A spanning approach to bilingual representations: Initial explorations

David Natvig¹, Erin Pretorius², Michael T. Putnam³,⁴, & Matthew T. Carlson³
University of Stavanger¹, University of the Western Cape², Penn State University³
& University of Greenwich (CREL)⁴

April 27, 2022

Abstract

Generative-oriented research has served as a useful tool in advancing our knowledge of aspects of bilingual acquisition, maintenance, and attrition across the lifespan. These analyses are, however, usually grounded in frameworks that focus on the (monolingual) ‘ideal speaker-hearer’. In this chapter we attempt to flip the script, adopting and elaborating on elements of current generative theory that are amenable to a system of bilingual primacy. We demonstrate that a neo-constructivist version of generative theory is capable of functioning as the basis of a bilingual system of grammatical representations.

1 Introduction

To date, there has been a wide array of formal models and various axiomata developed and employed in research on bilingual grammars. In most instances, linguists engaged in these endeavors select and apply theories, or particular elements of them, that were originally developed with the “ideal speaker-hearer” (Chomsky, 1965, 3) as the assumed norm. This body of theoretically-oriented work on bilingual populations has undoubtedly advanced our understanding of the linguistic competence and performance of these populations. Formal research on bilinguales, which originated in the domain of second language acquisition (SLA) (Slabakova, Leal, Dudley, & Stack, 2020; White, 2003), has now been extended into research on heritage language speakers (Polinsky & Scontras, 2020; Putnam, 2019, 2020a), studies on
language attrition (Cuza & Pérez-Tattam, 2016; Putnam, Perez-Cortes, & Sánchez, 2019; Putnam & Sánchez, 2013), and research involving multilinguals and L3-acquisition (Rothman, González Alonso, & Puig-Mayenco, 2019; Westergaard, Mitrofanova, Mykhaylyk, & Rodina, 2017). In spite of these acknowledged advances, bilingualism (and, by extension, multilingualism) presents specific challenges for theoretical explanation. Among these are:

- How bilinguals can possess knowledge of more than one grammar (Cook, 1991),
- How one grammar can impinge on another, not only in cases of so-called transfer of L1 structures to a later-learned L2, but also effects of second languages on the first (for a collection of relevant work see the contributions in Cook (2003)),
- Why bilinguals appear to treat potentially contrasting structures as equivalent, or partially equivalent, in their languages,
- Why L2 learners face substantial difficulties in learning certain structures, and
- How bilinguals regulate their language systems, both in monolingual communication and when they mix languages (Green & Abutalebi, 2013).

These questions give us reason to pause and reevaluate the validity and reliability of theoretical models that have in essence been directly imported from frameworks focusing on the ‘ideal’ monolingual. By starting with a theory of monolingual language, these questions often become conundrums that must be solved by adapting the theory to the ‘special’ case of bilingualism. We should, in contrast, approach human language from the opposite perspective, taking bi- or multilingualism as the norm. This, of course, does not make the aforementioned puzzles go away. But by treating the ability to maintain knowledge of multiple grammatical systems, the need to navigate multiple, potentially conflicting representations, the apparent invisibility of certain structures to some users, and so forth as basic properties of human language, these questions cease to be complications that must be fit into a theory, and we can seek a more encompassing theoretical account. Moreover, under this view, monolingualism becomes a subset of human language conditions, rather than an exceptional case to be fit in with theoretical tweaks.

This reversal is not a new suggestion, and in fact a literature has grown around the idea of multi-competence, first suggested by (Cook, 1991) in the context of a principles and
parameters framework, and now robustly represented through multiple approaches (see the contributions in Cook & Wei, 2016). However, the bulk of these approaches focus on psycholinguistics or usage-based approaches and, while these are valuable, we think that a more explicit formalism also has an important contribution to make. To this end, we explore recent generative approaches to language (following the lead of Amaral & Roeper, 2014; Roeper, 1999). Although we, too, draw on monolingual-oriented theorizing as a starting point (namely, neo-constructivist theories of grammar), we think these ideas offer a useful way of thinking about bilingualism as the basic instantiation of human language. We discuss how reinterpretating these approaches through a bilingual lens both opens the way to greater understanding of bilingualism and identifies specific areas for further theoretical development.

In this chapter, we revisit the ontology of generative, here meaning *formal* and *explicit*, research on bilinguals (broadly construed) by shifting the focus to the primacy of modeling the *bilingual* mind and language. Our goals here lie not only in developing a generative account of bilingual grammars (which has already been pursued for decades now), but in proposing a systematic way to deal with instances of representational conflict and other situations that are common to the bilingual experience across the lifespan. In our discussion of bilingual morphosyntactic patterns within the verbal domain, the following two questions guide our investigation:

**Q1:** Is it possible to model bilingual grammars in a systematic and constrained way?

**Q2:** What sorts of architectural changes/adjustments from those assumed for monolingual grammars are necessary in order to achieve this goal?

More concretely, as a starting point we adopt the theoretical perspective that the mind creates, or generates, mental representations that consist of hierarchical structure. These atomic elements of information, or **features**, are associated with (morpho-)phonological forms. For bi/multilinguals, features must be **reassembled** from their L1 (or default) setting to conform more to target L2 forms. In other instances, bilinguals face the challenge of integrating similar and dissimilar features into combined representations. Because languages do not make use of all possible contrasts (whatever those may be) and not all distinctions that are perceptible are encoded in the grammar (e.g. Dresher, 2014), bi/multilinguals are tasked with perceiving and acquiring potentially new distinctions and features, as well as differentiating, accessing, and implementing them in appropriate contexts. The integration
of multiple features lies at the heart of formal theories of mental representations, and is a particularly relevant issue for bi/multilinguals.

2 Theoretical Preliminaries

A valid model of bilingual grammar and its accompanying representations must encompass the broadest definition of what it means to be ‘bilingual.’ First, bilinguals are not “two monolinguals in one mind” (Grosjean, 1989), and are best understood as possessing a combined multi-competent, integrated architecture (Kroll & Gollan, 2014; Pickering & Garrod, 2013; Putnam, Carlson, & Reitter, 2018). Second, this integrated system leads to the dual activation of both, sometimes multiple, languages and potential for conflict between competing representations (Melinger, Branigan, & Pickering, 2014; Van Heuven, Schriefers, Dijkstra, & Hagoort, 2008) as well as shared ones (Gámez & Vasilyeva, 2020; Hartsuiker & Bernolet, 2017). The conflict within and between representations indicates probabilistic and gradient connections among them and the information (features) they contain (Bod, 2009; Chater, Clark, Goldsmith, & Perfors, 2015; Christiansen & Chater, 2016; Goldrick, Putnam, & Schwarz, 2016). Third, in line with these assumptions, successful language production and comprehension, as well as language attrition, is the result of having mastery (or the lack thereof), of inhibitory control (Green & Abutalebi, 2013; Green & Wei, 2014; Morales, Gómez-Ariza, & Bajo, 2016). Crucially, however, inhibitory control in isolation cannot explain or predict certain outcomes, rather conditions on representations and how they are formed also play an essential role (Gollan & Goldrick, 2016). Fourth, these factors taken together can affect the quality and composition of representations that bilinguals access, especially in studies focusing on online production and comprehension. Any attempt to model bilingual representations must minimally account for these conditions.

There are two particular developments in generative theorizing about bi- and multilingual grammars that have helped develop a divergent path from research (exclusively) on ‘ideal’ monolingual speakers and populations. The first of them concerns a committed movement away from parameters as having explanatory value to being regulated to second-order epiphenomena that emerge from more basic axioms and experience (Boeckx, 2014a; Eguren, Fernández-Soriano, & Mendikoetxea, 2016). Today, Universal Grammar (UG), i.e., the core set of linguistic knowledge not acquired through experience, is argued to consist of very little other than a structure building operation (call it Merge) that generates hierar-
chical (syntactic) structure (Boeckx, 2014b; Brennan & Hale, 2019; Getz, Ding, Newport, & Poeppel, 2018; Murphy, 2016) and some mechanism for acquiring formal representations (Biberauer, 2019; Cowper & Hall, 2014; Dresher, 2014). Research carried out by those who embrace the idea that language acquisition is guided by a limited number of domain-specific cognitive mechanisms that cannot be inferred from experience alone (i.e. UG), has also sought prediction based on some version of statistical learning and probabilistic parsing (Lidz & Gagliardi, 2015; Yang, 2016). In this vein, Lightfoot (2020) calls for an ‘open UG’ and espouses a version of the “learning to parse = acquiring a language” perspective. This view is common in more usage-based approaches, which postulate emergent categories and/or constructions (Tomasello, 2006). We adopt this version of an ‘open UG’ in our approach that follows.

An internally consistent and coherent account of mental representations must model how identified and internalized contrasts from ambient linguistic input, or Primary Linguistic Data (PLD), can be integrated in a system that generates longer connected spans (which is typically the byproduct of another acquired/emergent trait). These spans of information are hierarchical in nature, with information being shared within and between local units. The challenge for both simultaneous and sequential bilinguals is differentiating contrasts that are unique to each source grammar as well as recognizing ‘similarities’ and ‘differences’ between them. Making sense of these systemic attributes require a formal theoretical system (Jackendoff, 2017). In the remainder of this section, we establish the core desiderata of our system by outlining (i) the function of features, (ii) architecture of our model, and (iii) the concept of spans.

2.1 Features

Fundamentally, linguistic representational systems link FORMS to MEANINGS, and various traditions differ in how they conceptualize that architecture. Generative frameworks, either implicitly or explicitly, posit interfacing grammatical domains, or modules, with distinct representations and computations. Hall (2020, 248) conceptualizes these modular relationships as an arch of abstract systems that mediate between FORMS and MEANINGS (Figure 1), where representations toward the top are more abstract and discrete than those on the bottom. Syntax, the most abstract system, sits at the top of the arch and interfaces with phonology and semantics, each of which is respectively more abstract than a given (set of)
A central question for language acquisition concerns what the content of these various representations are and where they come from. Following a shift in perspective from parameters to features, and coinciding with a movement away from a ‘rich’ UG, feature content can be understood as being shaped by domain-specific mechanisms for creating abstract representations from linguistic input (Biberauer, 2019; Dresher, 2014; Hall, 2020). From this view, features are emergent results of the language acquisition process, specifically through the representation of grammatically relevant distinctions:

The ability to search for systematic contrast in the linguistic input, by correlating differences at various levels, is the only mechanism required to account for the abstract building blocks that make up those mental structures: the formal features of grammatical systems. (Cowper & Hall, 2014, 161)

Dresher (2014) maintains a similar view and argues that UG further requires that features be organized in a hierarchical relationship that reflects language-specific contrasts and patterns of activity. It is accordingly hypothesized that linguistic computations – whether they be (morpho-)phonological or (morpho-)syntactic – operate using these contrast-marking features (Dresher, 2009, 2014; Hall, 2020), i.e. the contrastivist hypothesis (Hall, 2007). Biberauer (2019) employs this perspective and proposes that hierarchical organizations of
emergent features account for different distributions of variable phenomena based on the scope of a feature over a given property. From this perspective, it is not only the features themselves, but also their relationships to each other in a particular system, that shape patterns of variation, both within and across languages (Biberauer, 2019; Dresher, 2014; Hall, 2020; Mackenzie, 2011, 2013; Natvig & Salmons, 2021).

Turning to what this means for a bilingual grammar, let us take an example from phonology. Bilinguals necessarily build representations based on PLD from multiple languages. A variety of language-external factors (e.g., age of acquisition, social contexts for acquisition and use over the lifespan, and many others) influences how these structures take form, resulting in heterogeneous representations of meaningful contrasts. For example, Natvig (2021) discusses situations where individuals build phonologies from source languages with different laryngeal (voicing, aspiration, etc.) contrasts. Spanish and English are clear examples, where, following a laryngeal realist framework (Honeybone, 2002, 2005; Iverson & Salmons, 1995; Natvig, 2021; Rainy, 2021), the Spanish phonemes /b d g/ are phonologically voiced ([slack] vocal folds) and voiceless /p t k/ lack laryngeal features (∅). English, on the other hand, contrasts aspirated, or [spread] glottis, /pʰ tʰ kʰ/ (spelled p t k) with plain, voiceless /p t k/ (spelled b d g). This distribution of features captures the language-specific differences in the phonological activity, based on either voicing or devoicing/aspiration assimilation, where the marked feature spreads to members of the unmarked series (Avery & Idsardi, 2001; Iverson & Salmons, 1995; Natvig, 2020). These two types of laryngeal systems – VOICING and ASPIRATING – are shown in (1), with (1a) demonstrating a Spanish-like system and (1b) and English-like one.

(1) Voicing and Aspirating laryngeal representations

<table>
<thead>
<tr>
<th></th>
<th>a. VOICING</th>
<th>b. ASPIRATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>/consonant/</td>
<td>{b d g p t k}</td>
<td>{p t k pʰ tʰ kʰ}</td>
</tr>
<tr>
<td>[slack]</td>
<td>∅</td>
<td>[spread]</td>
</tr>
<tr>
<td>{b d g}</td>
<td>{p t k}</td>
<td>{p t k}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>{pʰ tʰ kʰ}</td>
</tr>
</tbody>
</table>

Certainly, both systems are capable of representing a two-way distinction, such that a
Spanish speaker may speak English with a Voicing system (and vice-versa) in, for example, the early stages of L2 acquisition. However, as exposure and use of the L2 increases, so do the capabilities to map voiced, voiceless, and aspirated consonants to distinct categories (e.g., Casillas & Simonet, 2018; García-Sierra, Ramírez-Esparza, Silva-Pereyra, Siard, & Champlin, 2012; Gonzales & Lotto, 2013), consistent with the acquisition of both types of laryngeal systems. A bilingual Voicing-Aspirating system is shown in (2), where the /p t k/ series that lacks feature-marking is a shared category across both languages (Natvig, 2021). What this means is that bilinguals who acquire the requisite abstract features for both systems have multiple potential strategies for implementing those contrasts across languages so long as a distinction is made. For example, this could include variation between voiceless and aspirated consonants in Spanish or between the voiced and voiceless sets in English. Indeed, depending on patterns of use over time, access may be more automatic for one representational subset over the other, such that productions vary markedly from those of monolinguals (Natvig, 2021). In principle, however, an integrated system will contain both shared and distinct representations across all languages involved.

(2) Integrated representations for Voicing and Aspirating systems

The comparison of monolingual and bilingual representations in (1–2) demonstrates that acquirers can arrive at the same ‘goal’, i.e. representing contrast, but by taking different paths. One way that patterns in monolingual and bilingual language may differ is in feature composition, even if those features represent all the relevant language-specific distinctions. A typical monolingual Spanish-speaker will not have representations for phonologically aspirated consonants, and many Spanish-English bilinguals will, even if only the voiced and voiceless are used while speaking Spanish. Fully appreciating and operationalizing
potential variations in feature and representational content – even when bilinguals perform consistently with monolingual norms – advances the generative enterprise of modeling the architecture of language.

In the next section, we put in place some assumptions for a model morpho-syntax, where features are based on the same systematic representation of contrasts illustrated above. This model seeks to operationalize possible variations in speakers’ representational content in a formal and explicit way.

2.2 Architecture of the model

The aim of this section is to set out a formal model of the grammar that we believe reflects a (re-)orientation towards bilingual primacy. Generally speaking, the model falls in the is neo-constructivist family of approaches. This means that the grammatical properties of units of language are not predetermined (stored), but emerge from the structural environment in which they are embedded. The model assumes a maximally impoverished UG and emergent (natural classes of) formal features as described above. These formal features are merged in syntax into hierarchical structures called spans (see section 2.3 below), and the units of particular languages are mapped in an explicit and well defined way onto these spans. The components and organization of the grammar are compatible with the T-model of grammar (Chomsky, 1995), but adjusted to reflect the neo-constructivist view that lexical insertion takes place after syntax, and that syntax is fed by a formal feature inventory (Borer 2013; Halle and Marantz 1993; Starke 2009). This is illustrated in Figure 2.
The contents of (A) in Figure 2 come about through a combination of experience and a UG-supplied abstract template for acquiring formal features (Biberauer, 2019). When this template is activated in combination with general cognitive principles which alert acquirers to meaningful contrasts in the input, (A) becomes populated with formal features.\(^1\) The model assumes that bi-/multilingual individuals have an integrated feature inventory. In other words, formal features acquired through experience of a language/context X are housed in the same repository as features acquired through experience of a language/context Y. The integrated nature of the inventory means that a formal feature acquired through experience of a language/context X can (but need not) be extended, recycled or otherwise put to use in regulating the grammatical system of language/context Y.

In (B), Merge is the only UG-supplied mechanism and is responsible for organizing (a subset of) the formal features into hierarchies called spans (again, see §2.3 below). Given the myopic nature of syntax and the integrated formal feature inventory, there is no syntax-internal constraint on merging features originally acquired through experience of different languages/contexts. We also do not assume that the order in which features are merged is constrained by UG. In principle, features can be merged in any order, and the well-formedness of syntactic spans is constrained by interface conditions on the interpretability

\(^1\)For Biberauer (2019), also following the work of others, such cues to meaningful contrasts in (morpho-) syntax include (i) doubling, (ii) silence, (iii) multifunctionality, (iv) movement, and (v) category recursion (e.g. N-N compounds).
of structures (Ramchand & Svenonius, 2014). This yields a very limited number of combi-
natorial possibilities but, importantly, leaves room for variation.

The modular component of language in (C) houses two distinct types of information. The first are Units of Language (UoLs; following Wiltschko 2014). Our understanding of UoLs is that they represent listed associations between constellations of formal features (spans), (morpho-)phonological information, and semantic content. UoLs are mapped onto syntactic spans on the basis of a matching algorithm, detailed in §2.3 below. As with the formal feature inventory in (A), the model assumes that bi-/multilingual individuals have an integrated lexicon. For example, UoLs acquired through experience of a language/context X may compete for insertion into spans containing formal features originally acquired through experience of language/context Y, and vice versa. Furthermore, we expect (some) UoLs to be ‘mixed’ in their featural content, i.e., associated with formal features that (may have) originated through encountering input from different languages/contexts. The exact nature of the competition and ‘conflict’ between UoLs which are externally associated with different languages/contexts remains an open question for formal models of bi-/multilingual grammars.

The second type of information housed in (C) is a set of emergent category labels. A label can only be assigned to a UoL once the matching algorithm has mapped that UoL onto a span in the context of a particular derivation. In other words, it is the features comprising the syntactic span, and not the features encoded on the lexical items, that determine the grammatical properties of a UoL in context. Features encoded on lexical entries thus serve the sole purpose of determining the eligibility of a UoL to spell out a given span. This formalizes, for our approach in particular, the general neo-constructivist view that UoLs are a-categorial and categories can only be determined based on structural context. Emergent category labels (i.e. as syntax-eternal objects) is a topic that must be set aside in the present paper (but see Wiltschko (2014) for detailed discussion). As Figure 2 indicates, we assume that derivations proceed in a cyclic fashion, so components (A)-

\begin{footnote}
UoLs may be understood as roughly equivalent to *lexical items* in traditional minimalist frameworks, or to *vocabulary items* in Distributed Morphology.
\end{footnote}

\begin{footnote}
As Wiltschko (2014, 1) points out, categories (in the taxonomic sense) differ cross-linguistically since not all languages have the same categorial inventory (e.g. nouns, verbs, adjectives, adverbs, prepositions, and any number of functional categories). Furthermore, categories in descriptive traditions actually differ in kind: they may be words, morphemes, clause-types or constructions. This entire ‘heterogeneous set of categorizable entities’ (Wiltschko, 2014, 1) falls under our understanding of UoLs.
\end{footnote}
(C) may be accessed iteratively in the construction of a representation. Finally, although components (D) and (E) and their associated interfaces also fall outside the scope of this discussion, we assume that they operate under the same fundamental principles. For (E), see Rainy (2021) on the status of features and their roles in phonological representations.

Given our focus on components (A-C) in this paper, there are three places where variation may occur in the grammars of a bi-/multilingual speakers. Firstly, the contents of (A) may differ. Since formal features must be postulated by the acquirer, we do not expect different speakers necessarily to postulate identical features, with identical scope over relevant grammatical contrasts. Essentially, two superficially identical expressions may be underlyingly composed of different formal features. Secondly, variation might occur at (B) if, given an identical set of formal features, there is more than one merge-sequence that will yield an interface-interpretable structure. Third and finally, we also expect variation in terms of how UoLs map onto syntactic spans. In general, one and the same UoL may spell out more than one underlying structure. Importantly, we also expect that different speakers may have associated different constellations of formal features to a (morpho-)phonologically identical UoL. To make this fully explicit, the next section offers a discussion of spans and the matching algorithm that maps UoLs to structure in the post-syntax lexical insertion procedure.

### 2.3 Spans

As outlined in the previous section, spans mediate between abstract formal features and UoLs. The term span has been defined in various ways in recent literature. The definition in (3) is from Blix (2021, 7):

\[
\text{(3) Span:}
\]

An n-tuple of heads \( < X_n, ..., X_1 > \) is a span in a syntactic structure \( S \), iff \( X_{n-1}P \) is the complement of \( X_n \) in \( S \).

In prose, this means that any continuous set of heads in a single complementation line is a span. Since our representational system is one in which “syntax operates on individual features (and not on prefabricated bundles)” (Starke, 2018, 239), each formal feature is a head in syntactic structure. We assume furthermore that no head may be specified

\[4\]Cf. Svenonius (2020) and Blix (2021) for two examples.
with additional formal features, meaning that syntax adheres to a One Feature-One Head (OFOH) design (Kayne, 2005; Putnam, 2020b; Stroik & Putnam, 2013).

Once syntax has constructed a span through the successive merging of formal features (an S-span), that structure interfaces with the lexicon where it is matched to stored feature constellations (lexical L-spans) encoded on UoLs. A given UoL is a match for an S-span (i.e., it qualifies and competes with other UoLs to realize the S-span) when its L-span contains a superset of the features in the S-span. Thus, competition is governed by the Superset Principle:

(4) **Superset Principle:**

In case a syntactic span does not have an identical match in the lexical repertoire, select a UoL which contains a superset of the features present in the syntactic span. (adapted from Fábregas and Putnam 2020, 40)

In other words, the L-span of a UoL need not be a one-to-one match for the S-span in order to realize that S-span. What the set of competitors share is that their L-spans contain all of the features present in the S-span (they may additionally contain features not present in a particular S-span). The competitor whose L-span contains the fewest irrelevant features (in the sense that there is no match for them in the relevant S-span) is the ‘winner’. Thus, it may (but need not) be the case that the UoL which constitutes the best match for an S-span is an exact one-to-one match. Figure 3 offers an abstract illustration of the system.
Figure 3: Matching spans to UoLs

<table>
<thead>
<tr>
<th>Syntactic span</th>
<th>Lexical span</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Sa)</td>
<td>(La) UoL₁ ← F₁ F₂ F₃</td>
</tr>
<tr>
<td>(Sb)</td>
<td>(Lb) UoL₂ ← F₁ F₂ F₃</td>
</tr>
<tr>
<td>(Sc)</td>
<td>(Lc) UoL₃ ← F₁ F₂ F₃</td>
</tr>
<tr>
<td>(Sd)</td>
<td>(Ld) UoL₄ ← F₁ F₂ F₃</td>
</tr>
</tbody>
</table>

Suppose that (Sa-Sd) in Figure 3 are syntactic spans at the syntax-lexicon interface over which UoLs will compete for insertion. In the case of (Sa), the lexical spans encoded on the UoLs in (La) and (Ld) qualify and compete because they contain all the features comprising (Sa), i.e., F₁, F₂, and F₃. The UoLs in (Lb) and (Lc) do not qualify since they do not contain all the features comprising (Sa). Of (La) and (Ld), the winning UoL is (Ld); it contains fewer irrelevant features than (La) in this scenario. Thus, UoL₄ is the best match and is inserted to express (Sa).

In the case of (Sb), (Lb) is the only qualifying UoL since only its lexical span contains all the features comprising (Sb). Similarly, (La) is the only UoL that qualifies in the case of (Sc). Finally, (Lb) is again the only UoL that qualifies in the case of (Sd) since it alone contains a lexical span with all the features comprising (Sd).

It is worth pointing out that, in Figure 3, the features comprising the syntactic spans and the lexical spans are not obligated to ‘follow sequence’ (e.g. in (Sc) F₂ is ‘missing’; in (Lc) F₅ is ‘missing’). In a syntactic span, a feature may be ‘missing’ because a particular language does not incorporate that feature in its formal feature inventory (i.e. acquirers receive no
evidence in the ambient linguistic input for contrasts involving this feature). Another reason for ‘missing’ features in a syntactic span is that the composition of a particular expression simply does not involve a given feature. In the following section, we build on these general principles and illustrate our model for bilingual spanning. In particular, we analyze two case studies that demonstrate patterns of L2 acquisition in the verbal domain.

3 Representing bilingual grammars

This section demonstrates the potential of the model sketched out in section 2.2 to account for bilingual language phenomena from a perspective of normalised bilingualism. We review two case studies from the literature, each presenting a phenomenon which is (largely) unique to bilinguals, namely (i) shared syntax and cross-linguistic influence (§3.1), and (ii) missing categories (§3.2). As discussed in the introductory section, the level of detail attained in each particular analysis remains at present necessarily course-grained. Our aim is simply to highlight the affinity of the proposed model to account for these data. We hope that future research will continue to refine these first approximations.

3.1 Shared representations & Cross-linguistic influence (CLI)

It would be inaccurate to characterize CLI as mere transfer from the L1, wherein L2 words are fit into an otherwise L1 construction (although this of course happens, particularly when proficiency is low). Instead, many CLI phenomena are better characterized in terms of patterns arising from ‘pooled’ resources. In other words, information is shared between the language systems of a bi-/multilingual individual, and this gives rise to properties that are neither straightforwardly L1 or L2 properties. One such phenomenon is shared syntax, where contrasting structures originating in different language systems become, on the basis of sufficient similarity (Hartsuiker, Beerts, Loncke, Desmet, & Bernolet, 2016), connected in some way. This connection tends to strengthen with increasing proficiency.\(^5\)

\(^5\)There is an extensive literature on the concept of shared syntax in the psycholinguistic literature that supports the view that the representations in a bilingual mind that are ‘similar enough’ can correspond and directly compete other related representations (Hartsuiker et al., 2016; Hartsuiker & Bernolet, 2017; Hartsuiker, Pickering, & Veltkamp, 2004; Hwang, Shin, & Hartsuiker, 2018). Based on our understanding of these findings, they can be accounted for straightforwardly in the system we introduce here.
Despite evidence for shared syntax, however, our understanding of the nature of the shared information and how similarity is assessed by the language system is still lacking. The architecture we outlined in §2.2 offers a useful way to formalize this. Information can be shared (which also implies that it may vary between individual language systems) at different levels of the grammar. In the model sketched above, representations can be shared/may differ at the level of the formal feature inventory, at the level of syntax (i.e., the merging of S-spans), and at the level of the lexicon (i.e. the inventory of UoLs and the mapping of these elements onto S-spans). We also expect shared information/points of variation at the (interfaces with the) semantic and phonological components (see Natvig (2021) for a discussion of the latter).

To showcase the effectiveness of this model in understanding the phenomenon of shared syntax, the remainder of this section is dedicated to the findings of McManus (2015), who tested the effects of different L1 mapping patterns in German and English on the acquisition of aspect in L2 French learners. Simple past tense constructions are generated in English (5a), French (6a = imparfait), and German (7a = Preteritum). Additionally, each language also has a periphrastic construction involving an auxiliary and present participle (see the English example in (5b), French passé composé in (6b), and German Perfektum in (7b)).

   b. Jack has played football. [ENGLISH]

(6) a. Jacques jou-ait au foot
   Jacques play-IMP to.DEF.MASC football
   ‘Jacques used to play football.’
   b. Jacques a jou-é au foot
   Jacques has play-PC to.DEF.MASC football
   ‘Jacques played football.’ [FRENCH; McManus 2015, 162]

(7) a. Jack spiel-te Fußball
   Jack play-PRET football
   ‘Jack played/used to play football.’
   b. Jack hat Fußball ge-spiel-t
   Jack has football PERF-play-PERF
   ‘Jack played/used to play football.’ [GERMAN; McManus 2015, 162]
In spite of the availability of these two constructions in English, French and German, the three languages differ in how aspectual meanings map onto the structures. Whereas French has a 1:1 relationship between the simple/periphrastic form and habitual/perfective meanings, respectively, English and German do not. In German, aspectual distinctions are not encoded by these constructions (both forms can express either aspect, with Preteritum being generally used in literary and Perfektum in spoken contexts). In English, simple past can encode both habitual and perfective meaning, whereas the periphrastic construction can only encode perfective meaning.\(^6\) This information is summarized in Table 1.

\[
\begin{array}{|c|c|c|}
\hline
 & \text{Simple past} & \text{Periphrastic past} \\
\hline
\text{English} & \text{Habitual, Perfective} & \text{Perfective} \\
\text{French} & \text{Habitual} & \text{Perfective} \\
\text{German} & \text{NA} & \text{NA} \\
 & (\text{past, literary}) & (\text{past, spoken}) \\
\hline
\end{array}
\]

Table 1: Aspectual meaning x simple/periphrastic past in English, French & German

McManus (2015) compares four groups of L2 French learners, two with an English L1 background and two with a German L1 background. Of the two English L1 groups, one group was in year 1 and the other in year 4 of French instruction. The aim of the study was to test how the aspect mapping patterns of English and German might affect the acquisition of the French habitual/perfective aspect mappings onto the simple/periphrastic past constructions. The participants produced an oral narrative from a picture-based task that elicited the use of past aspectual forms.

The results demonstrate clear L1 effects on the production of target forms for both language background groups in all conditions (McManus, 2015, 167–172). English speakers were slightly more accurate in year 1, although their performance was highly variable in selecting the correct forms in both the perfective and habitual conditions. In contrast, year 1 German speakers overused the periphrastic form for both the habitual and perfective meanings. Both L1 groups in year 4 produced target forms at rates of 80% or above (McManus, 2015, 186–179), consistent with the 1:1 mapping of the French semantic associations.

\(^6\)English has the further possibility of expressing habitual meaning through other periphrastic constructions, such as with used to and would.
These results seem to reflect the fact that English grammatically encodes the distinction between habitual and perfective aspect (the English learners in year 1 were more accurate, and made use of both constructions), and also that the English simple past form is ambiguous between these aspectual meanings (the English learners in year 1 were highly variable in targeting the correct French forms). Furthermore, German does not employ simple/periphrastic past forms to encode the relevant aspectual distinction (the periphrastic form was highly preferred by the German learners, regardless of the target aspectual meaning). This bears out the prediction that the English speakers would initially outperform the German speakers because the former already have a grammatical ‘toolkit’ (features, S-spans, and UoLs with matching L-spans) in place for encoding these aspectual distinctions. For English speakers, the acquisition task involves acquiring new, distinct UoLs, isolating features, and associating those features to forms (i.e., building distinct L-spans for the newly acquired UoLs). By contrast, German speakers must constitute this toolkit ‘from scratch’, i.e., acquire representations for the aspectual contrast as well as the appropriate mapping processes (McManus, 2015, 164–165).

The remainder of this section aims to demonstrate that these results can be captured in an explanatory analysis using the model we have put forth in section 2.2. As a starting point, we will provide separate analyses of the French, English, and German systems, using an adapted version of Julien’s (2001) proposal for the structure of periphrastic tenses. In essence, periphrastic tenses are bi-clausal, both the main verb and the auxiliary are marked for tense, and every clause contains two temporal heads (a higher T(past) and a lower T(future)). Importantly, morphological markers on main verbs in periphrastic tense constructions are realisations of tense (not aspect), and the progressive, perfect, and prospective are tenses (not aspects). In particular, they are nonfinite present, past, and future, respectively (Julien, 2001, 132). So Julien’s system captures what is traditionally considered to be aspectual contrasts using only tense related features. The structures in (8) and (9) offer, using Julien’s system, representations of a simple past expression like He played and a periphrastic past expression like He has played, respectively.
A few comments about (8-9) are in order. First, perfective aspect is analyzed as non-finite past. This is reflected in the lower clause in (9), where [-FIN] embeds [+PST]. Second, present tense results from a combination of [-FUT] and [-PST]. Third, Julien (2001, 137) follows Roberts (1998) in assuming that be and have are “realizations of Vs that have no content except for formal features”, and thus the auxiliary has in (9) is the V of the higher clause, inflected for present tense (i.e., [-PST]; [-FUT]).

For our purposes, we adapt Julien’s (2001) analysis into a system with privative features and spans. Here we highlight the key differences. First, privative features do not come in ‘types’ (of which there may be subtypes), nor do such types/subtypes of features cluster on syntactic heads with which they associate (e.g., [PST] on T(past) and [FUT] on T(future)). Rather, privative features define OFOH systems (as described in §2.3 above), in which the features themselves are inherently ‘contentful’ and constitute the heads comprising the S-span. In our adaptations of structures like (8) and (9), features like [PST] are not ‘bundled’ into heads like T, but are themselves the syntactic heads constituting the structural spine. On an OFOH design, then, which is a fine-grained approach to syntactic structure, ‘T’ is better understood as a structural ‘zone’ than a syntactic head. Secondly, privative features do not have an attribute:value structure. Rather, the presence of a feature indicates a positive value and the absence of that feature is synonymous with a negative value. We represent a positive feature value by the presence of that feature in the S-span; a negative feature value is signalled by the absence of that feature in the S-span.

With these elements in place, we suggest that the French imparfait and the English simple past in expressions like (6a) and (5a) have the S-span in (10). Furthermore, the French passé composé and the English periphrastic in expressions like (6b) and (5b) have the S-span in (11). These structures have been substantially simplified, representing only the feature-heads comprising the clausal spines.

---

7For our purpose, we classify features as ‘contentful’ to the extent that they establish an index with a semantic reading/interpretation. See Fisher, Natvig, Pretorius, Putnam, and Schuhmann (2022) for a discussion of ‘purely’ morphological features in L-spans.

8As a point of clarification, ‘VP’ and ‘V’ should be understood as generalizations for event/verbal structures that contain additional detailed underlying structures.
In (10), the S-span denoting the mono-clausal spine of the French *imparfait* and the English simple past contains the features [FIN] (the clause is finite) and [PST] (the tense associated with the clause is past tense). In (11), the [FIN ['V'\textsuperscript{2}]] portion of the S-span corresponds to the higher of the two clauses in the bi-clausal spine of the French *passé composé* and the English periphrastic past. Note that ['V'\textsuperscript{2}] is the (set of) head(s) that correspond(s) to the ‘contentless’ verb that Julien (2001) and Roberts (1998) argue is realised as the auxiliary *have*. There are no tense features in this portion of the S-span because in Julien’s 2001 system [-PST, -FUT] indicates present tense. In the present system, where features with negative values are simply absent, [+FIN] merges directly with ‘V'\textsuperscript{2}, yielding the ‘default’ present tense. Still in (11), the [PST ['V'\textsuperscript{1}]] portion of the S-span corresponds to the lower clause. Note that [PST] is not embedded by any [FIN] feature. This yields a non-finite past, i.e., perfect aspect, which embeds the lexical (i.e., ‘contentful’) verb.

Let us now consider the L-spans, given in Figure 4, encoded on French and English UoLs that compete for insertion when S-spans like those in (10-11) are spelled out. \footnote{We assume that spell out is cyclic, but we leave the exact nature of how that interacts with our model, including the extent to which bi/multilingual grammars may define varying sizes of spell out domains, to additional work.}

---

\textsuperscript{9}We assume that spell out is cyclic, but we leave the exact nature of how that interacts with our model, including the extent to which bi/multilingual grammars may define varying sizes of spell out domains, to additional work.
When the S-span (10), denoting (finite) simple past, is spelled out in the grammar of an (ideal, monolingual) French speaker, only (a) of the competing UoLs in Figure 4 is a match. It alone contains the both of the features [fin] and [pst] required to lexicalize the *imparfait*. The set of nodes represented under ‘V’ are associated with √ and will be spelled out by √-containing UoL, here a lexical verb. For (6a), ‘V’ is spelled out by *jouer* and [FIN [PST]] as -ait or some other appropriate conjugated form. The same applies to lexicalizing this structure in English: only the UoL (d) in Figure 4, namely -ed, is selected. Accordingly, ‘V’+√ are lexicalized by the lexical verb, and [FIN [PST]] by the simple past ending.

When the S-span (11), denoting a bi-clausal, periphrastic tense in which the finite present embeds non-finite past (i.e. perfect), is spelled out in the grammar of an (ideal monolingual) French speaker, only the auxiliary (b) in Figure 4 matches the portion of the clausal spine associated with the higher clause, i.e. its L-span contains the sequence [FIN [‘V’]] (no √). Similarly, ‘V’+√ in the lower clause is lexicalized by a lexical verb, leaving only [PST] to be spelled out. The L-span of the *passé composé* verbal ending in (c) is a perfect match, resulting in a division of labor between the auxiliary *avoir* and the verbal ending é (or some other appropriate conjugated form) for giving expression to all the formal features in the S-span.

It is crucial to note that French has two separate UoLs that spell out the feature [PST]: the *imparfait* form, in conjunction with [FIN], yields finite (simple) past. The *passé composé*...
posed ending is, however, comparatively underspecified, with its L-span contains only [PST].
Recall that, on the model developed here, the L-span of a UoL must contain all of the features in the S-span in order to compete for insertion. So the passé composé verbal ending does not qualify as a candidate in expressions like (10). Although the imparfait ending is an insertion candidate for the [PST] feature in expressions like (11), the passé composé ending is optimal because it is an exact match, whereas the imparfait ending contains the superfluous feature [FIN]. Importantly, the imparfait ending cannot spell out [PST] and [FIN] in structures like (11) because these features are non-adjacent in this S-span; UoLs can only lexicalize features that are adjacent in the S-span. This produces the pattern observed in the French data, namely that there is a 1:1 mapping between verbal endings associated with perfect and imperfect meanings. English, in contrast to French, has no UoL whose L-span contains only [PST]. As such, English ‘recycles’ the simple past marker as the participle marker in the present perfect. This captures the pattern observed in the English data, namely that verbs marked with -ed are ambiguous between perfect and imperfect meanings. The task for English learners of French is to acquire the passé composé ending as an additional UoL, encoded as an L-span containing only [PST]. English learners of French also need to acquire the French morpho-phonological forms associated with the existing L-spans of the English UoLs. We may assume that the ‘combined’ French/English lexicon of an advanced English learner of French approximates Figure 5.

<table>
<thead>
<tr>
<th>French/English L-spans</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) <strong>Imparfait</strong> verbal ending / Simple past -ed ↔</td>
</tr>
</tbody>
</table>
| ![Diagram](image)
| (b) Auxiliary *avoir* / *have* ↔ |
| ![Diagram](image)
| (c) **Passé composé** verbal ending ↔ |
| ![Diagram](image) |

Figure 5: Combined French/English lexicon
We may hypothesize that in the early stages of acquiring the passé composé ending, English learners of French first acquire the morpho-phonological form and ‘piggyback’ the L-span of that UoL on the simple past -ed of English. In such a case, the L-spans of the imparfait and the passé composé endings are identical, are equally matched when it comes to spelling our the portion of the S-span underlying the verbal endings in both simple tense and periphrastic contexts. This is consistent with the high degree of variability of the English learners in achieving target forms. It is at this stage when there is a 50% probability of selecting the correct ending. As acquirers advance, the L-span encoded on the passé composé ending loses the \[\text{fin}\] feature, allowing the endings to become formally distinguished and thus viable forms for insertion in the correct structural contexts.

Let us now turn to German. Our basic assumption is that because German does not grammatically encode habitual/perfect meanings, but conveys viewpoints only implicitly, that German grammar does not create bi-clausal S-spans of the kind in (9) and (11) for the purpose of grammatically encoding perfective meaning. Thus, German only creates mono-clausal S-spans like (8) and (10). We view the perfective/habitual meaning that may come to be derived from that structure as a syntax-semantic (i.e., LF-interface) phenomenon. Importantly, though, periphrastic tenses occur frequently in German, so it would be incorrect to assume a 1:1 mapping between syntactic structure (the S-span) and form (the realization of the S-span by UoLs). Thus, we propose that although German, like French and English, has both a periphrastic and simple means of expressing tense, the structure underlying the German periphrastic differs from that of French and English in the sense that it is mono-clausal. The German L-spans are given in Figure 6.

In simple past expressions like (7a) above, the Preteritum ending (a) in Figure 6 expresses both \[\text{fin}\] and \[\text{pst}\] in the mono-clausal S-span (10), as shown in (12) below. In periphrastic expressions like (7b), we are still dealing with the S-span in (10), but it is spelled out by a different set of UoLs: the auxiliary hat (b) in Figure 6 expresses \[\text{fin}\], and the participle circumfix ge-...-t (c) expresses \[\text{pst}\], as shown in (13) below.

\[(12)\]
\[
\text{FIN} \quad \text{PST} \quad \sqrt{V}+\sqrt{V} \quad \text{spiel}
\]

\[(13)\]
\[
\text{FIN} \quad \text{hat} \quad \text{PST} \quad \sqrt{V}+\sqrt{V} \quad \text{spiel}
\]
Recall that, in Julien’s (2001) system, non-finite past is equivalent to perfect aspect. Thus, in line with traditional analyses, the participle circumfix `ge-...-t` in (13) is an expression of perfect aspect and not, as it may seem at first glance, past tense (see also, Musan 2002). It is in merging `[pst]` with `[fin]` that past tense (finite) emerges, as expressed by the Preteritum ending `-te` in (12).

The task facing German learners of French is thus not ‘simply’ to acquire new morpho-phonological forms and to encode them with the correct L-spans (as is the case for English learners of French). German learners must also acquire the ability to generate the bi-clausal S-span which is associated with the French (but not the German) periphrastic expression.

In this section, we illustrated the value of our model for capturing distinct trajectories of the acquisition of L2 aspect in a cohesive set of bilingual representations. We hope that future research will sharpen the proposed analyses, and propose new and exciting questions to test and hone our model. In the next section, we elaborate on how this perspective understands the concept of “missing categories”, and the implications that has for understanding L2 acquisition and grammatical representation.
3.2 Missing categories

Bilinguals may, at least on the surface, appear to lack a category in one language that is present in another, especially when compared to a monolingual ‘control’ group. This represents a legitimate confound in L2 acquisition. Mandarin Chinese is a ‘tenseless’ language in the sense that it does not overtly express past, present, or future with any special exponent overtly dedicated to tense-marking. Consider the examples in (14), which show that present (14a) and past tense (14b) are not distinguished morphologically from one another (data from Sybesma 2007).

(14) a. Zhāngsān zhù zài zhèr.
   Zhang San live at here
   ‘Zhang San lives here.’ [Mandarin: Present Tense]

   Zhang San 1989 year live at here
   ‘Zhang San lived here in 1989.’ [Mandarin: Past Tense]

Instead, Mandarin conveys temporal information through adverbials and discourse context (J. Lin, 2003, 2006; Smith & Erbaugh, 2005). Additionally, tense information is ‘piggy-backed’ on aspectual markers that often serve the dual function of also conveying temporal information (Li & Thompson, 1981). Four such morphemes are commonly discussed in the literature. Two of these, namely zhe and zài, characterize a particular situation as one that has not yet been completed (imperfective) (see (15)). The other two, le and guò, mark bounded situations (see (16)).

(15) a. Tā liú zhe yī tóu cháng fà
   he wear ASP one head long hair
   ‘He wears his hair long.’

   b. Lǐsì zài xǐzǎo
   Lisi PROG take-bath
   ‘Lisi is taking a bath.’ [J. Lin 2006, 13-14]

(16) a. Wò shuāidiàn-le tuǐ
   I break-ASP leg
   ‘I broke my leg (it’s still in a cast).’

   b. Lǐsì hē guò jiǔ
   Lisi drink ASP wine
   ‘Lisi drank wine.’ [J. Lin 2006, 13-14]
Although *le* is typically analysed as a perfective marker, (17) shows that it can combine with terminated (i.e. bounded, though incomplete) events as well as completed ones (as in (16a) above).

(17) Wò zúotiān xiě-le yì-fēng xīn, kěshì méi xiě-wán
    I yesterday write-ASP one-CL letter but not write-finished
    ‘I wrote a letter yesterday, but did not finish it.’ [Slabakova 2015, 287]

Furthermore, J. Lin (2003, 266) points out that although *le* usually marks complete or terminated situations, it is sometimes compatible with a present continuative interpretation. This is illustrated in (18).

(18) a. Ta yang-le yì-tiao jīnyú
    he raise-ASP one-CL goldfish
    ‘He is raising a goldfish.’

b. Wo (zài Boston) zu-le yì-jian gōngyù
    I in Boston rent-ASP one-CL apartment
    ‘I am renting/ have rented an apartment (in Boston).’ [J. Lin 2003, 266-267]

For J. Lin (2003), these facts suggest that *le* is not an absolute past tense marker. In contrast, situations marked by *guò* are always terminated past events, as in (19). Like *le*, *guò* does not solely mark past tense since it is compatible with future interpretations (see (20)). However, as J. Lin (2003) points out, it is a consistent marker of relative anteriority and therefore more like a past tense marker than *le*.

(19) a. Ta yang-guò yì-tiao jīnyú
    he raise-ASP one-CL goldfish
    ‘He raised a goldfish (before).’

b. Wo zài Boston zu-guò yì-jian gōngyù
    I in Boston rent-ASP one-CL apartment
    ‘I rented an apartment in Boston (before).’ [J. Lin 2003, 269]

(20) Deng nǐ ting-guò ta tān gǎngqín yīhou, nǐ jìu huì zhīdào ta de jiqiào nǐ yǒu duō hǎo
    After you have heard him play the piano, you will know his skill is how good
    ‘After you have heard him play the piano, you will know how good his skill is.’
    [J. Lin 2003, 269]
There is ongoing debate over whether Mandarin has a syntactically represented (although phonologically null) Tense projection, or whether the Tense projection is missing altogether (see i.a. J. Lin 2003, 2006; T. Lin 2015 and Grano 2017). Previous research on L1 Mandarin speakers acquiring L2 English highlights the difficulty for these speakers in acquiring English tense morphology. For example, Lardiere (2007) reports that her informant, Patti, with purported ‘fossilized grammar’, produced the expected -ed associated with the past tense of weak verbs only 34% of the time in obligatory contexts. This pattern sometimes persists even into advanced levels of L2 acquisition. For instance, Chan (2012) reports a lack of sensitivity to online English sentence processing of sample sentences lacking obligatory past tense morphology.

One hypothesis suggested by DeKeyser (2005) and others is that the English pattern is difficult for Mandarin speakers to acquire due to the lack of transparency between form and meaning pairings in Mandarin. This hypothesis adheres to the notion of a phonologically null Tense projection in Mandarin. An alternative hypothesis, put forward by Smith and Erbaugh (2005), is that learners rely on a ‘universal deictic pattern’, where “aspectual lexical class and view point aspect convey information that allows speakers to locate the situation in time” (Slabakova, 2015, 286). The deictic pattern determines that unbounded situations are located in the present and bounded situations are located in the past. This hypothesis adheres to the idea of a ‘missing’ Tense projection in Mandarin, on which the deictic pattern enables speakers to locate situations in time without establishing a new grammatical category for tense in the process of acquiring English.

Recall that a basic assumption for Julien (2001) is that viewpoint aspect in English (and, on our extension of that analysis, French and German as well) is simply non-finite tense. In other words, when tense features are embedded finite structure, the interpretation is tense-related whereas when tense features are embedded in non-finite structure, the interpretation is aspectual. Thus, only tense and finiteness related features comprise S-spans conveying viewpoint aspectual meanings in these languages. We argue that the same analysis should not apply wholesale to Mandarin.

Evidence from object fronting, wide scope of object-position quantifiers, and a scopal interaction between modals and aspect leads T. Lin (2015) to argue that Mandarin clauses exhibit finite/nonfinite contrasts. Since finiteness is a property of Tense, he concludes that Mandarin S-spans must contain tense-related features/projections. However, Grano (2017) argues that the contrasts identified by T. Lin (2015) can all be adequately
accounted for without appealing to Tense, and that Mandarin clauses may be said to “exhibit a finite/nonfinite distinction, but only if ‘finiteness’ is construed broadly as a cluster of properties that enable a clause to stand alone as a syntactically unembedded assertion” (Grano, 2017, 259). Because the model of grammar and the acquisition process put forth in this paper relies crucially on observable contrasts in the linguistic data, we follow Grano’s 2017 assertion and take the features comprising the S-span in Mandarin to be rooted in contrasts relating to viewpoint aspect, as opposed to tense. Following Slabakova (2015) and others, we consider tense interpretations in Mandarin to be a cognitive category arising from the interaction between grammatical aspectual and discourse-linked properties.

On Julien’s (2001) analysis, non-finite tenses (i.e. viewpoint aspect) must be realized as part of a bi-clausal structure. This is why expressions involving viewpoint aspect in English are periphrastic (i.e. bi-clausal). In contrast, since Mandarin viewpoint aspect is grounded in aspectual (i.e. not tense) features, we understand viewpoint aspect in Mandarin to be ‘basic’ and thus typically realised in the context of mono-clausal structures. In what follows we sketch out how an L2-learner grammar may arrive at these approximate features and their distributions. Consider the tree structure in (21):

(21)

\[
\begin{array}{c}
\text{FIN} \\
\text{‘V’}^2 \quad \emptyset \\
\text{zài} \\
\text{TERM} \quad \emptyset \\
\text{CONT} \quad \emptyset \\
le \quad \text{guò} \\
\end{array}
\]

The exponent zài is the only element that exists before ‘the verb’, which indicates that it is the only one of its kind that is a contentless verb. This is represented with ‘V’^2 in the decision tree. Of the remaining forms, both le and guò receive the feature [TERM] (inative), here used to indicate bounded events. This leaves both zài and zhe unspecified for TERM, rendering an imperfect interpretation. Finally in (21), we include the feature [CONT] (inuative), which for our purposes simply means “terminated, but not necessarily completed” Slabakova (2015, 288) and is consistent with a ‘past’ reading, and the ‘present continuative’ interpre-
tation when there is no [TERM] feature in the S-span. This is necessary in order for us to classify and capture approximations to actual aspectual features. The feature [CONT] also provides the additional function of distinguishing le from guò that also facilitates a somewhat wider distribution for le. Figure 7 offers a diagrammatic representation of the L-spans that we propose to be associated with the four aspectual markers in the grammar of an L1 Mandarin speaker.

<table>
<thead>
<tr>
<th>L-spans (L1 Mandarin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) zài ↔</td>
</tr>
</tbody>
</table>

Figure 7: L1 English and L1 Mandarin tense and aspect spans

In what follows we explicate the underlying structure that L1 English-L2 Mandarin learners with respect to their S-spans. Guò in (22) marks the bounded situation shown in (16b) Lǐsì hē guò jiǔ ‘Lisi drank wine’ because it spells out the TERM feature. Likewise, le spells out [TERM] in (23) in the sentence Wǒ shūàiduàn-le tuǐ from (16a) ‘I broke my leg (it’s still in a cast)’ in an addition to [CONT], which contributes to the understanding the leg at still broken. In contrast to (23), le in (24) spells out [CONT], but not [TERM], in Ta yang-le yì-tiao jīnguī ‘He is raising a goldfish’ from (18a) that provides a continual, but not completed (past), meaning. The representations presented in these trees (S-spans) capture the distinctions between le and guò in their bounded contexts, a le contains the additional cont feature. Furthermore, the larger L-span stored with le modes its a wider distribution, particularly in spelling out a non-bounded, i.e., non-[TERM], S-span, while still matching

---

10 J. Lin (2006, 12-13) proposes an alternative semantic account to address the ‘continuative’ interpretation of eventualities. This proposal assigns the assignment of result state aspectual semantics to all eventualities, even those that are inherently non-terminal (such as activities). We do not pursue a detailed discussion of the merits and challenges of this proposal here, but simply wish to point this out as an additional alternative to addressing the ‘continuative’ nature of Mandarin ‘tense’ for the interested reader.
to an S-span with a [CONT] feature in (24), corresponding to (18a).

In comparison to le and guò, neither zài nor zhe in our model spell out specific aspectual features generated in the S-span. The interpretation of these structures that are unspecified for aspect is that they relate to unbounded events. Again, the difference between the two in their underlying representations is that zài is a contentless verb and finiteness marker (25), whereas zhe only spells out finiteness (26). In Lísi zài xǐzào, ‘Lisi is taking a bath,’ zài both precedes the verb xǐzào and indicates a progressive aspect, like English through biclausal structure. However, for Tā liú zhe yī tóu cháng fā ‘He wears his hair long,’ there is no biclausal structure or progressive aspect. The unbounded meaning from the lack of [TERM] in the S-span reflects a present interpretation of the event. For all four markers, we argue that the ‘tenses’ they indicate are derived from the interactions of the aspectual features they spell out over larger syntactic structures, i.e., spans.
Assuming that our analysis is on the right track, we see that it possesses a number of challenges for Julien’s (2001) assumption that aspect = non-finite tense is too strong of a claim. In our view, the outcome that only tense and finite features can be projected in the syntax is too limited to achieve descriptive adequacy. With in the desiderata of the framework developed here, tense and finite are present in the syntax (i.e., S-spans) for English, whereas aspectual features are UoLs that comprise larger spans over multiple ‘VPs’. Of course, the features that are represented in the S-span and those that are interpreted over larger structures should be language-specific and result from the acquisition process described above. Along these lines, ‘tenses’ in Mandarin may be interpreted from how the larger collection of aspectual features in the syntax relate to one another. In many respects, this is the exact opposite scenario of what we propose. This, in our view, is both an expected and desired outcome of our model. Since we interpret features as the representations of meaningful contrasts, how these are organized and distributed within the syntax is based on language-specific properties. Because the primary contrast in English is tense, but aspect in Mandarin, we understand these to be the relevant features for each language’s core syntactic projections. How those features contribute to additional nuances in meaning in combinations over larger structures is furthermore dependent upon the acquisition and representation of language-specific phenomena.

Returning to bilingual phenomena, Slabakova (2015) analyzed the L2 acquisition of temporal distinctions in Mandarin by L1 speakers of English. Slabakova (2015) sought to test whether or not L1 English learners of Mandarin fared better in their acquisition of the Mandarin system. She found that English learners of Mandarin performed quite well on three tasks; (i) a Temporal Interpretation Choice task, (ii) a Stories task, and (iii) a Translation task, arguing that they make use of the universal deictic pattern. Furthermore, advanced- and intermediate-L2 learners also displayed categorical and variable behavior that roughly matched an L1 control group.

To understand what formal feature inventory and what lexical spans would already be available to L1 English-L2 learners of Mandarin, recall from Figure 4 (d) that the past tense verbal suffix /-d/ has the formal feature [pst]. Based on the universal deictic pattern, L1 Mandarin speakers tend to interpret le and guò as past tense and zài and zhe as present tense in the absence of any adverbials encoding temporal information. Our proposal is that the Mandarin S-spans shown in (25–26) are consistent with English biclausal and monoclausal present tense structures, i.e., (10–11) without the feature [pst]. Furthermore,
English L2 learners of Mandarin, instead of postulating a new formal feature ‘recycle’ the \[\text{PST}\] feature from their English inventory into the lexical spans of their L2 Mandarin, in precisely those places where L1 Mandarin speakers have postulated a \[\text{TERM}\] feature as in (27), leaving aside the distinction between the two. L1 English - L2 Mandarin Chinese acquirers possess a repertoire of UoLs that contain tense features and UoLs for aspect requiring multiple verb clauses. From these representations, they have prior experience creating lexical spans for the Mandarin particles even though their feature compositions differ from those in L1 Mandarin speakers. By drawing on their English representations, these speakers are able to generate (very close) approximations of the L2 target forms.

(27)

The structure in (27) opens the door to the possibly that L2 - a bi/multilingual grammars more generally - may exhibit different ‘levels’ of UoLs, according to which a feature projected into the syntax, or interpreted over a collection of features (i.e., spans), may be postulated. Admittedly, the exact nature of ‘levels’ of UoLs in bilingual grammars requires thought and attention far beyond what we can offer at the conclusion of this chapter, however, the mechanism for generating them would in principle remain the same as those we introduced here. We leave these exciting possibilities for future research.

The notion of Mandarin Chinese and other languages as being tenseless is therefore somewhat of a misnomer (see Ritter and Wiltschko 2009, 2014, for a related treatment). Although they may not have dedicated features and UoLs specifically for temporal distinctions, speakers nevertheless draw on the available feature combinations to express those contrasts. The English-speaking learners here show that L2 acquirers can then implement L1 tense categories as approximations to L2-target expressions and subsequently create new categorial distinctions.

4 Conclusion

Our primary objective in this paper was to take a critical look at recent trends in generative theory with the goal of modelling bilingual language and cognition as the assumed
default of human language. Here we return to the two research questions that anchored our investigation.

**Q1:** Is it possible to model instances of conflict resolution in bilingual grammars in a systematic and constrained way?

**Q2:** What sorts of architectural changes/adjustments are necessary in order to achieve this goal?

We hope to have demonstrated that we can answer Q1 in the affirmative. Moreover, the neo-constructivist model developed here and applied to a number of empirical contexts extends beyond current analyses that focus predominantly on constraining and guiding exponency selection, providing a tentative treatment of a wider array of phenomena unique to bilingual grammars. Our answer to Q2 is equally optimistic, but must be answered in two separate parts. First, our neo-constructivist model follows the larger trend towards an impoverished UG. Even though we hypothesize that the concept of *features* and the way they are structured may not be extrapolated from experience alone, their values and combinations are emergent. This approach provides support for Lightfoot’s (2020) recent appeal to an ‘open UG’. An additionally attractive facet of our model here is that it is capable of capturing the adaptive and dynamic nature of bilingual grammars through the lens of a modular architecture of grammar.

Second, we hope to have provided ample food for thought that the traditional strong divide between emergent and generative approaches is one of degree rather than category in many respects (Carlson, Fábregas, & Putnam, 2021). Certainly, some will interpret this as an improvement, others as a compromise, and perhaps for some as a gradual abandonment of the generative research program altogether. Perhaps we take undue solace in the assessment that we are following the invitation set out by Chomsky (1995, 225) to pursue a truly minimalist computational system:

> In pursuing a minimalist program, we want to make sure that we are not inadvertently sneaking in improper concepts, entities, relations, and conventions. The point of the occasional forays into formalism below is to clarify just how closely C_HL\(^{11}\) keeps to minimalist conditions, with principles and conventions derived where valid.

\(^{11}\)This stands for the Faculty of Human Language.
Our neo-constructivist model built on the foundation of bilingual primacy represents structural conflicts with minimal machinery that requires only “occasional forays into formalism”. The questions we raise here substantially outnumber the tentative answers we can provide at this time, however, this acknowledged caveat notwithstanding, we hope to have provided ample food for thought regarding how bilingualism can shape formal linguistic analysis, and in turn, how these formal analyses can contribute to research in bilingualism moving forward.

**Acknowledgments**

Many thanks to those who have provided us feedback on earlier versions of this analysis. We’re particularly grateful to the Synners group @ Penn State for critical conversations and suggestions. Lastly, a special thanks to Robert Klosinski for proofreading.
References


Christiansen, M. H., & Chater, N. (2016). *Creating language: integrating evolution, acqui-
sition, and processing. Cambridge, MA: MIT Press.
Goldrick, M., Putnam, M. T., & Schwarz, L. S. (2016). Co-activation in bilingual grammars:


trends in language acquisition within the generative perspective (pp. 1–23). Dordrecht: Springer.


tics, 49, 713–756.


