual language development that may involve transfer on some occasions but not on others, and for some bilinguals but not for others.

In the research programme described and exemplified here we offer an account of how multiple key factors interact in bilingual language acquisition and use, as these have been revealed through a variety of empirical studies. These include the typological relationship between L1 and L2, general principles of language processing (production and comprehension), individual psychological factors (such as age of acquisition, proficiency and dominance) and social factors involving social demographics and prestige and the general environment for learning and interaction (e.g., single or dual language use, with or without code-switching).

Precisely because there are so many different factors affecting bilingual outputs, one of the main problems we see in the research conducted hitherto on transfer has been a failure to control for all the variables other than those being explicitly and ostensibly tested. Different types of bilinguals with different learning histories and different habits of language use have been tested under very different conditions and with different task demands, and they exhibit very different outputs as a result. Equally possible are the same outputs by different types of bilinguals as well as different outputs by the same bilingual under different circumstances, one of which may involve manifestations of transfer and one which does not.

The main novelty of this article lies in its specific two-pronged focus on (a) the theoretical disagreement between the Competition Model and Processability Theory that the CASP model helps resolve, showing where each framework fares well and also where each fares less well, and (b) on providing the critical connections between different empirical findings that involve a wide variety of cross-linguistic data, which were previously not considered together and which are supportive of the theoretical tenets of the CASP model, thus adding to its empirical foundations. We felt that this kind of discussion, tying many threads together, was needed and warranted because the arguments along these lines were dispersed across the literature and the relevant connections were just not being made, especially in the theoretical realm. Our discipline seems to have become more centred on methodology and on methodological diversity than on the (empirically supported) theoretical core.

In Section “Previous Literature: Critical Comparisons,” we discuss a number of relevant studies, with detailed reference to previous theoretical accounts, and in particular, we focus on two critical case studies, concerned with word order and pro-drop, that have been discussed in earlier

frameworks (MacWhinney's Competition Model and Pienemann's Processability Theory) and which illustrate why a new explanatory model is needed. Section "CASP for Bilingualism: A Multi-Factor Model of Bilingual Language Acquisition and Use" introduces the unifying platform of the CASP for Bilingualism Model, which enables us to solve some of the apparently persistent puzzles in the field, and in Section "CASP at Work: Solving the Puzzles," we put the CASP model into action and reconcile the conflicting findings presented in the previous literature. Section "Conclusion" concludes with numerous suggestions for future research and for further testing of CASP's predictions.

Previous Literature: Critical Comparisons

The competing views on the role of transfer are best exemplified by the competing arguments of MacWhinney (1992, 2005) on the one hand, and Pienemann et al. (2005) on the other. MacWhinney's Competition Model is based on learners tracking and validating different linguistic cues (e.g., word order or case marking) as reliable indicators for different structures and meanings in both monolingual and bilingual language acquisition. In the context of bilingualism, L1 transfer is of central importance for L2 acquisition according to the Competition Model (MacWhinney, 1992, 2005; MacWhinney & Bates, 1989) because the L1 cues can either transfer positively into the L2, if the same cues are shared, or they transfer negatively if the same cues are not there in the L2 or not equally reliable. Typological proximity versus distance between two languages is therefore significant, albeit not always the crucial or most reliable predictor of whether a positive or negative transfer will occur. According to Processability Theory, by contrast (Pienemann, 1998; Pienemann et al., 2005), L1 transfer is constrained by general principles of language processing, and the typological distance between L1 and L2 is not of key importance at all (see also Kellerman, 1983; Zobl, 1980, for the origins of this idea). Rather, this theory proposes that L1 transfer is determined by a general and universal processability hierarchy rather than by the typological proximity versus distance in cue validity between two languages. In other words, typological similarities versus differences do not guarantee the occurrence of positive versus negative transfer and instead the degree of, and ease of, processability is the main determinant of what will transfer and what will not.

In spite of many recent advances in the research on transfer (see MacWhinney & Gao, 2025, for a succinct overview), a unifying theoretical account is still lacking that can explain some apparently contradictory findings and solve some of the most persistent transfer puzzles. The main goal of the present article is to address these puzzles and to argue for such a unifying theory, CASP for Bilingualism, which builds on both the Competition Model and Processability Theory and solves many of these puzzles by integrating a broader range of contributing factors showing, in a principled way, how they interact so as to produce transfer with one pair of languages and in one set of circumstances versus the lack of transfer in other languages and circumstances.

A crucial part of our motivation for the CASP model involves paying more attention to cases where there is an apparent lack of transfer. Demonstrating this is not straightforward. It is much easier to notice where transfers have occurred, either positive or negative, across languages. Identifying properties or structures that did not transfer is trickier, because at any one point in time, most of the contrasting properties between L1 and L2 will not, in fact, transfer from one language to the other. Learners of an L2 are constantly striving to learn that L2, with its distinctive sounds and structures, and transfers from L1 to L2, when they occur, are actually generally the exception rather than the rule within the overall totality of properties that comprise a given L2 (their many sounds, grammatical morphemes and constructions, etc.) that need to be acquired. The theoretical challenge posed by looking for, and trying to explain, non-transfers is therefore to find structures in the L1 that could plausibly have been transferred, given universal phonological and grammatical constraints

permitting certain properties to co-occur within a single linguistic system, but that weren't actually transferred, and to then explain why they weren't in the language pair in question.

In addition to the phonological or grammatical compatibility of adding an L1 feature to an L2 grammar, the plausibility for transfer could come precisely from the fact that the feature in question did transfer from L1 to L2 in other language pairs. Or else the feature in question could be such a prominent and central part of the grammar of L1 (with high "cue validity" in MacWhinney's terms), but not in that of the L2, so that there would be no motivation or plausibility for its transfer, leading to no transfer. Or alternatively, the feature could be such a basic part of the processing routines in L1 that its abandonment would possibly cause learning difficulty in L2 (per Pienemann), leading to negative transfer. It is important for any theory of L2 learning to try and explain both transfer and non-transfer, just as it is important for any theory of language change to explain both when changes were "actuated" in certain languages and at certain times, and when they were not, as Weinreich et al., (1968) originally pointed out (see further De Smet et al., 2025, for a general discussion of the actuation problem in historical linguistics, and Hawkins & Filipović, 2025 paper in that volume).

Studies on Word Order

A notable absence of transfer has been recorded in a number of studies on word order. One such is McDonald and Heilenman (1991, p. 331), who report that English learners of French L2 abandon their English word order strategies early and do not transfer them into French, particularly in non-canonical orders. Another study (Kawaguchi, 1999, 2002) found that English learners of Japanese did not transfer their VO word order when learning Japanese, which is an OV language, even at the initial stage. Similarly, Japanese learners of English acquire VO very rapidly without transferring their own rigid and exceptionless OV (Rutherford, 1983). MacWhinney (2005, p. 60) argues that it is the input that enables L2 learners of Japanese to fix the OV word order in the very early stages of acquisition. It may indeed be the case that input plays a key role here, but there are also many occasions when word orders are nonetheless transferred from L1s into L2s whose native speakers do not use these L1 word orders, for example when the distinctive word orders of L1 Spanish, L1 Italian and L1 Chinese are transferred into L2 English (see Hawkins & Filipović, 2012, for examples and review). We cannot assume that the Japanese learners have superior types of input or cue recognition strategies in their L2 English compared to Spanish, Italian or Chinese learners. Moreover, L1 Spanish and L1 Italian learners of English do indeed transfer word orders from their L1s that are more difficult on Pienemann's processability hierarchy and which, according to Processability Theory, should not be transferred, at least not at the initial stage of acquisition (e.g., Verb-Subject order from L1 Spanish is used ungrammatically in L2 English, as in **Yesterday came my boyfriend*; see Filipović and Hawkins (2013, 2019). Di Biase and Kawaguchi (2002) also claim, in support of the Processability Theory, that structures that are more difficult on the processability hierarchy are never transferred at the initial state, regardless of typological constellation, and yet these Verb-Subject transfers into early L2 English counter-exemplify this claim.

It seems that we need a more subtle explanation for what is going on here that neither Processability Theory nor the Competition Model has captured in full. Obviously, it is not just L1 transfer or just processability that is at play, it is their *different interactions in different structures* and among *different language pairs* and *under different conditions* of learning and use that is the key. As we shall see in Section "CASP at Work: Solving the Puzzles," we can reconcile these two apparently opposing views on the primacy of L1-L2 difference versus processability within the context of an integrated and more multi-factor theory and offer a more nuanced explanation for when and why transfer may or may not occur.

Null Versus Overt Subjects

Another area in which the effects of transfer are constrained by, and interact with, other factors in second language acquisition, is exemplified by null versus overt subjects. For instance, Pienemann et al. (2005) cite many studies with L1s of either the overt subject or pro-drop type and with L2 being a pro-drop language (e.g., Liceras & Díaz, 1999; Phinney, 1987), and they note that all these studies reported a consistent early appearance of null subjects (i.e., pro-drop) in all learners regardless of L1. However, *early appearance does not equal early full acquisition*, as Hawkins and Filipović (2012) point out. And it appears that at least some variation in the degree of success in acquiring pro-drop rules is indeed driven by the L1 as well as by ease of processing.

First, with respect to the processing of this alternation, dropping a pronoun in the subject position of a main clause in L2 acquisition is not hard to explain – an omitted subject in this position is easily recognized and acquired by all learners (see Section “To Pro-Drop or Not to Pro-Drop” for more details). For example, a study by Pladevall Ballester (2010) on the acquisition of L2 Spanish by different L1 English children and youth learner groups (aged 5, 10 and 17) showed that all learners found pro-drop more acceptable in the main clause than in subordinate clauses. Thus, it seems that null subject acquisition is generally easier and faster in some positions but more difficult in others, and this has been shown to hold for different learners with both pro-drop and non-pro-drop L1s acquiring Spanish, a pro-drop language, as an L2; see Lubbers Quesada (2015) for a thorough critical review.

Another factor that makes a difference for the successful acquisition of pro-drop is proficiency. Montrul and Rodríguez Louro (2006) predicted that less proficient L2 learners would start producing null subjects early but would show a lower percentage of usage than more advanced learners and native speakers. Their results with L1 English/L2 Spanish learners at three different levels of proficiency (intermediate, advanced and near-native) supported this. All three groups of learners acquired the “pro-drop vs. overt subject” patterns but to a very different degree: the intermediate group produced more overt than null subjects compared with the advanced and near-native groups, reflecting the overt subject grammar of their L1 English. The two higher proficiency groups performed quantitatively and qualitatively like native speakers of Spanish. These authors also noticed that when the intermediate group did use null subjects accurately, this usage was especially visible in structures in which the grammar of English also allowed it (e.g., when there was a null subject in coordinate deletions with identical subjects in each clause; see also Miyamoto & Yamada, 2019, for a related finding with L1 German transfer into L2 Japanese).

Similar findings showing subtle transfer effects have been reported in Sorace et al. (2009), whose study included English-Italian (from the United Kingdom and Italy) and Spanish-Italian (from Spain) bilingual children (aged 6–7 and 8–10), as well as monolingual controls (from the United Kingdom and Italy). The younger children living in the United Kingdom were more likely than those living in Italy or Spain to select an overt pronoun in non-topic-shift contexts, revealing cross-linguistic influence from English. Their study also showed that certain usage constraints governing null and overt options in Italian are more difficult to learn and develop over time. Again, it seems that both L1 factors and general processing ease have a role to play in explaining the full L2 data.

We also need to consider reports from heritage language acquisition studies. They show a rare agreement when it comes to “pro-drop resetting” among immigrant communities in the United States, for example (Heine & Kuteva, 2005: 99). Numerous studies cite overuse of overt subjects across the board in heritage languages spoken in an English-dominant environment, for example, in heritage Russian (Schmitt, 2000), Spanish (Myers-Scotton, 2006), Serbian (Savić, 1995), Polish (Rappaport, 1990), Tamil (Polinsky, 1995) and Hungarian (Fenyvesi, 1994). This clearly indicates language transfer from the stronger L1 (English).

Finally, it is also important to look at studies that report on acquisition in the opposite direction, when the L2 has a strict overt subject rule, as in English. For instance, L1 Spanish/L2 English learners often transfer their pro-drop pattern into early L2 English (e.g., **is a beautiful country* instead of *it is a beautiful country*; see Hawkins & Filipović, 2012, for many examples and discussion). Processability Theory actually predicts L1-independent early acquisition of the correct English forms here because null subjects are placed at the same level as pronominal subjects in the processability hierarchy (Di Biase & Kawaguchi, 2002; Pienemann et al., 2005), and yet we see a very clear L1 transfer effect. Second language learners of English with a first language that is a pro-drop omit their subjects in L2 English ungrammatically, that is, L1 Spanish learners of L2 English use pro-drop in their L2 even though it is ungrammatical (Filipović & Hawkins, 2013, 2019; Hawkins & Filipović, 2012), which seems to contradict the prediction of the Processability Theory for early L2 acquisition of overt subjects in this case.

A related unexplained puzzle is that many studies of Romance pro-drop languages as L1s have shown significantly more difficulty for their learners of English than L1 Japanese and L1 Chinese speakers do when learning the overt subjects of English, even though Japanese and Chinese are also considered pro-drop languages (Roebuck et al., 1999). Studies that have compared the different proficiency levels of L2 learners (such as Roebuck et al., 1999) have shown that L1 Chinese learners performed much better than the L1 Spanish speakers at both lower and higher proficiency levels in L2 English in this regard. Thus, it seems that learning overt subjects is more difficult from the very beginning for L1 Spanish speakers than for L1 Chinese learners – this surely is a sign of different transfer effects from different L1s. These effects can be expected to be gradient (Neeleman & Szendrői, 2005) – languages will often allow pro-drop to the extent that their verbal agreement paradigms (as in Romance languages) or discourse processes and topicalization (Chinese and Japanese) enable local recovery of the content of dropped arguments (see also Xu & Yuan, 2022, for the role of pragmatics in the acquisition of null subjects in L2 Chinese by L1 English speakers). Thus, typological proximity versus distance, and the grammatical profile of the L1, are extremely important here, and the explanation for these different outcomes in L2 English must take into account, and can in part be traced back to, differences among these features of the L1. Crucially, typological proximity needs to be tackled within a finely grained approach. As Guo and Yuan (2024) demonstrate in the context of Mandarin L3 acquisition by English-Cantonese bilinguals, the similarity between Cantonese and Mandarin can lead to a detrimental effect due to a lack of exact feature overlap. They explain that because the question word *ma* [嗎] has two functions in Cantonese (a question and a surprise reaction) and only one (the former) in Mandarin, both these functions get transferred from Cantonese (the stronger language, L1 or L2) into Mandarin (L3) (see also Filipović, 2017, 2019, on partial typological overlap of features in bilingual acquisition and use).

More generally, by delving deeper into the respective grammars and usage patterns of the different language pairs, we can better understand the different outputs in L2 English, through a mix of transfer, processing ease, language typology and performance data as found in the relevant language types. Spanish and Italian license only null subjects, but Chinese and Japanese permit both null subjects and null objects. Because of their much wider pro-drop practice, Chinese and Japanese have been classified as “radical pro-drop languages” (Neeleman & Szendrői, 2005). The null subjects of Spanish and Italian can often be recovered and identified on the basis of their rich verbal inflections. Such is not the case with Chinese and Japanese, however, as these languages lack subject–verb agreement. Instead, null subjects/objects in Chinese and Japanese are reconstructed by discourse context and “topic chains” (Huang, 1984; Roebuck et al., 1999). For this reason, some scholars consider null arguments in Chinese and Japanese to be instances of “topic deletion” rather than “pro-drop” (Roebuck et al., 1999, p. 256). Therein may lie the explanation as to why different

L1 learners behave differently in the same L2. Grammatical licensing of pro-drop in an L1 may be a bigger hurdle than the discourse-licensed L1 pro-drop for the acquisition of a non-pro-drop L2.

Insights from heritage bilingualism are again relevant here. They indicate that the respective heritage speakers of English never fail to produce overt subjects in English, regardless of the stronger L1 (i.e., the language of their country of residence). Nor is there any reverse effect of oversupplying overt subjects in their dominant pro-drop L1. Pro-drop stays pro-drop, and heritage English maintains its overt subjects (Polinsky, 2018). This linguistic behaviour is in stark contrast to L1 speakers learning L2 English as a foreign (and not a heritage) language, who do seem to transfer their pro-drop patterns into the overt subject L2 English in certain cases, as we have seen, and therein lies the key difference between when we do and do not have transfer (namely, it depends on the bilingualism type, the level of proficiency, the language pair, and on whether the grammatical feature in question is typologically permissible or not).

CASP for Bilingualism: A Multi-Factor Model of Bilingual Language Acquisition and Use

In this section, we showcase the model, CASP (*Complex Adaptive System Principles*) for Bilingualism, which was inspired by the literature reviewed in the previous section and which has been supported by extensive and diverse datasets, including the largest learner corpus of English as an L2 (the Cambridge Learner Corpus—CLC) and experimental data elicited in a number of psycholinguistic studies (Filipović, 2019, 2025; Koster & Cadierno, 2019). All of these datasets and methodological approaches point to the same unifying idea, namely the need to view bilingualism as a complex adaptive system, which we have argued for on multiple occasions, as have other scholars (e.g., Larsen-Freeman & Cameron, 2008).

In the case of CASP the theoretical basis came from two sources, from Murray Gell-Mann's work on complex adaptive systems in the natural sciences (see Gell-Mann, 1992; Hawkins & Gell-Mann, 1992) and from Hawkins' (1994, 2004, 2009, 2014) Theory of Efficiency which has been extensively tested and applied to multiple studies in theoretical linguistics, language typology, psycholinguistics and computational linguistics (see, for example, Kirby, 1999; Levshina, 2023; Newmeyer, 1998; Song, 2011). The basic premise of Hawkins' theory, which was developed using cross-linguistic grammatical patterns and correlating patterns of usage and preference within individual languages, is that efficiency is paramount in any act of communication. Efficient transmission of information means transferring a message from the speaker to the hearer quickly and with minimal effort and maximal gain. In other words, maximize the communicative benefits for minimum processing costs (see especially Levshina, 2023, for a recent book-length general summary). When the speaker needs to process two languages at the same time, the importance of efficiency becomes even greater.

Note that *efficiency is not the same as economy*. Economy as a driving force in general (as well as bilingual) language usage has been championed most notably by Muysken (2013), among others. But economy does not always lead to efficiency. We are sometimes more efficient when we use less economical means. For example, if somebody wants to identify a particular referent for the hearer in the most efficient way, they might say "*The professor we talked about yesterday showed up in my class today!*" instead of just "*The professor showed up in my class today*". The longer, less economical phrase can, on various occasions, be more efficient because it describes and identifies the referent faster and avoids follow-up questions and additional interaction (e.g., with the hearer having to ask "Which professor exactly are you talking about?"). It is also important to note that efficiency can sometimes be suppressed, for example, when a speaker purposefully delays conveying information for some kind of special effect, such as humour, poetry or even seduction (see

Filipović, 2019, p. 59) for more details on this point. Such uses of language are much less common than adherence to efficiency in the general transfer of information in the normal case. Efficiency, as mentioned above, has been supported as the explanatory mechanism behind different types of data, in the grammatical conventions of different languages and typological patterns of variation (Hawkins, 2014), in language processing (Gibson et al., 2019), in monolingual and bilingual language acquisition (Filipović, 2014, 2019), and also in recent physiological studies on the bilingual brain (Gracia-Tabuenca et al., 2024). This general view of language and language acquisition belongs within the rich emergentist tradition in psycholinguistics (e.g., O'Grady, 2005, 2008) and neurolinguistics (Blanco-Elorrieta & Caramazza, 2021; Hernandez et al., 2015, 2019; Stoco, 2019).

The CASP model consists of five general principles that together form the mechanisms behind both bilingual language learning and processing. The earlier version of the model (Filipović & Hawkins, 2013) targeted unbalanced bilingualism and second language learning, but the more recent version (Filipović & Hawkins, 2019) captures all types of bilingual learning and processing.

The main difference between the 2013 and 2019 versions lies in the addition of the fifth principle, Maximize Common Ground, which is unique to bilingual as opposed to monolingual learning. This new principle is especially relevant for issues of transfer, positive and negative, and for explaining effects, in conjunction with the other general principles, that were previously subsumed under various derivative subprinciples (Maximize Positive Transfer, Permit Negative Transfer and Block Negative Transfer).

The five principles are shown below. They enable us to capture the different types of bilingual outputs, for different types of bilingual speakers under different social circumstances of interaction. They collaborate and reinforce each other sometimes, and sometimes they compete. There is no *a priori* hierarchy among the principles and their relative strengths are modulated by two sets of factors, *internal* (such as proficiency, age of acquisition, dominance) and *external* (such as who the bilingual is talking to and on what occasions, for example, formal vs. informal). These internal and external factors condition the strength and the ultimate outcome of the push-pull relationships that the five principles enter into under different conditions and they underscore the resulting dominance of one or more of the principles that work together towards the same goal. These two groups of factors are precisely the reason why the same bilinguals can produce different outputs on different occasions and why different types of bilingual speakers can produce the same output on some occasions. In other words, understanding the interactions among the principles under the pressure of different internal and external factors enables us to explain why sometimes we see transfer and sometimes we do not, even when we are testing one and the same bilingual speaker!

CASP for Bilingualism Principles:

1. Minimize Learning Effort: Effort is reduced when grammatical and lexical properties are simpler and more frequent – they will be learned earlier in both monolingual and bilingual acquisition stages, in the latter case especially when these properties are shared in both languages (see Principle 5 below). Bilingual speakers will favour such properties and use them even more frequently when both languages are active.
2. Minimize Processing Effort: Simpler structures are processed (i.e., recognized, accessed and produced) more quickly than complex ones by both monolingual and bilingual speakers, and thus preferred within one language and across two languages of a bilingual if they exist in both.
3. Maximize Expressive Power: Both monolingual and bilingual speakers want to formulate and express all their thoughts equally well in one or both their languages respectively; this

principle is partially opposed to the previous two principles because it motivates the eventual learning and use of more complex and less frequent linguistic structures, which are harder to learn, process and use.

4. **Maximize Efficiency in Communication:** Monolingual and bilingual speakers prefer to maximize efficiency in communication; this principle reflects a trade-off between minimizing learning and processing effort and maximizing expressive power in different communicative situations. Although speakers tend to use simple linguistic structures, sometimes they need more complex structures to communicate their intended information more successfully or sooner (Hawkins, 2004; Levshina, 2023). For efficient communication, speakers use a simpler grammar or lexicon whenever possible and greater complexity when necessary (e.g., if no prior conversational context is available, a referent, *the professor*, can be identified more successfully and faster by adding the relative clause as in *the professor we talked about yesterday* cited earlier).
5. **Maximize Common Ground:** Bilingual speakers maximize common grammatical and lexical representations in their two languages. When two languages have the same or very similar grammatical structure or lexical items, bilinguals prefer to use those shared properties more frequently (which would fall under positive transfer). When the languages do not share common properties, bilinguals create common ground by introducing new properties from one language into the other, leading possibly to ungrammatical structures or lexical conventions in the relevant language (previously negative transfer) or by avoiding the use of these non-shared properties altogether when they can.

The first four principles characterize any language learning and processing event, whether monolingual or bilingual. Principle (5) is the bilingualism-specific principle.

These principles can reinforce each other, for example, the frequently occurring items that lead to early learning by principle (1), Minimize Learning Effort, are also often structurally simple, so minimizing processing effort by principle (2). Sometimes the principles compete and produce variable outputs and alternative interlanguages. For example, principle (3), Maximize Expressive Power, necessitates the learning of less frequent and more complex structures in some cases, at variance with the minimization principles (1) and (2). Positive transfers (“good translations”) are maximized on account of general principle (1), that is, they are always advantageous for learning, and they are further motivated by principle (2) Minimize Processing Effort, and also principles (3) Maximize Expressive Power and (4) Maximize Communicative Efficiency (see Filipović & Hawkins, 2013, for full discussion). Principle (5), Maximize Common Ground, reflects the drive towards *processing efficiency* among bilingual learners and it captures a generalization that needs to be recognized and that is not always apparent, given the different labels “positive” and “negative” transfer: the same general efficiency-based processing mechanism of Maximize Common Ground underlies both positive transfers and negative transfers (“the bad translations”); see Filipović & Hawkins (2019), and also the discussion that follows.

Maximize Common Ground enables us to understand phenomena that have been variously described as “convergence,” “positive transfer,” “bidirectional influence,” “bilingual-specific language use” or “in-between performance” (see Pavlenko, 2014, for discussion). Different terminology has been used here to capture what are essentially the same phenomena, differing only (e.g., for convergence vs. transfer) based on who is providing the bilingual outputs (balanced bilinguals vs. L2 learners; see Athanasopoulos, 2011). In CASP, we propose that the underlying principle guiding these outputs is the same: Maximize Common Ground is the most efficient option, and all bilinguals are doing it, with variable success in terms of form, meaning and usage conventions, and to different degrees depending on their interlocutors and purpose. We know that bilinguals share

representations as much as possible at all levels of linguistic analysis (e.g., lexical and syntactic; see Filipović, 2019, for a thorough critical overview). If proficiency is high in both languages, then bilinguals will maximize common ground in a way that is grammatical in both languages, and if not they may do so in an ungrammatical way in the weaker language. Recent neurolinguistic studies on shared representations and activation support the behavioural findings in focus here (Blanco-Elorrieta & Caramazza, 2021).

It is vital to emphasize that the five general principles we have put forward are modulated by both internal (psycholinguistic) and external (sociolinguistic) factors. For instance, the same bilingual speaker will produce different outputs depending on proficiency level in each language and also depending on whether the interlocutor is another bilingual speaker of the same two languages or a monolingual of one of the two languages (see also the discussion in Muysken, 2013, p. 714 on different factors that impact outputs in language contact situations). Long-term language change is also fundamentally dependent on these two factors because, for example, when certain types of bilinguals, for example, L2 users, interact with monolingual speakers of the bilinguals' L2, and crucially, if the bilinguals outnumber the monolingual speakers in a community, then the language of the monolinguals may change under the influence of those who speak it as an L2 (see Trudgill, 2010, 2011).

Internal factors may include age of acquisition, proficiency, type of input, frequency of use, which can all sway our predictions in different ways and make bilingual outputs variable (see especially Jarvis & Pavlenko, 2007, for a detailed discussion of a number of these factors). External factors are driven by the inherently adjustable nature of bilingual linguistic behaviour, which depends on the interlocutor types involved (i.e., who bilinguals are talking to) or the type of communicative situation a bilingual is involved in (e.g., formal vs. informal; see Dewaele, 2001). Grosjean (1992) defined these different occasions of use as *language modes*, whereby a monolingual mode is characterized by the higher activation of one of the two languages of the bilingual because only one language is being spoken in a particular communicative situation. The alternative, bilingual mode, is when a bilingual has to maintain the same or similar level of activation of both languages because the communicative circumstances require this.

A related idea comes in the form of the adaptive control hypothesis (Green & Abutalebi, 2013), which distinguishes *single language*, *dual language* and *dense code-switching* conditions. The single language is the equivalent of the monolingual mode, and the two other conditions are related to the bilingual mode with the crucial difference that the dual language condition occurs when a bilingual is talking to two monolinguals in the same communicative situation (cognitively the most demanding situation type) and there can be no code-switching, while in the dense code-switching condition the bilingual is talking to other bilinguals who share the two languages. Since code-switching is allowed there is less need for a very high level of control for one of the two languages. These different types of external conditions will affect the outputs significantly. Dewaele (2001) showed that external situation factors regulate the amount of code-switching – there was much more code-switching in an informal class conversation situation compared to a formal exam situation. We can also expect that a more formal situation may result in fewer instances of maximizing common ground than an informal one since the mechanisms of cognitive control that constitute the “output monitor” (see de Groot, 2011, for details) may be on a higher alert in the former rather than the latter. Finally, these situation-driven outputs interact strongly with the internal factors such as relative proficiency: bilinguals with balanced proficiency in both languages are likely to be more successful in output monitoring than those whose proficiency in one or both languages is uneven (i.e., L2 learners or heritage speakers; see Filipović, 2019, for further discussion and empirical confirmation).

In a nutshell, CASP's general principles are there to help us formulate our initial predictions based on both typological factors (e.g., grammatical features in the two languages in focus) and general

processing and learning factors as reflected in our five principles (minimizing learning effort and processing effort while achieving expressive power, communicative efficiency and processing efficiency), which are then moulded into final predictions for our bilingual outputs by (1) the linguistic profile of our bilingual speaker and (2) who our bilingual is talking to. CASP predicts that instances of negative transfer due to typological contrasts that do not affect communication may either be eliminated at later stages or persist throughout acquisition (e.g., the transfer of Verb-Subject word order from L1 Spanish into L2 English, or article usage in L2 English by speakers of different L1s; see Filipović & Hawkins, 2013; Hawkins & Filipović, 2012, for details), depending on the length and purpose of L2 learning. Errors that do significantly impact communication (such as the transfer of SOV word order from L1 Japanese into L2 English) will be blocked early. It is important to highlight yet again that these CASP predictions will be modulated by the all-important internal (e.g., proficiency) and external (e.g., social circumstances of use) factors. More proficient bilinguals will not transfer their L1 Spanish word order (or will transfer it less often) when they speak L2 English and the transfer, if it appears, will be less frequent if the communicative situation is monolingual (English-only) and the bilingual relative proficiency in both languages and control capability is high.

CASP at Work: Solving the Puzzles

Word Order Transfers, Sometimes

As discussed in Section “Previous Literature: Critical Comparisons,” the reported lack of transfer in L2 acquisition for a number of syntactic features, such as word order or pro-drop in some language pairs, combined with their different degrees of transfer in others, has been a long-standing puzzle in the field for which neither the Competition Model nor Processability have provided fully satisfactory solutions. Especially puzzling is the fact that speakers of a language that has much typological similarity with English when it comes to basic word order (e.g., Spanish) can transfer their un-English word orders from Spanish, while speakers of a typologically distinct language like Japanese, whose word order is very different from English, do not transfer their L1 word order into English L2. Negative transfer is permitted in the case of L1 Spanish learners of English, but not in the case of the L1 Japanese learners of English.

What is the reason for this apparently disparate linguistic behaviour? In the case of L1 Japanese learners of L2 English, the CASP Principles (1), (2) and (5) [Minimize Learning Effort, Minimize Processing Effort and Maximize Common Ground] are trumped by the other two Principles, (3) Maximize Expressive Power and (4) Maximize Efficiency in Communication. All things being equal, learners want to be engaged in less effortful acquisition and production activities and use shared structures whenever possible, which would motivate transfer. However, the impact that the imported word order would have on communicative success and communicative efficiency in the L2 is too highly compromised. Transferring a Japanese rigid SOV into L2 English would result in extreme communicative inefficiency: speakers importing these diametrically opposed Japanese head-final orders into English would simply not be understood (cf. e.g., **I a few days ago the cinema to went*; see Hawkins, 1983, 2014). So it is not just a matter of how many principles are acting together on one side, but what the relative cost of their “victory” would be as well as the size of the benefit they bring to the outcome. The prevailing two principles in this case bring a much higher benefit – the speaker will be understood if the transfer does not happen, and this secures the ultimate goal of communication – to be understood. By contrast, the strength of these two winning principles in the Japanese case is weakened in Spanish, where transfer of L1 Spanish word orders into L2 English does happen. Namely, Spanish-type variants of English head-initial word orders such as **Yesterday came my boyfriend* and **I read yesterday the book*, for which similar

Verb-Subject and Direct Object Postposing parallels already exist in English (e.g., *Away ran the boy* and *I finally read yesterday the book that I had been trying to find for months and months*, see Emonds, 1976), are negatively transferred. This is because (a) these transfers are “cheaper”: they do not prevent successful communication so the cost of the transfer is low and (b) the CASP principles of (1) Minimize Learning Effort, (2) Minimize Processing Effort and (5) Maximize Common Ground operate as stronger and more established in Spanish because Spanish and English do share a lot of word order options, so the general paths of word order transfer are already “well-travelled” and learners just continue along them until the internal factor of proficiency and the external factor of monolingual interactions in L2 English correct this and “curb the enthusiasm” of the transfer-inducing principles (1), (2) and (5). If the two factors, proficiency and L2 monolingual mode exposure, do not happen, fossilization sets in. Further examples are found in the acquisition of L2 English by L1 Chinese (**I by bus go to university*) and L1 Italian speakers (**I like very much sweets*; see Hawkins & Filipović, 2012 for details).

This appeal to communicative efficiency and success can be extended to many other transfer puzzles that are related directly or indirectly to word order. For instance, errors like the overuse of topicalization by Chinese learners and by French learners (see Trevis, 1986) or underuse of the passive by Hebrew learners (cf. Seliger, 1989) can be explained by the fact that they do not affect the ability of learners to make themselves understood by their hearers, that is, they do not impact efficiency and success in communication so Principle (4) Maximize Efficiency in Communication is again strong here, and is further supported by CASP’s other principles, namely Minimize Learning and Processing Effort, and Maximize Common Ground (i.e., the relevant forms and structures exist in both languages and are used in both albeit with different frequencies). Since communication is not impeded, there is little incentive to reduce or increase their frequencies of use compared with the L1. With all these four principles (1, 2, 4 and 5) acting in unison, Maximize Expressive Power (Principle 3) is not strong enough on its own to drive the incentive of acquiring the native-like usage patterns.¹

By contrast, L1 Chinese prenominal relative clauses are not overused and transferred into L2 English (i.e., **the woman loves whom man* does not feature in L2 English instead of *the man whom the woman loves*) because this is a complex and typologically marked structure in Chinese (Hawkins, 1999, 2004). Complex lexical or constructional meanings in an L1 without an L2 equivalent will not generally transfer negatively if they seriously impede communication. As in the case of L1 Japanese word order in L2 English before – direct transfer would lead to reduced communicative power and result in extreme communicative inefficiency, which Principle 4 actively blocks. However, again, internal factors, such as proficiency and stage of development (adult vs. child learners) can affect the predicted outcome here because they can modulate the collaborative non-transfer output that we find in adult data. In fact, we do get exactly this type of transfer in child bilingual acquisition with L1 Cantonese/L2 English, as documented in Yip and Matthews (2007) (e.g., in the L2 output **You buy that tape is English?* which is directly imported from Cantonese grammar and corresponds to the English post-nominal relative clause *the tape that you bought*). This is not altogether surprising since adult versus child language acquisition differences have been well-documented (Saville-Troike, 2009, p. 82) – adults seem to have an edge when it comes to the initial phase of learning. It is also the case that the bilingual parents of the children who produce these transferred structures provide positive feedback because they do understand what the children mean and they respond to them appropriately, so extreme communicative inefficiency does not occur. With different external factors (e.g. monolingual mode of communication in English-only), the power of Principles (3) and (4) would have been enhanced and the outcome would have been different, comparable to the adult ones in this context. Unlike the Verb-Subject word order transfers from L1 Spanish to L2 English discussed above, the incentive to eliminate this

L1 Cantonese structure from L2 English is now high because the use of Cantonese prenominal clausal modifiers in English is almost unprocessable and totally impedes communication – adult L2 learners evidently “know” this and have the requisite sensitivity to their audience, but children do not yet.

Thus, we can conclude, that L1 word order will sometimes transfer into an L2 and sometimes not, and that both outcomes are guided by general CASP principles applied to different pairs of languages and cooperating and competing in the ways we have illustrated, subject to internal (psycholinguistic) and external (sociolinguistic) factors that ultimately determine their relative strength and applicability. The different outputs in different types of bilingualism are a consequence of typological proximity versus distance (which results in lower vs. higher communicative failure when a transfer is made), internal factors (such as age or proficiency), and external factors (whether both languages or a single language is used; whether code-switching is possible or not, Dewaele, 2001; Green & Abutalebi, 2013; whether social demographics and prestige favour transfer, Trudgill, 2011;). Many of CASP’s predictions will accordingly be gradient rather than discrete, reflecting the extent to which, and the strength with which, these different factors apply (see Filipović & Hawkins, 2013, 2019; Hawkins & Filipović, 2012 for further discussion and extensive illustrations; see also Filipović, 2019 for details on how to make predictions using CASP while ensuring their falsifiability).

To Pro-Drop or Not to Pro-Drop

Another transfer puzzle that was touched upon in the literature review (Section “Null Versus Overt Subjects”) involved the acquisition (or lack thereof) of pro-drop. Pienemann et al. (2005) state that pro-drop rules in L2 are acquired early regardless of whether the L1s do or do not have pro-drop, and they explain this as being due to the general early accessibility of this feature (pro-drop vs. overt subject). By this logic, speakers of pro-drop languages should not have a problem acquiring the corresponding rules of an L2 in which there is no pro-drop, since this feature of the processability hierarchy is claimed to be universal regardless of L1-L2 typological differences. However, as we mentioned in Section “Null Versus Overt Subjects,” this seems not to be the case. The Spanish pro-drop pattern is indeed transferred into early L2 English by L1 Spanish speakers (e.g., **is a beautiful country* instead of *it is a beautiful country*). Furthermore, heritage Spanish speakers whose stronger language is English tend to shun the pro-drop in Spanish and overuse the pronouns as overt subjects under the influence of their stronger language, English (Montrul, 2004, 2008, 2015). Both of these different outputs, dropping the pronoun when you should not (in L2 English) and using it when you shouldn’t (in L2 Spanish), are clear transfer effects that are not captured by Pienemann’s Processability Theory. And where there is a lack of transfer, in this and other areas, this is problematic for the Competition Model. Both of these transfer puzzles can be explained by CASP because CASP is a multi-factor model, and its multiple factors, as opposed to the single factors or foci of other models, operate at all times in bilingual acquisition and use. Principle (3) Maximize Expressive Power on its own is not strong enough here to encourage learning of the contrasting structures because the cost of not having this knowledge is not that high and the other principles collectively support this transfer-based outcome. Principle (5) is very active here, and it supports Principles (1), (2) and (4) because the outcomes are efficient, the message gets through as it is, both with excessive pronoun use and when it is dropped ungrammatically.

Let’s look at the interaction of these principles in more detail. In the case of L1 Spanish speakers who transfer pro-drop ungrammatically from their L1 into L2 English, we see that Principle (5): Maximize Common Ground supports this negative transfer because the cost to communicative efficiency is not high enough (thus Principle 4 is not acting strongly against it). In other words, the

removal of the subject in their L2 English does not impede communicative success, and this negative transfer is thus permitted. Specifically, the erroneous pro-drop output by L1 Spanish learners of L2 English is supported by CASP Principles (1) and (2) involving minimal learning and processing effort, because learners are simply transferring structures that they already have into the L2, and it is further supported by Principle (5) while Principle (3) Maximize Expressive Power remains inactive here because the incentive for learning is low – the message already gets through.

The contrary output, that is, the overuse of overt subjects in Spanish by heritage speakers or second language learners who have English as the L1 (stronger) language and Spanish as the L2 (weaker) language is also motivated by the drive to minimize learning and processing effort (i.e., adhering to the simple and frequent patterns and avoiding the learning and use of different patterns in L1 vs. L2). All learners regardless of their L1 are incentivized to use pro-drop patterns in a pro-drop L2 in salient environments (e.g., main clauses), on account of its structural simplicity, frequency, lower processing effort and high communication efficiency (CASP Principles 1–4), though a full acquisition in all positions may be delayed and modulated by the relevant L1 rules and usage, as previous research has argued (Montrul & Rodríguez Louro, 2006; Sorace et al., 2009). The initially active Maximizing of Common Ground that encourages overuse of overt pronouns (because doing so is not ungrammatical in pro-drop languages) eventually gets “toned down” with increased proficiency.

In sum, all of the bilinguals, whether their stronger language is of a pro-drop type or not minimize learning and processing effort and optimize processing efficiency by maximizing common ground (using the same patterns in both languages) and they achieve communicative efficiency of getting the message across fast and with least effort, albeit in two different ways with either overuse or erroneous use of pro-drop in their weaker (heritage or L2) language, which is due to the internal factor of relative proficiency in the L1 versus L2. We see no incentive here to maximize expressive power (Principle 3) because it does not add much to the outcome of communicating the message with minimal effort and maximal gain, until the key *internal and external factors* act upon it and sway the interaction of the principles in another (grammatical or native-like usage) direction (with acquired higher proficiency or through exposure to monolingual mode interactions).

It is important to highlight here that heritage bilinguals and advanced L2 learners are more likely to know when and how to maximize common ground in a way that is grammatical in both languages compared with early or intermediate L2 learners (see Hawkins & Filipović, 2012; Montrul & Rodríguez Louro, 2006). For example, early L1 Spanish/L2 English learners have not yet “worked out” how best to maximize common ground in their weaker (L2) language. By contrast, balanced and unbalanced proficient bilinguals and proficient heritage bilinguals usually have this “figured out”, or at least they do it better than late bilingual learners (see Filipović, 2019, for further details and many other examples).

Again, it should be emphasized that Principle (5) Maximize Common Ground operates by encouraging bilinguals to use a common pattern, but the extent to which this will be implemented in their outputs will be gradient in response to the internal and external factors that modulate all of these factors: more in dual language activation (when both their languages are equally active, Green & Abutalebi, 2013) than in single language activation; and in proportion to proficiency (more monolingual-like outputs and less common ground with higher proficiency and in the single rather than dual language condition).

In addition, the acquisition of pro-drop will be universally modulated by structural factors that determine the ease or difficulty of using specific instances of pro-drop and of recovering the reference of the null subject – morphosyntactic cues within simpler main clause structures that are easy to process and salient (see Montrul & Rodríguez Louro, 2006) will be acquired more readily (following Principles (1) and (2) that support minimal learning and processing effort) and sooner than those involving language-specific and complex syntactic or discourse pragmatic properties that are less

straightforward to master (of the type that CASP's Principles (3) and (4) apply to involving the acquisition and use of more complex features). All of these factors operating together can lead to different outcomes among different pairs of bilinguals, distinguished according to the kinds of internal and external factors enumerated above, namely grammatical use of pro-drop, overuse, and underuse in L2, by different bilinguals as well as by the same types of bilinguals under different circumstances.

Conclusion

The view that “anything that can transfer will” as advocated by the Competition Model needs to be revised to “anything that can transfer will under the right circumstances.” In other words, whether transfer will happen or not depends on multiple factors: the typology of the languages in question, universal processing factors, (internal) psycholinguistic factors pertaining to the individual learner, and (external) sociolinguistic factors, all of which condition bilingual learning and processing. We now have a unifying theory (of efficiency; Hawkins, 2004, 2014; Levshina, 2023) and a model of bilingualism (CASP) that helps us make predictions about when we do and when we do not expect transfer, and crucially, that enables us to explain why it is sometimes there and sometimes not for different types of bilingualism and different circumstances of use (single vs. dual vs. code-switching conditions) as well as for the same types of bilingualism under different conditions of interaction.² Considering multiple factors at the same time is necessary because their presence, relative strength and possible interactions can impact our predictions for the occurrence or blocking of transfer.

We have known for a while that both L1-specific and general processing constraints are in operation in L2 acquisition but what this article has brought to the fore, which was not available before, is the multi-factor account of how these constraints interact and why they result in apparently contradictory bilingual outputs. We now also have more persuasive evidence that if transfer-driven production impedes communication in the L2, the relevant outputs will either not be found to begin with (as exemplified by the non-transfer of L1 Japanese rigid verb-final word order into English L2) or they will be eliminated and calibrated to the degree of communication impediment, learner motivation and educational level in the L2. We have seen in this article that the absence of transfer is driven by multiple factors, not just processability and not just the L1 component properties and their contrasts with L2, and that it leads to different outcomes for different types of bilingual speakers of the same language combinations (see in particular Barking et al., 2023/2025, for transfer effects due to schematicity of constructions and individual speaker differences in bilingual production).

CASP also defines numerous predictions that can test, and possibly refine, these multiple factors and their interactions. For example, we predict more maximizing of common ground in dual than single language conditions within a communicative situation, and more in the weaker versus the stronger language if proficiency is unbalanced, but less maximizing of common ground (i.e., less transfer) if communication is impeded due ultimately to typological distance. How these factors play out in different L1-L2 combinations must be tested empirically, and is being tested in studies that have already provided numerous insights (e.g., Filipović, 2025; Hijazo-Gascón, 2021; Lewandowski & Özçalışkan, 2021; Nicoladis et al., 2024) and that will hopefully continue to refine the theory in the future with different L1-L2 combinations being learned and used by different types of bilinguals in different social contexts.

Maximize Common Ground can lead to both the gain and loss of linguistic features over time. Some distinctions are universally easier to “lose” than others, however. Bilingual speakers are found to avoid lexical and constructional meaning distinctions that are available in only one of their two languages (see Pavlenko, 2014, for a detailed overview). Mirror image and opposite word orders (i.e., SVO vs. SOV), on the other hand, cannot readily be transferred across languages, as we have seen, though such transfers do occur through gradual language change and under

extensive long-term language contact and bilingualism (Ross, 1996, 2007). This happens over longer timescales, in historical language change, and therein lies another application of CASP—explaining when, where and why we do or do not have bilingualism-induced language change (see Hawkins & Filipović, 2024). Historical linguistic research has also shown that losing features is more common in cases of unbalanced (adult, late) bilingualism (Bentz & Winter, 2013) while gaining them is more likely with balanced (child, early) bilingualism (Trudgill, 2010, 2011). But gaining features in one language from another can also happen sometimes in adult late bilingualism, if the social circumstances are right (see Filipović, 2019, for details). Finally, there are universal laws that underlie the order of L2 acquisition and that contribute to shaping the way in which an L2 develops in individuals and undergoes change through language history (see Hawkins & Filipović, 2024, 2025). These laws account for anaphoric definite articles being transferred into article-less languages before those with generic reference, and for the failure of a voiced stop to transfer in the absence of a corresponding voiceless one. Projecting future language changes based on what we know about the effects of bilingualism on the changes documented already is a thrilling new prospect in this field, as is computational modelling of the multiple factors we have made explicit in this article, with a view to formulating precise predictions for their empirical effects, along the lines that was done for computational modelling of monolingual language evolution by Kirby (1999) using Hawkins' (1994) theory of efficiency.

Acknowledgements

We owe our gratitude to two anonymous reviewers who provided helpful insight and constructive criticism that resulted in numerous significant improvements of this paper. We would also like to thank the journal editor, Ad Backus, for his supportive handling of this submission. Any remaining errors rest solely with us as the authors.

ORCID iDs

Luna Filipović  <https://orcid.org/0000-0001-9460-9797>

John A. Hawkins  <https://orcid.org/0009-0005-3938-7462>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors received financial support for the research, authorship, and/or publication of this article beyond general University of California Davis faculty research funding.

Declaration of Conflicting Interests

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

Notes

1. We are thankful to one of our anonymous reviewers for inspiring us to bring more clarification to this point.
2. See Filipović (2019: 112) for a visualization of CASP predictions and for more discussion in this regard related to different grammatical and lexical domains.

References

- Athanasopoulos, P. (2011). Cognitive restructuring in bilingualism. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 29–65). Multilingual Matters.

- Austin, L., Hernandez, A. E., & Schweiter, J. (2019). *Proficiency predictors in sequential bilinguals: The proficiency puzzle*. Cambridge University Press.
- Barking, M., Mos, M., & Backus, A. (2023). Investigating language transfer from a usage-based perspective. *International Journal of Bilingualism*, 29(2), 415–438. <https://doi.org/10.1177/13670069231175629> (Original work published 2025)
- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, 3, 1–27.
- Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition*, 213, 104625.
- de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press.
- de Smet, H., Inglese, G., & Rosemeyer, M. (Eds.). (2025). *Acting on actuation*. Language Science 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). Multilingual Matters.
- Di Biase, B., & Kawaguchi, S. (2002). Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second Language Research*, 18(3), 274–302.
- Emonds, J. E. (1976). *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. Academic Press.
- Fenyvesi, A. (1994). *Language contact and language death: The case of Hungarian* [MA dissertation] University of Pittsburgh, Pittsburgh, PA.
- 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). John Benjamins.
- Filipović, L. (2017). Applied Language Typology: Applying typological insights in practice. *Languages in Contrast*, 17(2), 255–278.
- Filipović, L. (2019). *Bilingualism in action: Theory and practice*. Cambridge University Press.
- Filipović, L. (2025). Bilingual advantage for memory of events: Why now you see it, now you don't? *Brain and Language*, 268, 105615. <https://doi.org/10.1016/j.bandl.2025.105615>
- Filipović, L., & Hawkins, J. (2019). The Complex Adaptive System Principles model for bilingualism: Language interactions within and across bilingual minds. *International Journal of Bilingualism*, 23(6), 1223–1248.
- Filipović, L., & Hawkins, J. A. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics*, 51, 145–176.
- 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). Addison-Wesley.
- Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5), 389–407.
- Gracia-Tabuenca, Z., Barbeau, E. B., Kousaie, S., Chen, J.-K., Chai, X., & Klein, D. (2024). Enhanced efficiency in the bilingual brain through the inter-hemispheric cortico-cerebellar pathway in early second language acquisition. *Communications Biology*, 7, 1298.
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530.
- Grosjean, F. (1992). Another view of bilingualism. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 51–62). North Holland.
- Guo, Y., & Yuan, B. (2024). Disentangling cues of different domains in transfer and development in L3 acquisition: An investigation of L2/L3 Mandarin yes-no questions. *Bilingualism: Language and Cognition*, 27, 901–913.
- Hawkins, J. A. (1983). *Word order universals*. Academic Press.
- Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge University Press.

- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press.
- Hawkins, J. A. (2009). An efficiency theory of complexity and related phenomena. In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable* (pp. 252–268). Oxford University Press.
- Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. Oxford University Press.
- Hawkins, J. A., & Filipović, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge University Press.
- Hawkins, J. A., & Filipović, L. (2024). Bilingualism-induced language change: What can change, when and why? *Linguistic Vanguard*, 10(2), 115–124.
- Hawkins, J. A., & Filipović, L. (2025). Contact-induced language change and its constraints in relation to actuation. In H. De Smet, G. Inglese, & M. Rosemeyer (Eds.), *Acting on actuation* (pp. 15–39). Language Science Press.
- Hawkins, J. A., & Gell-Mann, M. (Eds.). (1992). *The evolution of human languages: Santa Fe Institute Studies in the sciences of complexity, Volume XI*. Addison-Wesley.
- Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge University Press.
- Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., Castilla-Earls, A. P., Sun, L., Weiss, S. D., & Young, D. R. (2019). Neuroemergentism: A framework for studying cognition and the brain. *Journal of Neurolinguistics*, 49, 214–223.
- Hernandez, A. E., Greene, M. R., Vaughn, K. A., Francis, D. J., & Grigorenko, E. L. (2015). Beyond the bilingual advantage: The potential role of genes and environment on the development of cognitive control. *Journal of Neurolinguistics*, 35, 109–119.
- Hijazo-Gascón, A. (2021). *Moving across languages: Motion events in Spanish as a second language*. De Gruyter.
- Huang, C. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15, 531–574.
- Jarvis, S., & Pavlenko, A. (Eds.). (2007). *Crosslinguistic Influence in language and cognition*. Routledge.
- Kawaguchi, S. (1999). The acquisition of syntax and nominal ellipsis in JSL discourse. *Proceedings of the Third Pacific Second Language Research Forum*, 1, 85–93.
- Kawaguchi, S. (2002). Grammatical development in learners of Japanese as a second language. In B. di Biase (Ed.), *Developing a second language* (pp. 17–28). Language Australia.
- Kellerman, E. (1983). Now you see it, now you don't. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112–134). Newbury House.
- Kirby, S. (1999). *Function, selection and innateness*. Oxford University Press.
- Koster, D., & Cadierno, T. (2019). The effect of language on recognition memory in first language and second language speakers: The case of placement events. *International Journal of Bilingualism*, 23(2), 651–669.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems in applied linguistics*. Oxford University Press.
- Levshina, N. (2023). *Communicative efficiency: Language structure and use*. Cambridge University Press.
- Lewandowski, W., & Özçalışkan, Ş. (2021). The specificity of event expression in first languages influences expression of object placements events in second language. *Studies in Second Language Acquisition*, 43(4), 838–869.
- Liceras, J. M., & Díaz, L. (1999). Topic-drop versus pro-drop: Null subjects and pronominal subjects in the Spanish L2 of Chinese, English, French, German, Japanese and Korean speakers. *Second Language Research*, 15, 1–40.
- Lubbers Quesada, M. (2015). *The L2 acquisition of Spanish subjects: Multiple perspectives*. De Gruyter.
- MacWhinney, B. (1992). Transfer and competition in second language learning. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371–390). Elsevier.
- MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.
- MacWhinney, B., & Bates, E. (Eds.). (1989). *The crosslinguistic study of sentence processing*. Cambridge University Press.
- MacWhinney, B., & Gao, Z. (2025). Six advances in research on bilingualism. *Brain & Language*, 268, 105614.
- McDonald, J., & Heilenman, K. (1991). Determinants of cue strength in adult first and second language speakers of French. *Applied Psycholinguistics*, 12(3), 313–348.

- Miyamoto, Y., & Yamada, K. (2019). On the interpretation of null arguments in L2 Japanese by L1 German speakers. *Journal of Monolingual and Bilingual Speech*, 1(2), 312–332.
- Montrul, S. (2004). *The acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and in adult L2 acquisition*. John Benjamins.
- Montrul, S. (2008). *Incomplete acquisition in bilingualism. Re-examining the age factor*. John Benjamins.
- Montrul, S. (2015). Dominance and proficiency in early and late bilingualism. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalisation* (pp. 15–35). Cambridge University Press.
- Montrul, S., & Rodríguez Louro, C. (2006). Beyond the syntax of the null subject parameter: A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In V. Torrens & L. Escobar (Eds.), *The acquisition of syntax in Romance languages* (pp. 401–418). John Benjamins.
- Muysken, P. (2013). Language contact outcomes as the result of bilingual optimisation strategies. *Bilingualism: Language and Cognition*, 16(4), 709–730.
- Myers-Scotton, C. (2006). *Multiple voices: An introduction to bilingualism*. Blackwell Publishers.
- Neeleman, A., & Szendrői, K. (2005). Proceedings of the 24th West Coast conference on formal linguistics (pp. 299–307). Cascadilla Proceedings Project.
- Newmeyer, F. (1998). *Language form and language function*. MIT Press.
- Nicoladis, E., Tewelde, D., & Zeschuk, V. (2024). Bilingual and monolingual adults' lexical choice in storytelling: Evidence for the weaker-links hypothesis? *The Mental Lexicon*, 18(3), 401–416.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Lawrence Erlbaum.
- O'Grady, W. (2008). The emergentist program. *Lingua*, 118, 447–464.
- Pavlenko, A. (2014). *The bilingual mind and what it tells us about language and thought*. Cambridge University Press.
- Phinney, M. (1987). The pro-drop parameter in second language acquisition. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 221–238). Reidel.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. John Benjamins.
- Pienemann, M., Di Biase, B., Kawaguchi, S., & Håkansson, G. (2005). Processing constraints on L1 transfer. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psychological approaches* (pp. 128–153). Oxford University Press.
- Pladevall Ballester, E. (2010). Child L2 development of syntactic and discourse properties of Spanish subjects. *Bilingualism: Language and Cognition*, 13, 185–216.
- Polinsky, M. (1995). Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics*, 14(1–2), 87–123.
- Polinsky, M. (2018). *Heritage languages and their speakers*. Cambridge University Press.
- Rappaport, G. (1990). Sytuacja językowa Amerykanów polskiego pochodzenia w Teksasie. In M. Władysław (Ed.), *Język polski w świecie: Zbiór studiów* (pp. 159–178). Państwowe Wydawnictwo Naukowe.
- Roebuck, R. F., Martínez-Arbelaiz, M. A., & Pérez-Silva, J. I. (1999). Null subjects, filled CPs and L2 acquisition. *Second Language Research*, 15(3), 251–282.
- 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 University Press.
- Ross, M. D. (2007). Calquing and metatypy. *Journal of Language Contact*, 1(1), 116–143.
- Rutherford, W. (1983). Language typology and language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 358–370). Newbury House.
- Savić, J. (1995). Structural convergence and language change: Evidence from Serbian/ English code-switching. *Language in Society*, 24, 475–492.
- Saville-Troike, M. (2009). *Introducing second language acquisition*. Cambridge University Press.
- Schmitt, E. (2000). Overt and covert codeswitching in immigrant children from Russia. *International Journal of Bilingualism*, 4, 9–28.

- Seliger, H. (1989). Semantic transfer constraints on the production of English passive by Hebrew-English bilinguals. In H. Dechert & M. Raupach (Eds.), *Transfer in language production* (pp. 21–33). Ablex.
- Song, J. J. (2011). *The Oxford handbook of linguistic typology*. Oxford University Press.
- Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, 119, 460–477.
- Stoco, A. (2019). Commentary on Neuroemergentism: A framework for studying cognition and the brain. The neurocomputations of neuroemergentism: Long-term memory + reinforcement learning =language? *Journal of Neurolinguistics*, 49, 248–251.
- Trevis, A. (1986). Is it transferable, topicalization? In K. Kellerman & M. Shanvood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 186–206). Pergamon.
- Trudgill, P. (2010). *Investigations in sociohistorical linguistics: Stories of colonisation and contact*. Cambridge University Press.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.
- Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics: A symposium* (pp. 95–189). University of Texas Press.
- Xu, L., & Yuan, B. (2022). Dependency resolutions of null and overt subjects in English speakers' L2 Chinese: Evidence for the cue-based model. *Second Language Research*, 40(2), 301–325.
- Yip, V., & Matthews, S. (2007). *The bilingual child: Early development and language contact*. Cambridge University Press.
- Zobl, H. (1980). The formal and developmental selectivity of L1 influence on L2 acquisition. *Language Learning*, 30, 43–57.

Author Biographies

Luna Filipović is an Affiliated Faculty at the University of California Davis. Her specialization is in bilingualism, forensic linguistics and experimental psycholinguistics. She discovered when, how and why language-specific effect on memory for witnessed events may or may not occur and she is the author of *Bilingualism in Action: Theory and Practice* (2019), published by Cambridge University Press.

John A. Hawkins is a Distinguished Professor of Linguistics at the University of California Davis and Emeritus Professor of English and Applied Linguistics at Cambridge University. His current research programs focus on language typology and grammatical universals from a processing perspective, language change and historical syntax, and bilingualism and second language learning. He originally proposed the empirically supported Theory of Efficiency in language processing, which accounts for typological variation patterns across languages, as presented in his seminal publications, *A Performance Theory of Order and Constituency* (1994) by Cambridge University Press, and *Efficiency and Complexity in Grammars* (2004) and *Cross-linguistic Variation and Efficiency* (2014), both by Oxford University Press.