

Multiple Grammars and Second Language Representation

Second Language Research

2014, Vol. 30(1) 3–36

© The Author(s) 2013

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0267658313519017

slr.sagepub.com

**Luiz Amaral and Tom Roeper**

University of Massachusetts Amherst

Abstract

This paper presents an extension of the Multiple Grammars Theory (Roeper, 1999) to provide a formal mechanism that can serve as a generative-based alternative to current descriptive models of interlanguage. The theory extends historical work by Kroch and Taylor (1997), and has been taken into a computational direction by Yang (2003). The proposal is based on the idea that any human grammar readily accommodates sets of rules in sub-grammars that can seem (apparently) contradictory. We discuss the rationale behind this proposal and establish a dialogue with recent research in SLA, multilingualism, L3 acquisition, and L2 processing. We compare the Multiple Grammars explanation to optionality in L2 to other current proposals, and provide experimental results that can demonstrate the existence of active sub-grammars in the linguistic representation of L2 speakers.

Keywords

Multiple Grammars, Interlanguage, Universal Bilingualism, Second Language Acquisition, UG.

Introduction

Multiple Grammars (MG) is a theory of representation and acquisition that was originally proposed by Roeper (1999) to explain how idiosyncratic, incompatible rules could co-exist in adult monolingual grammars, and how they played a role in child first language acquisition. The extension of a model that was also called *Universal Bilingualism* to describe the interlanguage representation in adult second language learners, and bilinguals in general, seems to be an obvious next step, although the consequences of this extension are not necessarily trivial.

Different from traditional learnability theory, MG does not presuppose that new input requires the speaker to change a rule in the grammar in a way that either information is

Corresponding author:

Luiz Amaral, University of Massachusetts Amherst, 300 Massachusetts Avenue, Amherst, MA 01003, United States.

Email: amaral@spanport.umass.edu

Move2
step1Move1
step2Move3
step1

added to the rule or the whole rule is replaced¹. The core of the original proposal suggests that natural language grammars readily create parallel rule-sets, and the speaker has to decide which of the rules are generally productive and which have definably limited productivity, such as being limited to a lexical class, or a single idiosyncratic item. Roeper (1999) argues that any language contains properties of several recognizable language types, i.e. the grammar of a language L_1 can have elements that form sub-grammars compatible with L_2, L_3, L_n . An additional L2 challenge arises if a rule is productive in one grammar and lexically limited in another.

When linguists try to define the general characteristics of an L_x , they usually look for a convergent set of rules that defines the prominent properties of this language. These prominent properties are normally defined by the productive rules that can be used across the board with a wide variety of lexical items. For instance, when linguists say that English is an SVO, non-pro drop language, they are looking at the productive properties of the English language that can be generalized to a vast array of constructions, and they do not take into consideration the small set of the examples we present in the section *Multiple Grammars Theory* below. However, when observed closely, any grammar (G_x) also contains rules that allow for the existence of sub-regularities or idiosyncratic constructions that are (for the most part) lexically triggered. As we show in the sections *Minimalist rules and feature variation and Parameters*, the English language presents several examples of constructions that follow a V2 rule (like German) or a pro-drop one (like Spanish). Even if those examples are lexically limited, not entirely productive, and not very numerous, the grammar of a native English speaker has to allow for them to exist at the syntactic level. In the next section, we provide more details of how it is possible for these (in principle) contradictory properties to co-exist in the same grammar, and some of the consequences for a theory of acquisition based on parameter or feature value setting. It is important to notice that when we refer to ‘grammar’ in the term *Multiple Grammars*, we are actually talking about these subsets of rules (or sub-grammars) that co-exist in G_x .

We will orient our approach to previous proposals in L2 acquisition, and develop technical examples of how contradictory input is resolved under our Minimalist Principle: ‘avoid complex rules’. We then differentiate production grammars from comprehension grammars, and provide an experiment that demonstrates how the MG approach can explain optionality in L2.

This proposal follows the spirit of Kroch and Taylor (1997) and Yang (2003). Kroch and Taylor argue that as we see a mixture of grammars in the process of historical change, speakers must have had two representations in their minds and, given gradual shifts, they could calculate statistical rates of changing preferences. Yang, like Roeper (1999), argues as well that given contradictory evidence, a person can support and register evidence on both sides of a parameter. If the parameter is well-defined, then it is not difficult to tabulate how much evidence each side receives. We illustrate the details and complications for these perspectives with examples from V2 grammars below.

In sum, while Roeper (1999), Kroch and Taylor (1997), and Yang (2003) argue that the constant presence of incompatible sub-grammars in human language is responsible for dialectal variation in adult grammars, diachronic language change, and variation in L1 acquisition; we argue, in this paper, that it is also the primary source for *optionality* in all stages of adult L2 acquisition.

Multiple Grammars Theory

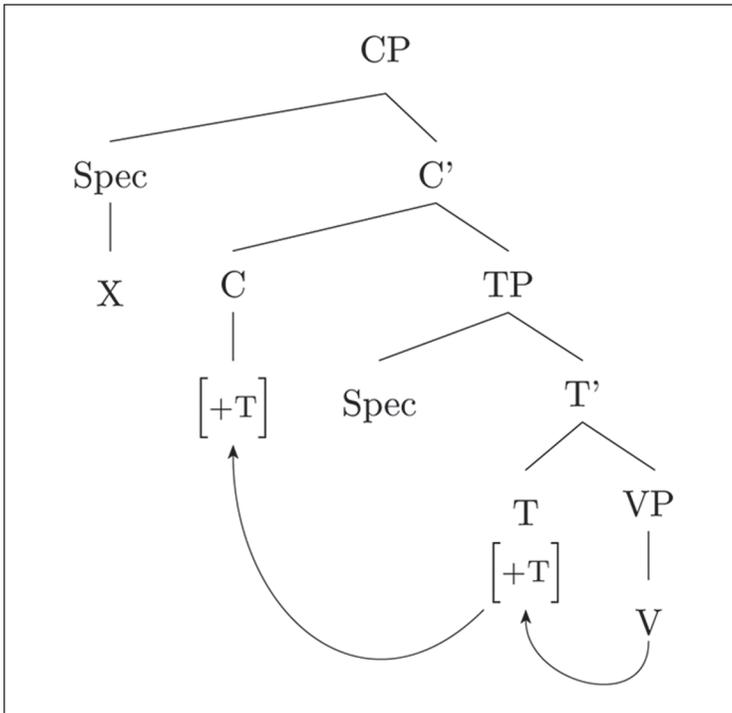
The argument that we are all bilingual (*universal bilingualism*) has intuitive, empirical, and technical dimensions. Several factors in modern linguistic theory make Multiple Grammars inevitable: (i) *Modularity* – a variety of modules with different primitives: thematic theory, binding theory, movement theory; (ii) *Minimalist Rules* – a vision of the statement of rules that requires minimal representations; (iii) *Lexicon* – a lexicon whose information is both idiosyncratic and carries partial representations of many modules; (iv) *Interfaces* – interface requirements which must obey their own restrictive principles. All of these factors will lead to the postulation of distinct sub-grammars within a given language.

It is valuable to bear in mind a non-idealized empirical dimension as well. It is a fact that the majority of children in the world hear more than one language and must assimilate, somehow, essentially incompatible information. Moreover, even monolingual societies have languages that (a) are in transition and so carry ingredients of different grammar types, and (b) contain lexical items that carry idiosyncratic information with the seeds of other grammars within them. In a word, the child is always faced with ‘contradictory information’ which must be resolved. We argue that it is never the case that a child (as an idealized model) builds a uniform set of rules that captures all of the information. In fact, the minimalist proposal makes this virtually impossible. For instance, the suggestion by Chomsky (1995) that there are no optional rules will immediately force a child to posit independent rules, rather than generate a single rule with an optional part to capture two related phenomena. If there are two rules, then one might ask how they can be both in the same language or in the same grammar. The theory of Multiple Grammars responds in part to an important formal requirement: *Avoid complex rules*. This is in the spirit of modern minimalism. It means that rules with subcategories and complex exceptions are difficult or impossible to formulate and therefore one’s grammar must reject them. They favor, we argue, access to two sub-grammars within a grammar. Ultimately, we expect to disallow other representational devices such as angled brackets and indices (as proposed by Reuland, 2011). We will illustrate this approach with four core sources of multiple sub-grammars that show that the notion of Multiple Grammars can be completely explicit, although many areas of the linguistic theory are themselves not explicit enough to carry out this promise.

Minimalist rules and feature variation

If rules are minimal in character, then they should be abstract and carry few exception features. German has a very simple rule of V2 that applies to the CP. It moves the verb to the second position after any constituent. It could hardly be simpler. It occurs after the verb has picked up a tense morpheme either in final or middle position, see illustration in (1).

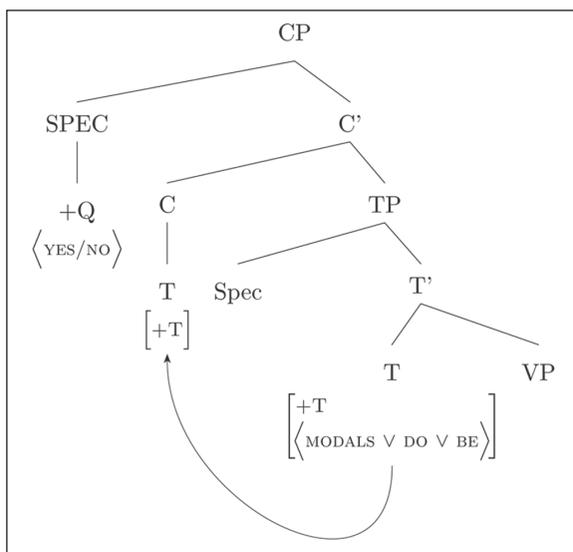
(1) German V2 Rule



Modern English does not move the main verb at all, so we never have a Shakespearean kind of question: “*plays he baseball?*” (except for idioms like: “*say you so?*”). Instead tense attaches to a modal or do-insertion occurs. Then, in what is called residual V2, it can move into CP after a question word: “*What did he say?*” This is also a very simple rule of movement of Tense, as we can see in the rule in (2).

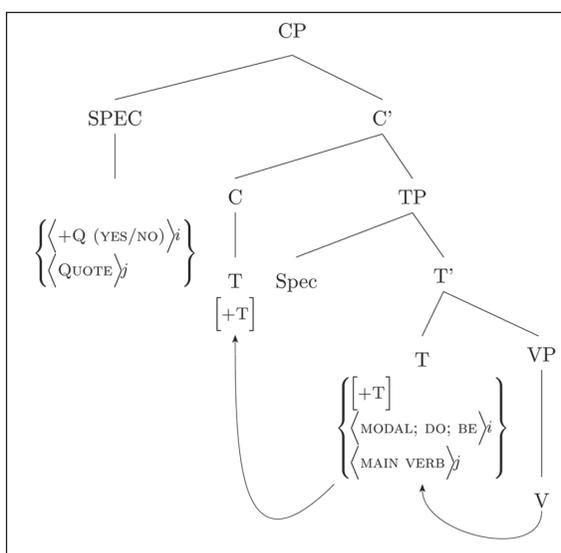
The English rule is limited to moving Auxiliaries and limits the initial XP to a Q-morpheme, but it has one important exception: it cannot occur with quotations, or else we would have: “**‘nothing’ did Bill say*”, instead of “*‘nothing’ said Bill*”. This means that even in modern English the main verb can move in this occasion, just like in German.

(2) English Auxiliary Rule



In order to capture this exception, we could provide a single complex rule in which either a verb or an auxiliary moves to the Tense position, and if it is a main verb with a *Quote-feature* or an auxiliary with a *Question-feature*, then it moves higher to the CP. Since they both involve movement to the C position, we can conceive of them within a single rule². If we decided for this approach, the rule would look like (3) (overlooking a number of details):

(3) English Complex Rule



In this rule, we now have a dependent choice: <+ Question (yes/no)> or <Quote>. This means that we have two kinds of dependency inside of each other in the statement of the rule. It is important to notice that historically English favors this formulation because it expresses the fact that the two rules are related. However, this approach runs against the spirit of minimalism and the notion of “*avoid complex rules*”. The MG approach keeps each of these rules separate and simple, but it allows a diacritic to attach to verbs of quotation making it a feature on a lexical item, linking to a simple rule of V2 languages, which has a much wider application in German. This is an alternative way to maintain the connection to historical evolution. It says that the earlier grammar continues to exist in a simple form, but it is lexically limited to a verb class, instead of being expressed with the categorial generality of V2. Now, in effect, we still have exactly the German rule that allows the *main verb* to move to the CP, but it is linked as a diacritic to a lexical class.

We believe that separating the two rules may provide several advantages from a descriptive perspective. Two simple rules facilitate the predictions about possible relations between L1 and L2, and allow us to more explicitly create testable hypotheses. For example, suppose an L1 English speaker learning German transfers only the “Quote rule”; this would demonstrate an instance of positive transfer. Then, we could hypothesize that the next step would be for the speaker to generalize this rule moving in the direction of the target German L2 rule. However, it is not clear to us how a model that presupposes only one complex rule that combines both the “Quote” and the “Auxiliary” parts, such as in (3), would explain the same phenomenon. We would have to assume partial rule transferring, which would create enormous descriptive problems, since we would have to test all possible hypotheses about places where rules could be divided to be partially transferred. In other words, a simple rule can be easily applied to any new sentence from either language. A complex rule, by contrast, does not apply so straightforwardly.

In reality, the historical situation is far more complex, and suggests the presence of lexical and frequency factors. Pintzuk (1999) notes that variation between a verb-final and verb-medial INFL position has some predictive value as to whether question formation involves movement of the main verb into C for Germanic V2. Thus, as long as there is a lexical link to particular verbs, this is what we predict. When both verb-final and verb-medial become categorical rules, then some competition must occur to allow one to dominate, perhaps in concord with other factors.

From this perspective, the concept of optionality begins to dissolve into more precise representations with formal predictions that are more easily testable. This does not seem to be the case if we were to present a descriptive observation that a single construction is optional at the rule level. We believe that many factors which are discussed under the notions of *transfer* and *optionality* can be captured with greater precision, and should be captured with precision in this manner (see sections *MG and optionality and Transfer*). This gives us a small inroad into the large empirical domain of *variation*, where these factors are present, but do not exhaust the domain. The presence of *social registers* is another manifestation of how the differences among grammars may be used in a socially differentiated manner. They are, still, primarily examples of multiple sub-grammars that can apply within a single language.

Meaning options

While any meaning can be expressed in any grammar, some features of grammars directly encode meaning, while in others it requires paraphrase. The strongest hypothesis is that every distinct grammatical form is ideal for the expression of some meaning.

To illustrate this point, let us imagine a Romance-L1 / English-L2 speaker whose language allows verb-raising, which separates the verb from the object, such as in (4). In English no raising occurs and one can produce a sentence like (5a), but not (5b).

- (4) Pierre lit rapidement le livre.
 Pierre reads quickly the book
 ‘Pierre quickly reads the book.’
- (5) a. Peter quickly reads the book.
 b. *Peter reads quickly the book.

However, as the examples in (6) show, adverb placement can alter readings. While (6a) can be interpreted positively, as if Japan is indeed gaining ground in the international arena, (6b) conveys an opposite message, where Japan is not catching up as fast as one might expect³.

- (6) a. Japan is slowly catching up with the West.
 b. Japan is catching up with the West slowly.

We can capture a broad distinction here by noting that there is a difference between modifying a verb by itself and modifying the verb+object. Thus one can easily imagine an English L2 speaker, who has a Romance L1, choosing the Romance grammar even while speaking English in order to capture a semantic difference that cannot be expressed within the English syntax⁴. Where (5a) can be used to mean that the event of “*Peter reading the book*” did not last very long, while (5b) allows *quickly* to modify the style of reading without including the object, meaning “*Peter gave the book a quick read*”. In other words, “*Peter read quickly the book*” could be paraphrased as “*the book did not seem to deserve careful study*”.

The use of another sub-grammar to capture a meaning option may in fact disregard the evidence that led to the rejection of that same sub-grammar, for instance, the role of inflection. To use French in English one must raise the verb to an Inflection position, although there is no inflection present motivating movement. Therefore we argue that meaning alone can be sufficient to motivate movement as a peripheral aspect of grammar⁵.

The lexicon

A third factor, the lexicon, quite overtly allows the representation of information from other grammars in limited ways. At an intuitive level, it is quite obvious that a language like English has both Anglo-Saxon and Latinate vocabularies and rules of productive morphology that are specified for language type. Thus, we add ‘-tion’ to *radiate* but not

to *push*: radiation, *pushion, which means that the words must be marked diacritically for their language origin. While many words have specialized meanings (transmission - 'a part of a motor'), alongside compositional meanings ('transmission of ideas'), only the compositional ones are also productively created by rules. Thus, we can say 'mis-comprehend' but not '*mispush'.

Anglo-Saxon words have their own set of productive affixes: -er, -ness, -ly. Interestingly, some of each type can cross-over and apply to the other system. We can say both *grammaticality* and *grammaticalness* (with slight differences in degree of abstraction), and we can say *misgivings* and *mistakes*, both of which involve Latinate affixes on Anglo-Saxon words. The first thing to notice is that both *misgivings* and *mistakes* actually block a compositional reading. They are permanently drifted (see Bauke and Roeper, 2011, for further discussion). It is not a misgiving if you *give something to the wrong person*, nor is it a mistake if you *take too much time*.

In those cases where the cross-over is productive, we have to change the representation to delete a class definition and leave only a category definition. We can substitute a broad category for one limited to V alone, noting that any V linked to V [other language] will not allow 're-' prefixation. Thus, if we allow ourselves to use a German word 'schlepp', it will not easily take an affix: '*mis-schlepp'. However, creativity is always a possibility and one might imagine someone inventing a new word or expression: *wish him 're-bon-voyage'*.

Another example can be found in Romance where the adjective comes after the noun, while in English they appear before the noun, as in (7). However, as we can see in (8), quantifiers in English appear to follow the Romance pattern.

- (7) a. Big politicians came to town.
 b. *Politicians big came to town.
- (8) a. *Big everybody came to town.
 b. Everybody big came to town.

Now we have a sub-regularity which we can represent in a lexical class of quantifiers, although it seems to have a deeper rationale that favors this syntax, namely the fact that quantifiers are less referential than adjectives⁶.

Each of these factors allows us to illustrate with some precision the form which Multiple Grammars can take. A number of semantic factors and interface factors may entail similar precise alternative formulations when those domains achieve formal precision. That is what is predicted by the MG theory.

Parameters

The exact status of parameters within the Generative Theory remains debated, and, here, we limit ourselves to a couple of comments about parameters and MG. Evidence suggests that parameters involve both specific triggers and 'clustering' of properties (see Holmberg, 2010). For the child, this means that, before all the impact of a cluster of properties is decided, she must keep track of multiple options. In early learnability

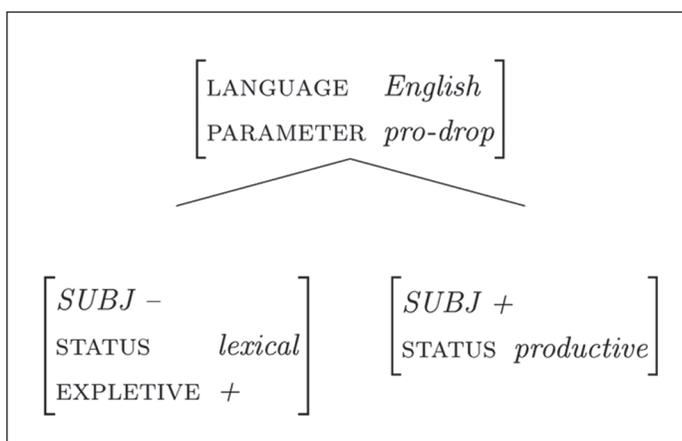
theory, the idea that the child projected multiple grammars and chose among them was created by the fact that there were in principle an infinite number of grammars. A simplicity metric choosing the shortest grammar was proposed by Yang (2003), who has suggested that frequency of each option for a parameter could help decide which was productive or dominant.

We can illustrate the idea in a simple and empirically valid way. The pro-drop parameter allows subjects to be deleted in many languages. A correlation with ϕ -features (number, gender, tense) exists in some instances, and a correlation with *there*-insertion exists in others. The fact remains that an English child hears the examples in (9). The presence of an expletive in (9a) and (9b) suggests that it is not pro-drop. The absence of the expletive in (9c), (9d) and (9e) suggests that it is. Therefore the child has evidence on both sides and starts both kinds of grammars.

- (9)
- a. It is good to bake bread.
 - b. There is a cat here.
 - c. Seems nice.
 - d. Looks good.
 - e. Raining, isn't it?

At first, the child would start out with both possibilities. Then, input would provide evidence that (i) in English only expletives are deleted, which occurs with a small number of verbs; and (ii) the majority of verbs can be added to the set that does not allow missing subjects. The second type of evidence would allow the child to substitute the group of verbs that require a subject by the whole category V. In sum, the child could see that while the two parameters are present, one is lexically limited, while the other is more productive; as illustrated in (10).

(10) Parameter Setting with Two Values



The figure in (10) allows us to illustrate the intuition of a parametric choice with two possible values established from the degree of productivity for each one of them in English. We will use this intuition to discuss more complex mechanisms to establish different combinations of feature values for different languages in bilingual/L2 speakers.

MG and L2 acquisition

To situate MG in L2 acquisition research tradition, we can state that it follows the steps of the Full Access/Full Transfer Hypothesis (Schwartz and Sprouse, 1996). Like any other natural language grammar, interlanguage is constrained by UG, i.e., the learner will not use properties, categories, or features that are inconsistent with the repertoire available to describe human languages. We follow Dekydtspotter et al. (1998) and White (2003a), who argue that *Full Restriction* (instead of Full Access) might be a better term to describe the role of UG in SLA. According to the version of MG presented here, the initial state of L2 acquisition is the stable⁷ state of L1, which means that all properties of L1 will be available to the L2 learner, and those properties will constitute the initial state of the L2 interlanguage.

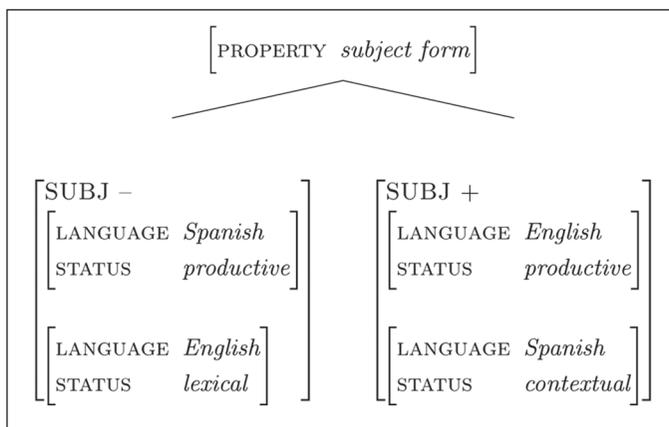
In the vast majority of formal models proposed to represent the nature of the interlanguage, researchers seem to take for granted the fact that the interlanguage is a system of representation that evolves by restructuring its properties in a very specific manner. White (2003b) presents a variety of graphical representations of L2 acquisition, in which the initial grammar changes into other grammars, as it restructures ‘away from the L1 grammar’ (White, 2003b, p.61). The linguistic description of what exactly is restructured changes according to the version of the theory used. In the 90s parameters were at the center of the debate. More recently features have taken a more prominent role due to their relevance in Minimalism (c.f., e.g. Lardiere, 2009). But in all cases the process of L2 acquisition seems to be based on the ability of the learner to eliminate or restructure the non-target properties of the interlanguage.

Differently from other generative theories of L2 acquisition, MG does not presuppose a resetting of parameters or restructuring/reanalysis of any linguistic feature or rule in an attempt to move away from the L1 towards the target L2. As we argued in the section *Multiple Grammars Theory*, any human grammar is made of multiple sub-grammars, and the interlanguage of L2 speakers will not be any different. There is no need to presuppose that one rule, feature bundling or representation will be replaced or deleted. Other features may be added to the interlanguage repertoire to play similar or distinct morphosyntactic roles to the ones that are previously existent, but the new ones will not replace the old ones. This idea is in sync with recent findings in L3 acquisition and L2 processing, as illustrated in the sections *Evidence from multilingualism and Evidence from processing*.

What is then the process of L2 acquisition, and how can L2 speakers attain native-like proficiency? We argued in the section *Multiple Grammars Theory* that the monolingual child receives contradictory input, starts developing multiple sub-grammars to handle the input, and is eventually able to use linguistic cues to figure out which of her sub-grammars are dominant and productive, and which ones are lexically motivated and idiosyncratic. The first language acquisition process is then defined by the child’s ability

to classify this (in principle) incompatible set of rules from the input, and develop mechanisms to appropriately access her multiple sub-grammars in different linguistic contexts. In L2 acquisition, this distinction between productive and idiosyncratic sub-grammars is much harder to achieve, exactly because they are language dependent. A language property that is productive in the L2 (say, pro-drop in Spanish), may be lexically based in the L1 (pro-drop in English). Any bilingual speaker (adult L2 learners included) will use a similar mechanism as the monolingual child, i.e., start multiple sub-grammars to accommodate the apparent contradictions in the input. However, the linguistic cues used to observe the productivity of each sub-grammar is going to be language specific, and consequently the bilingual speaker will have to undertake an extra step in the process: classify sub-grammars according to the language where the evidence came from. To illustrate this concept, let's compare the figure in (10) and the one in (11).

These two examples are obviously a (much) simplified version of all the features involved in the pro-drop phenomenon⁸, but they illustrate the basic idea behind our proposal. In (10) the language specification is done at the property level, because there is only one language involved (English). However, because of the lexically motivated exceptions in English, the monolingual speaker forms two independent sub-grammars that are classified according to their individual productivity. In (11), the language specification has to be instantiated at the sub-grammar level. The values *contextual* and *lexical* mean that those forms are restricted either by the sub-types of lexical items that accept them, or by the context in which they appear.



(11) Subject form for Spanish and English bilinguals

It is important to notice that the illustration in (11) reflects a simplified, idealized description of a sub-section of a bilingual grammar. In practice, a model that stipulates the existence of multiple sub-grammars allows us to test different configurations for the linguistic representation of bilinguals, without having to worry about encoding optionality at

the rule level (see section MG and optionality). Also, notice that in (11) the properties of English are similar to the ones in (10), which illustrates the idea that the interlanguage is not being restructured away from the L1, but simply building on top of it.

This specific characteristic of human grammars, namely the possibility of activating multiple parametric values and accommodating contradictory rules, has a direct consequence to the predictions on possible facilitative factors in adult L2 acquisition. In first language acquisition the *Language Acquisition Device* (LAD) supplies the child with an Initial State Grammar, which we can call a *Minimal Default Grammar* (MDG). This grammar then exhibits changes and modifications in response to input⁹. LAD is then responsible for certain stages in the acquisition, when we can see children producing non-adult like forms. The MG theory predicts that all sub-grammars that were activated during L1 acquisition will be available for life. That could explain why older children are not only able to understand baby-talk, but can also mimic their younger siblings, even if their grammars have evolved and are significantly different from the ones used by their brothers and sisters. These (previously activated) language properties could have a facilitative effect during L2 acquisition. For example, in the case of pro-drop, Liceras and Diaz (1999) show evidence that the Spanish L2 grammar of English L1 speakers contains null subjects in matrix and subordinate clauses at very early stages, while Galasso (2003) argues that Spanish speakers learning English have a tougher time imposing restrictions on null-subjects, and seem to follow a general “approach to L2 learning that is based on partial overt/covert language specific problem-solving procedures – lending credence to Transfer Hypotheses” (Galasso, 2002, p.1). Ever since Hyams (1983, 1986) presented the *parameter missetting idea*, we know that English kids go through a stage during acquisition that share some similarities with the L1 acquisition of pro-drop languages, such as Italian (Valian, 1990; Serratrice, 2005) and European Portuguese (Valian and Eisenberg, 1996). Despite the differences in the properties of constructions with pro-drop that surface during the acquisition of English and the acquisition of pro-drop languages (see Hyams, 2011), it seems clear that both English speaking kids and kids who speak pro-drop languages allow pro-drop constructions at a very early age. The MG prediction would be that this early development of (some sort of) pro-drop grammar could have a facilitative effect in L2. However, further studies would need to be carried out comparing the specific pro-drop constructions in L1 and L2 acquisition to allow us to understand the exact scope of this hypothesis.

When we consider specifically adult L2 acquisition, several of the sub-grammars to be used are already pre-defined (and pre-classified). L2 speakers may need to undertake two complementary actions in the overall process: (i) first, learners have to add new rules/features to their grammatical repertoire to account for language properties not contemplated in their existing grammatical knowledge, and (ii) learners have to reassess the productivity of the existing rules based on the input from the L2. The second part of the process may be much harder than the first one, since this assessment of productivity is language specific, and, as we saw in the section *Multiple Grammars Theory*, the input from any language is usually ambiguous due to lexically triggered exceptions.

The process of temporarily blocking a productive rule is completely different from eliminating or replacing that rule in the interlanguage representation. According to the MG theory, the learner never deletes a rule (or eliminates it in the process of

restructuring), because at some point s/he had evidence that the rule was needed, even if only to deal with very limited idiosyncratic lexical items, or to interpret input from another language. Non-productivity is not a reason to delete a rule from one's grammatical repertoire.

At this point, the theory seems to presuppose the existence of a unique repository for all sub-grammars a learner possesses, which would have rules that belong to more than one language. Though we believe this to be the case, it is not a necessary requirement for the MG theory to describe L2 representation. The alternative view would be to have two independent storage spaces, or completely separate grammars for the L1 and the L2, as suggested by the work with bilingual first language acquisition done by De Houwer (2005). However, if we take the L1 grammar to be the initial state of adult L2 acquisition, all L1 rules will already be present in the L2 grammar, and the same mechanism described above can be applied to describe the interlanguage. The reason for us to prefer a unique representation mechanism is because it allows us to explain different phenomena using a single theory of MG, such as instances of bidirectional transfer (e.g. Pavlenko and Jarvis, 2002), language attrition, and codeswitching Macswan (2005)¹⁰.

Finally, we would like to highlight that the example we used above in (11) illustrates a binary choice for explanatory simplicity. At this point, we are not committed to any view of how features are added or subtracted in developing grammars. It may happen through feature bundles, parametric decisions, or other properties of feature geometry that are currently not well understood. The ideas behind feature assembly in L2 acquisition presented by Lardiere (2009), and debated by Montrul and Yoon (2009), Liceras (2009) and Carroll (2009) deal with several controversial issues in current generative theory, and are beyond the scope of this paper. In principle, we do agree with the suggestion that parametric variations and feature reassembling are not necessarily incompatible ideas, such as proposed by Montrul and Yoon (2009), but we will leave the relationship between this discussion and the MG model for future work.

MG and optionality

There are two well-known approaches to optionality in the L2 literature that are worth mentioning in order to situate the MG proposal:

MOGUL. Truscott (2006) defines optionality in SLA as 'the simultaneous existence in a single speaker's grammar of two or more features, (where) each of which should normally exclude the other' (p.311). He provides the example of the French (L1) / English (L2) speaker who can produce both "*I quickly read the paper*" and "**I read quickly the paper*", in which the second sentence is generated with a French (non-target) rule. According to Truscott, optionality has been 'passed over' by generative researchers, and most of them 'have tended to relegate it to the often murky collection of performance phenomena, which competence theories are not expected to explain' (p.312). He, then, proposes to deal with optionality in SLA using an approach based on a framework called the 'Modular On-line Growth and Use of Language' – *MOGUL* (Truscott and Sharwood Smith, 2004; Sharwood Smith and Truscott, 2005; Truscott, 2006).

Before we describe the differences between the way MOGUL and MG deal with optionality, it is important to highlight that both approaches were motivated by a similar disagreement with the way grammatical development was depicted in traditional generative L2 literature. Sharwood Smith and Truscott (2005) review the data presented by Vainikka and Young-Scholten (1994) for word order, (the lack of) pro-drop constructions, and agreement markers in L2 German, and argue that a view of L2 language development where learners advance “through a series of discrete knowledge stages” (p.219) is not supported by empirical data. When describing the options of grammatical structures that arise from both the L1 and the L2, the authors rightfully point out that “(r)esearch has repeatedly found that development consists not of an abrupt switch from one option to another or from one option to free variation to the other option but rather of a gradual shift in the frequency of each option, both being used for extended periods, as if the grammar allowed optionality¹¹” (p.230).

Both MOGUL and MG approaches recognize that optionality is extremely pervasive in L2, to the point where it seems to define the grammar of second language speakers from the initial state to the ultimate/stable attainment of the target language¹². Second, both approaches agree that current alternatives to explain optionality in representational terms are too limited, such as the evidence discussed below that shows that interface phenomena can trigger optionality in L2 final attainment. Because optionality is a phenomenon that occurs throughout the acquisition process, we agree that it deserves a treatment that could explain it more thoroughly.

The MOGUL¹³ framework is based on the idea of a modular language faculty divided into two independent submodules, representing the phonological and syntactic structures, plus an external sub-module dealing with conceptual structures. These submodules are connected by interfaces, and there is one “process that operates on items stored in the (phonological, syntactic, and conceptual) lexicons” (Sharwood Smith and Truscott, 2005, p.234) – the ‘*activation level*’. According to Sharwood Smith and Truscott, many elements from L1 and L2 are activated during parsing, some are “ultimately used to achieve the final interpretation” (p.233) and some are just activated but do not survive. The *level of activation* of each item will determine its chances of making the final cut. “A particular item or configuration of items (for instance, N plus various syntactic features) may be more or less strongly activated. A strongly activated structure has a good fighting chance of being selected for association with structures elsewhere during the parsing.” (p.234)

According to (Truscott, 2006), optionality is then based on an existing imbalance in the activation levels between L1 and L2 items, which allows “the former to dominate use of the L2, especially in the early stages of L2 acquisition, when a relatively small amount of input has been processed and the appropriate value has not yet had a change to reach a high level. As more L2 processing goes on, this latter value should gradually rise to a point at which it can successfully compete” (p.321). The activation level of lexical properties uses a logarithmic model where ‘activation rises sharply in the early stages, after which the curve levels off to the point at which further input, no matter how extensive, has essentially no effect on the activation level’ (p.321).

The MOGUL framework presents a model for language development that explains the gradual changes in the frequency of use of language properties originating both from L1 and L2. It is interesting to notice that, without acknowledging explicitly, MOGUL

authors are taking for granted a model of representation compatible to what is proposed by the MG theory.

Differently from MOGUL, MG is not a theory of processing that aims at explaining the development of L2 grammars, nor in its current version it presents an explanation to the mechanisms that govern the choices regarding the degree productivity of each rule. The primary goal of MG is to describe how grammars (at any given point during the acquisition process) readily accommodate incompatible rules. It seems crucial to us to develop a descriptive model that can portray an L2 grammar at any given stage of development. A model of representation is a prerequisite to any theory that attempts to explain how linguistic knowledge evolves, or how humans use their grammars to parse sentences. Notions like processing can involve a lexical dimension, a structure-building dimension, or a memory dimension. Each of them refers to representations of some kind and the discussion will certainly advance if we know how to define human capacity beyond the representation¹⁴.

To MG, optionality is representationally possible because of the existence of multiple sub-grammars in any human grammar. The issue of how speakers decide on the productivity of each sub-grammar is an open empirical question to which activation (as proposed by MOGUL) can certainly be a possible answer. The version of MG presented here is not committed to such statistical models of input, and researchers with other inclinations may suggest that productivity be triggered by other causes, such as phatic factors along the lines of the research on motivation (i.e., Dornyei, 2009). One anecdotal example comes from a language teacher that would have to make his teenage English students use and repeat regular vocabulary items (such as *parts of the body*, *rooms of the house*, etc) many times before they could recognize and use them. However, if a student suddenly said in class a dirty word or slang and explained the meaning to his peers, the whole class would be using this word without any problems for weeks to come. An even funnier real life example happened to one of the authors, who always had problems remembering the gender of nouns in French. Once in a French Canadian restaurant he asked for cocaine instead of Coca-Cola, when instead of saying 'le Coke' (*masc*) he said 'la coke' (*fem*). After the waiter and his French speaking friends burst into laughter, he never again forgot the right gender for the non-alcoholic drink. This last story actually illustrates what is proposed by some interactionist approaches to language teaching, i.e., that feedback to errors made during attempts to communicate are much more effective than preemptively calling the learners attention to some language properties¹⁵ (Long, 2000; Doughty and Williams, 1998).

Another alternative to explain language activation could be related to the way certain linguistic properties are structured. For instance, word-finding difficulties have a great deal to do with how the lexicon is structured: can we find the appropriate paradigm to search? And where in the paradigm is the word? Thus, if we can find *blue* or *red* with ease, but *purple* is harder to locate, it could be a matter of pure frequency, or it could be the fact that we have two sets of colors, primary colors and more refined combinations. *Red* and *orange* might be same in frequency but one is in a subordinate lexical class while the other is not. An L2 speaker whose color system was not subdivided might say "sort of orange" where *red* would be correct, because there is no subordination. A native

speaker who says “sort of red or blue” instead of “purple” would make a more natural choice.

In sum, we do believe that MG as a generative representational theory is more suited for approaches such as the one proposed by MOGUL than other theories of representation that presuppose the restructuring of interlanguage properties away from the L1.

The Interface Hypothesis. The *Interface Hypothesis* – IH (Sorace and Filiaci, 2006; Sertratrice et al., 2006; Sorace, 2011) presents one of the most prominent proposals to deal with optionality in L2 acquisition. According to Sorace (2011), IH is ‘an attempt to account for patterns of non-convergence and residual optionality found at very advanced stages of adult second (L2) acquisition’ (p.1). The core idea behind this proposal is that structures ‘involving an interface between syntax and other cognitive domains are less likely to be acquired completely than structures that do not involve this interface’ (p.1). One interesting expansion of the original proposal establishes a connection between the difficulty of advanced L2 learners in acquiring structures that involve interface phenomena, and attrition in L1 bilingual acquisition. The use of the expression ‘not acquired completely’ by these authors suggest a connection to other approaches that contemplate the issue of ‘incomplete acquisition’ (Montrul, 2008).

Some of the evidence for the Interface Hypothesis comes from a sequence of experiments targeting the interpretation of null and overt subjects in Italian by native speakers of English¹⁶. The study done by Sorace and Filiaci (2006) asked advanced speakers of Italian L2 to choose a picture that could describe sentences similar to the one in (12). The researchers, then, checked the preferred interpretation for the null and the overt pronouns in the embedded clauses¹⁷. The results showed that while null subjects were interpreted in the same way by L1 and L2 speakers, the difference in interpretation of the overt pronoun by L2 speakers in relation to Italian L1 speakers was statistically significant. L2 speakers chose the subject in the matrix clause as the co-referent of the overt embedded pronoun more than three times as much as L1 speakers (26% vs 7.6%).

- (12) La mamma_i dà un bacio alla figlia_k mentre lei_{k/i/pro_i}
 the mother gives a kiss to the daughter, while she/*pro*
 si mette il cappotto.
 puts on the coat.

‘The mother kisses her daughter, while she/*pro* puts on her coat. (p.352)’

This connection between interfaces and optionality in L1 and L2 bilinguals suggest that whenever there is more than one language at stake, there are tangible consequences to the representation of the grammar of speakers. There are several critiques to the IH in the literature (cf., e.g., Montrul, 2011; White, 2009), and Sorace (2011) provides some interesting responses to some of the previous criticism.

From a MG perspective, the greater shortcoming of the IH is that it does not present a solution to the problem of which mechanism allows for the *representation* of the contradictory rules necessary to license the observed phenomenon. In order to explain how L1-English / L2-Italian speakers use the target rule to interpret embedded subjects when

they are null, and non-target rules to interpret them when they are overt, there has to be some way of encoding this contradictory information in the speaker's grammar. Sorace and Filiaci (2006) point out to two possible explanations to the problem. The first one involves some sort of 'un(der)specification at the level of representation' (p. 340). If this is the case, speakers should not be able to choose a possible interpretation to the null subjects in embedded clauses presented in the study, since the necessary rules to provide either the English or the Italian interpretation would be 'un(der)specified'. In fact, Sorace and Filiaci recognize this problem and show their preference to a second explanation, which is based on 'insufficient processing resources to integrate the multiple types of information involved at the interfaces between syntax and other cognitive domains' (p.340).

The issue of whether or not there are insufficient processing resources that prevent the speaker from assigning the appropriate co-referent in the case of overt embedded pronouns (but not in covert ones) is an empirical question that deserves future online tests, and we do not wish to dispute any conclusions based on such argument. We just want to point out that this explanation as stated does not address the description of the formal properties of interlanguage from a representational perspective. There has to be an underlying grammar that allows for this interpretation, even if it is so complicated to the point where L2 speakers cannot use it due to processing limitations. Besides, in Sorace and Filiaci's experiment L2 speakers do behave like native speakers 60% of the time, when selecting the object of the matrix clause as the referent for the overt pronoun in the embedded clause. We believe it is important to present a description of the underlying grammar being used in both cases.

Our core argument is based on the fact that processing accounts are usually inexplicit in their relation to representations, but since representations must be involved, it may well be the representation which holds the explanation. After all, a sentence like (13) is so short that it is hard to imagine a memory limitation or a heavy cognitive load as possible explanations for a preference.

- (13) The mother saw the girl while \emptyset running.

It is interesting to see that two of the most prominent approaches to optionality in the SLA literature (Truscott, 2006; Sorace, 2011) focus heavily on processing accounts. If, on the other hand, we try to explain optionality from a representational perspective, we have primarily two options. First, we can encode optionality into the grammar rules. For example, if a bilingual speaker of L1-Karaja (an SOV native Brazilian language) and L2-Brazilian Portuguese (an SVO language) is using both SVO and SOV word orders (say, for different communicative purposes, as illustrated in the section *Meaning options*), we could presuppose that the speaker has the representation of the rule in (14). This solution has obviously some unintended consequences; the most straightforward one being that this rule would allow a speaker to produce null object sentences and/or sentences with double objects. Since not all SVO and SOV languages have both these properties (null objects and double objects), it is perfectly conceivable that a speaker can learn the underlying word order of a language without learning to use null or double objects, which would make such approach to be descriptively inappropriate.

(14) S → Subj (Obj) V (Obj)

The second solution to the representational problem posed by optionality is the one presented in this paper: *Multiple (sub-)Grammars*. For example, if we assume, as argued with example (11), that in the case above both choices can be made at the same time, the same bilingual speaker would end up with both representations available in their interlanguage grammar, i.e. instead of having a rule like in (14), the speaker ends up with both rules in (15). This will obviously have direct consequences to the speaker's performance, as noted in the section *Multiple Grammars Theory and illustrated in section Evidence of MGs in the interlanguage*.

(15) S → Subj Obj V
S → Subj V Obj

Finally, we would like to make one final comment on the issue of optionality and MG in SLA. According to Sorace (2003), L2 optionality is different from L1 optionality in at least three respects: '(i) L2 learners have the L1 as an additional source of optionality; (ii) L2 optionality tends to persist at advanced competence levels; (iii) Residual optionality is found at ultimate L2 attainment' (p.139). We fully agree with Sorace on (i), there are indeed more sources for optionality on L2 grammar, since all properties of the L1 are present. However, we would like to point out, once again, that optionality is not a phenomenon exclusive to the final attainment of L2 speakers. As we saw in the section *Multiple Grammars Theory*, there is a great deal of optionality going on in adult monolingual grammars, triggered by idiosyncratic lexical specifications, sociolinguistic variation, etc. The main difference is that the optionality we encounter in L2 can lead to instances in production and/or comprehension that are considered to be 'non-target like' by monolingual standards. In sum, for MG, optionality is a byproduct of one of the core properties of the grammars of any human language (whether L1, L2 or L_x), and a formal representational approach that addresses this ubiquitousness of optionality can more elegantly help explain a more comprehensive range of phenomena.

Transfer

Odlin (1989) defines transfer as 'the influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired' (p.27). Whatever definition for 'transfer' one uses, this is a complex phenomenon with ramifications far beyond the scope of this paper¹⁸. In this brief section, we just want to highlight how a theory of language representation that allows for multiple sub-grammars to coexist in a single grammar can contribute to a formal approach to transfer. It is important to notice that we are referring to transfer as a rule-governed phenomenon, since our overall interest is on *systematic variation* (Romaine, 2003, p.410).

So, how can MG contribute to the way we formally view morphosyntactic transfer? If we accept the *Full Transfer/Full Access* hypothesis (by Schwartz and Sprouse, 1996) as

the initial state of L2 acquisition, it is not difficult to explain transfer at the very early stages, or even at intermediate levels of L2 development. However, the more an L2 speaker advances his/her mastery of the target structures, the trickier it gets for a theory based on parameter or feature resetting to explain why transfer may persist at the morphosyntactic level. If parameters and or features are completely reset in the interlanguage, transfer has to be caused by the interaction of two independent *grammatical representations*, i.e., two full-fledged independent grammars that exist in a speaker's mind. By adopting this representational model, the researcher would then need to present a formal mechanism for these two grammars to interact.

In the MG approach there is no need to presuppose the transfer of properties between two independent grammars, since all grammatical properties would permanently exist in the interlanguage. One advantage we see with this approach is the fact that, from a representational perspective, transfer is not caused by a mechanism that is unique to L2 speakers. Representationally, transfer could be caused by the speaker's uncertainty about the degree of productivity of each sub-grammar in his/her interlanguage, i.e. something similar to having both sides of a parameter activated and marked for “+ *productive*” (see figure in example (10)).

As we will see in the section *Evidence from processing* there is a high degree of interconnection between the grammars of bilinguals, and a growing number of researchers are encountering evidence on parallel activation of a bilingual's two languages (see Kroll et al., 2012, for details). That would suggest that one way of viewing transfer is through a speaker's inability to temporarily block the use of a productive rule from L_X . Notice that we are not making any claims about when each sub-grammar is used or activated by bilingual/L2 speakers in order to allow for transfer. We are simply formulating a theory of representation that allows for the multiple language properties that generate these variations to co-exist in the same grammar.

When it comes to the issue of optionality and transfer, we have to remember that not all types of optionality are caused by transfer, and they can actually be viewed as separate phenomena; but, as far as the MG theory is concerned, any type of morphosyntactic variation is motivated by the (natural) existence of multiple sub-grammars in the speaker's mind.

Evidence of MGs in the interlanguage

Sorace (2003) argues that a theory of optionality as presented by Multiple Grammars (Roeper, 1999) is impossible to be falsified because, according to her, ‘what evidence would unambiguously indicate that the speaker is using different grammars?’ To understand the nature of what we claim to be unambiguous evidence in support of the MG theory, we have to remember that we are assuming that the speaker's competence can be represented by a descriptive mechanism called *grammar*, and that for a given construction to be licensed, there has to be a corresponding rule in the speaker's grammar responsible for doing so. Without going too much into details on the differences between competence and performance, we want to emphasize that we do not believe that a given structure can be licensed by a speaker in a consistent way if there are no rules in the speaker's grammar that can generate it. By consistent, we mean frequently and exactly

with the same properties¹⁹. So, in order for a speaker (or a group of speakers) to consistently generate two representations for a given construction that are incompatible with one another, our model of grammar has to be able to accommodate the existence of two distinct (incompatible) rules²⁰.

Remember that when we say “*Multiple Grammars*” we are actually referring to the existence of a set of sub-grammars that may contain conflicting rules with one another. In L1 acquisition, speakers learn that these sub-grammars can be used in different ways. While some are lexically triggered, and can only be used with a limited number of lexical items in certain syntactic contexts, others are productive, and their use can be generalizable across the language. We also said that in L2 and bilingual acquisition this distinction between productive vs. idiosyncratic is harder to establish, because among other things, speakers may receive contradictory input regarding a certain rule (R_x) from the different languages, and be tempted to generalize the use of R_x across the board for both languages. Or, in the case of L2, speakers begin their acquisition process generalizing R_x from their L1, and later (hopefully) learn that R_x in L2 can only be used in very limited contexts (or not used at all). One way to present evidence for this type of language representation is to find productivity in rules that are not supposed to be productive, but rather lexically based.

We believe that even more compelling evidence for this type of multiple coexisting sub-grammars comes from situations where learners show strong preferences for two distinct conflicting rules in production and comprehension to represent similar semantic propositions. As we mentioned above, the reason we believe this dichotomous choice cannot be limited to a performance or processing realization is because we assume here a theory that explains the linguistic competence of speakers: the *grammar*. This grammar underlines all forms of language use, both in production and comprehension. This means that if a linguistic property is consistently found in production and/or in comprehension, we can assume that this property belongs to G_x , even if we find a contradictory rule in the same grammar. We turn now to an experiment that shows clearly that production and comprehension may access different grammars.

Conservatism in production vs comprehension

Many people have had the intuition that the use of certain rules in the grammar differs from production to comprehension (c.f. e.g., van Hout et al., 2010; Hendriks and Koster, 2010). Snyder (2007) has assembled a number of arguments to show that speakers command a great deal of lexical control not only over the choice of words and in particular verbs, but over their natural environments and subcategorizations. We find errors, but surprisingly few. For instance, we allow ourselves to say (16), but not (17).

- (16) Tell me the answer.
- (17) *Say me the answer.

It is not clear why a child would not make this extension, and allow for (19) as well as (18). In fact, if one searches on CHILDES the examples below, one would find 74 instances of (18), but none of (19).

- (18) Gimme.
 (19) *Give you.

So the double-object construction is simply not generalized by L1 children either to new verbs or to a minimal extension of a given verb²¹:

give + *clitic 'me'*
 *give + *clitic 'pronoun'*

The point here is that this is a case of a deletable object, similar to the one found in Brazilian Portuguese or Japanese. It does not generalize in English, but remains tied to an “exception”. The fact that it is an “exception” of course raises the question of how a child can recognize it as an exception rather than as a first instance of a productive rule, i.e., why would the child not be led to generalize it to “give you” ([V NP \emptyset])?

This is typical of the conservatism that production reveals. Exactly why production grammars – which generate novel sentences all the time – produce fewer overgeneralizations than one would expect remains a challenge. Snyder’s claims about conservatism apply particularly to production, but we argue that the opposite may hold for comprehension; and that this discrepancy between conservatism in production and lack of conservatism in comprehension opens a window for us to observe conflicting sub-grammars in action within the same speaker. We use a well-studied case in generative L2 acquisition to illustrate our point in the next section²².

The Spanish overt pronoun constraint

To illustrate the existence of (incompatible) sub-grammars in the interlanguage, we started with a well documented case that was primarily used to show access to UG by adult L2 learners: the acquisition of the *Overt Pronoun Constraint* (OPC). In a nutshell, the OPC (Montalbetti, 1984) is a phenomenon that occurs in pro-drop languages and prevents the co-indexation of an overt subject pronoun in embedded clauses with quantified noun phrases as its intrasentential antecedent, such as in (20a). This means that if the matrix subject is a quantified NP, the co-referent embedded subject must be a null pronoun (*pro*), such as in (20b).

- (20) a. Nadie_i dijo que él_{*i/j} ganará el premio.
 No one said that he will win the prize
 ‘No one_i said that he_j will win the prize.’
- b. Nadie_i dijo que (*pro*_i) ganará el premio.
 No one said that *pro* will win the prize
 ‘No one_i said that he_j will win the prize.’

Perez-Leroux and Glass (1997) and Kanno (1997) showed that adult speakers of English acquired such constraint in L2 Spanish and Japanese respectively, even without ever receiving explicit instruction on the topic.

Although we fully agree with the interpretation that UG access allows L2 speakers to acquire the OPC, we would like to challenge the notion of interlanguage restructuring away from the L1. Taking the work by Perez-Leroux and Glass (1997) as the point of departure, we designed an experiment to tap into the comprehension versus production dichotomy, as explained in the previous section. Our goal was to show that even after L2 speakers consistently use the OPC in production (with clear signs of stabilization), they could still allow an incompatible rule (from their L1) to interfere in the interpretation of utterances under certain pragmatic conditions.

So, the first step was to replicate Perez-Leroux and Glass (1997) production experiment, where participants had to translate sentences with intrasentential and extrasentential antecedents. We updated the situations and controlled for the number of fillers and test items. After reading the situation, participants had to complete the translation, where the first word (the quantifier) was given to them. The situations used were similar to the ones below.

- Intrasentential example:

People often cheat on their taxes. The reason for this is that they feel everybody does it, and besides, the amount owed is often very small. So, *nobody thinks that he is really committing a crime.*

Translation: *Nadie ...*

- Extrasentential example:

Pablo Picasso has influenced many artists in the course of the last century. Even years after his death, artists mention his influence. We interviewed 10 of the best contemporary artists and, *every artist says that he is the best painter of all time.*

Translation: *Todo ...*

We decided to follow the same protocol as the previous study because we wanted our results to be comparable to the ones found by Perez-Leroux and Glass (1997).

We ran the first experiment with 58 undergraduate students enrolled in an *Advanced Grammar* class at our university. All 58 participants were native English speakers, and were enrolled in a course that is required for the Spanish major. We used this first experiment not only to elicit the production of the OPC by these participants, but also to filter the group of students, and select only the ones that consistently used the restrictions imposed by the OPC in their production. We set our threshold to 80%, and only considered participants scoring above this mark as the ones who could consistently produce their sentences respecting the OPC, i.e., not allowing the co-referentiality between the embedded overt pronouns and the matrix subjects. Out of 58 participants 27 showed this consistent pattern of OPC use in their production, and, consequently, were invited to participate in the comprehension experiment as well. The table in (23) shows the results for these 27 participants.

The comprehension experiment consisted of a sequence of short stories, each of which followed by three sentences that could describe the preceding story or not. Participants were asked to read the stories and then, for each sentence, decide if it described the story or not, and provide a brief explanation of why the sentence did/did not describe the story, such as the one in (21).

- (21) “*Michael’s friends decided to throw a party to celebrate his 21st birthday. They decided that they would all bring something to drink, but that Michael would not have to bring anything, since it was his birthday.*”

There were 15 stories in total, and we tested three quantifiers: *cada* (each), *ningún* (no one), and *todo* (every). The sequences of sentences they saw in the comprehension experiment were similar to the one presented in (22), and we controlled for fillers, and the quantifiers that appeared in the same sequence of sentences; no quantifier appeared more than once in each sequence. In the example below we see a filler in (22a), and two test items in (22b) and (22c) for the story in (21).

- (22) a. Los amigos dijeron que iban _____ a traer una bebida.
The friends said that (they) were going to bring a drink.
- b. Ningún amigo dijo que él iba _____ a traer una bebida.
No friend said that he was going to bring a drink.
- c. Cada amigo dijo que él iba _____ a traer una bebida.
Each friend said that he was going to bring a drink.

To code participants’ responses, we used not only their judgment of whether or not a sentence described the situation, but also their justifications. If in their responses they showed they understood the sentence, we coded them as follows:

- In (22b):
 - *True* = OPC reading (él = Michael)
No friend said that Michael was going to bring a drink;
 - *False* = Non-OPC reading (él = ningún amigo)
No friend_i said that he_i was going to bring a drink.

- In (22c):
 - *True* = Non-OPC reading (él = cada amigo)
Each friend_i said that he_j was going to bring a drink;
 - *False* = OPC reading (él = Michael)
Each friend said that Michael was going to bring a drink.

The results showed that there was a statistically significant difference between the results in production and comprehension (Wilcoxon Rank Sum test: p -value < 0.000; Z -score = 5.96). While all participants respected the restrictions in production more than 80% of the time (since this was our threshold for their selection), the overall OPC responses in comprehension were only 24.33%. The tables in (23) show how participants behaved in each of the quantifiers tested in both conditions.

(23) Production:

Quantifier	# of tokens	OPC production		Non-OPC production	
		#	%	#	%
Ningún	29	27	91.1%	2	6.9%
Cada	58	56	96.5%	2	3.5%
Todo	29	26	89.7%	3	10.3%

Interpretation:

Quantifier	# of tokens	OPC interpretation		Non-OPC interpr.	
		#	%	#	%
Ningún	97	13	13.4%	84	86.6%
Cada	148	28	18.9%	120	81.1%
Todo	92	41	44.6%	51	55.4%

We interpret these results as evidence that L2 speakers of Spanish who have acquired the OPC, and use it consistently in production, also have a competing sub-grammar where the co-referentiality between the quantified subject of the matrix clause and the embedded overt subject pronoun is perfectly allowed.

Evidence for MGs from other areas

Evidence from multilingualism

One of the consequences of the MG theory to the repertoire of morphosyntactic rules that exist in any given L2 grammar is that it can only increase in size, since new rules (created from L2 input) do not replace the existing rules (from L1). This idea finds direct support on the data that comes from the research in third language acquisition. Zobl (1992) tested adjacency conditions of verbs and objects and found out that in grammaticality judgment tests multilingual speakers will accept sentences that presuppose a grammar that over-generates. According to De Angelis (2007), Zobl's results suggest that 'multilinguals seem to create less conservative hence more powerful grammars than monolinguals do' (De Angelis, 2007, p.58). Any formal model that explains L2 or L3 acquisition has to

deal with the fact that those grammars seem to be ‘more powerful’ than monolingual ones. In MG this is the result of the proliferation of sub-grammars that populate the grammar of the multilingual speaker. In the section *Evidence of MGs in the interlanguage* we show that this effect is particularly pervasive in comprehension, and that, as Snyder (2007) argued for in L1, production can be much more conservative.

More important to our theory than this extended generative power of multilingual grammars is the data that comes from research on the facilitative effect of the L2 in the L3 acquisition process. Investigating the effects of L2 transfer into L3, Foote (2009) looked into the acquisition of the preterite/imperfect distinction by advanced foreign language speakers of Romance whose L1 is English, a language that does not show such aspectual distinction. She demonstrated that L1 English speakers who acquired two Romance languages were statistically more successful in acquiring Romance aspect than those who only learned one language. Their gain was not only an effect of transfer from the L2 into the L3. Their knowledge in both foreign languages benefited from the fact that they were exposed to a third language. Foote’s test with verb morphology showed that EnglishL1/RomanceL2 speakers had 77.01% of correct answers, while EnglishL1/RomanceL2/RomanceL3 speakers had an accuracy rate of 91.43% in their L2 and 87.86% in their L3. Another compelling example comes from Klein (1995). She looked at the acquisition of preposition stranding by speakers of languages that form wh-questions through wh-movement, and allow pied-piping, but not preposition stranding. Her results showed that multilinguals had an advantage over monolinguals in the acquisition of preposition stranding in English, even when none of the languages previously spoken by either group showed that phenomenon.

Further evidence of this cumulative effect of rules in the grammar of multilingual speakers is provided by Flynn et al. (2004), and their *Cumulative-Enhanced Model for Language Acquisition*. They looked at the acquisition of relative clauses in English by three groups: L1Kazakh/L2 Russian/L3 English, L1 Spanish/L2 English, and L1 Japanese/L2 English. Kazakh and Japanese are head-final, left-branching language, where relative clauses appear to the left of the head. Besides, in Kazakh, there are no ‘overt wh-operators and overt complementizers in relative clauses’ (Flynn et al., 2004, p.10). On the other hand, Spanish and Russian are head-initial, right-branching languages, just like English. In their L2 acquisition experiment, Flynn et al. (2004) showed that there was a significant difference in the acquisition of relative clauses in English between L1 Spanish and L1 Japanese speakers, which the authors attributed to positive transfer on the part of Spanish speakers. When the results from Kazakh L1/Russian L2/English L3 speakers were compared to the two other groups, the researchers found out that they behaved like Spanish speakers, and not Japanese ones, in what regards their knowledge of English relative clauses. The authors concluded that: (i) ‘prior CP development can influence development of CP structure in subsequent languages’ (p.10); (ii) ‘L1 does *not* play a privileged role in subsequent acquisition’ (p.3); ‘experience in any prior language can be drawn upon in subsequent acquisition’ (p.10) . The goal of the MG theory is to present a formal model to describe how this linguistic knowledge of prior languages can play a role in the acquisition of subsequent languages, independently of the source of the this prior knowledge.

Evidence from processing

Recent research on bilingual and L2 processing seems to indicate a great deal of cross-language interference among the languages spoken by an individual²³. Some findings show that proficient bilinguals are not capable of isolating their L1 even during monolingual tests where their L2 is not required (e.g., Kroll et al., 2006). Kroll et al. (2012) suggest that the research done on bilingual processing for the last 20 years shows that there is no single answer to the question of whether or not there is one or two representations for the two languages in the mind. They state that ‘representations for the bilingual’s two languages might be shared in some circumstances and independent in others,’ and add that the issue of shared representations ‘seem to depend on the type of difference or similarity under investigation, the language history of the bilinguals, their proficiency, and the context in which the two languages are learned and used.’ (p.2)

As we mentioned in the previous section, even if there are two independent representations for L1 and L2, the fact that we take the approach of *full transfer* at the initial state of L2 acquisition guarantees that all L1 properties will be present in the interlanguage. However, it is interesting to notice that current research on processing is pointing towards a deep connection and dependency between the grammatical representations of the two languages in a bilingual mind. During their background overview of bilingual processing research, Kroll et al. (2012) point out to two interesting facts:

- ‘There is compelling evidence that shows that it is virtually impossible to switch off the language not in use and that the parallel activation of a bilingual’s two languages can be observed in reading, listening, and in planning speech.’ (p.2) (Costa, 2005; Dijkstra, 2005; Kroll et al., 2006, among others); and
- Researchers found evidence that not only the L2 is affected by L1 during the initial and subsequent stages of acquisition, but that some L1 forms converge towards the L2 (Ameel et al., 2005). This interference of L2 in the L1 has been found both in the lexicon (Linck et al., 2009) and in the grammar (Dussias, 2003).

Although a full review of recent research on bilingual and L2 processing is beyond the scope of this paper, it is important to highlight the fact that research in processing is increasingly indicating the interconnection and interference between L1 and L2 in the brain, which suggests ‘a high level of plasticity within the language system’ (Kroll et al., 2012, p.3). We believe that a theory of representation that provides a formal mechanism to explain how productive linguistic properties from different languages can co-exist within the same grammar provides a more suitable formalism to establish connections between research on L2 processing and representation.

Conclusion

MG is a generative theory of representation that addresses issues related to variation and optionality. The core idea is that any human grammar readily accommodates contradictory sets of rules in sub-grammars that are classified according to their productivity. We argued that even monolingual grammars show the existence of these distinct sets of

rules, and we believe that this property of grammars allows us to propose a representational mechanism that works for both L1 and L2.

The most direct consequence of such theory is that optionality becomes a central property of human language, and we should see it in all its manifestations. Optionality can be triggered not only by input from different languages, but also from different sociolinguistic contexts, historical change, and idiosyncratic properties of a single language. One of our major claims is that access to different sub-grammars is much greater in comprehension than in production. This is far less obvious and less researched, but perhaps one major factor that makes the L2 experience different from L1, since the L2 speaker must consider and eliminate potentially misleading interpretations that never occur to the L1 speaker. Technically, one very important contribution of the theory is to place optionality out of individual rules, and into co-existing sub-grammars. We believe this solution is not only more in sync with minimalist requirements, but it also reflects more clearly what we see from empirical observations, and it allows us to be more formally precise in our predictions. Overall, we believe MG greatly facilitates the formal description of observable phenomena.

We also believe that the MG theory has a great potential to interact directly with current research in transfer, L2 processing, and multilingual acquisition. At this point, besides the observations made throughout the paper, we prefer not to further speculate on which processing approaches could dialogue directly with MG. However, we would like to reiterate our position on the need to be explicit about the nature of the grammar before we analyze its capabilities. If we understand parsing as “taking an input and producing some sort of structure for it” (Jurafsky and Martin, 2000, p.57), and if we accept that our parser operates with rules described in a symbolic grammatical representation, then it is vital that the nature of the representation be compatible with our parsing algorithm, otherwise our parsing procedures are doomed to failure. This direct interconnection between the properties of the parser and the nature of the grammar have to be taken very seriously when implementing computer models, but it is sometimes overlooked and not explicitly mentioned when dealing with human models. Because our view is strongly based on symbolic (as opposed to stochastic) paradigms to describe human language, we favor parsing algorithms that are explicit in the ways they interconnect with the appropriate grammars.

As far as social dimensions are concerned, the MG theory presents a descriptive mechanism that approximates all human grammars, including the ones describing interlanguage and heritage languages. From an MG perspective all grammars are created equal. Each individual will have a grammar with a unique configuration of rules, and these individual configurations may converge or diverge from what is considered standard in a given language by different social groups. However, an individual grammar is never deficient in any way. In general, in the bilingual and L2 literatures we see a tendency to establish artificial gold standards based on idealized L1 models. The consequence is that heritage and L2 grammars are frequently described as ill-conceived versions of this idealized monolingual grammar. We believe that the suggestion that all grammars have to some degree structures that deviate from the prominent and productive properties of a given language highlight the fact that from a descriptive

perspective L1, L2, bilingual, and heritage grammars all share the core properties of human language.

One key question that needs to be addressed in future work is what elements trigger the classification of productivity for a given language property. One of our recent projects is looking into the connection between MG representation and the research on language mode, as proposed by Grosjean (2008), in an attempt to investigate if speakers operating in a bilingual mode would have a more difficult time assessing the degree of productivity of a given rule. We are also looking into how sub-grammars generated during intermediate stages of the L1 acquisition (but later discarded by the speaker) could interfere in the acquisition of an L2. Another interesting argument raised by Lawall (2012) is that language typology might play a role facilitating (or not) the ability of L2 speakers to change their classification for the productivity of a given rule. By looking at the acceptance rate of unaccusative constructions with “se” in Spanish by native speakers of English and Brazilian Portuguese, she found out that while BP speakers have an easier time accepting sentences with the unaccusative “se” in Spanish, they also have a harder time blocking ungrammatical sentences without this particle. It appears that once English speakers acquire the use of “se” as a marker for unaccusativity, they more easily block the productivity of their L1 rule that does not require the presence of any overt unaccusative marker. BP speakers, on the other hand, have an easier time acquiring the unaccusative “se”, but a much harder time blocking the productive rule (from their L1) that allows sentences without it. Lawall uses the MG framework to interpret these results as an example of language typology influencing the way speakers classify the productivity of their sub-grammars.

A key dimension of our claim follows from Minimalism: *rules must be simple*. If rules are simple, they can be utilized in another language. If rules are complex, then it becomes impossible to export a subpart of the rule to another language, or to integrate into the rule system of another language. However, evidence suggests that the lexicon allows specific links to simple rules from other grammars with a fixed and limited impact. The key to cross-grammatical effects lies in the notion of rule simplicity.

Finally, one of the overarching goals of generative SLA research has been to show that there is a common language faculty operating in the acquisition of both L1 and L2. Most L2 researchers have tried to accomplish this goal by designing models that attempt to approximate the L2 acquisition process to what the generative theory proposed for L1. One of the innovative ideas behind the MG theory is that, contrary to what was proposed so far, it approximates L1 and L2 acquisition by highlighting the observable properties of bilingual grammars that are present in the L1. The original proposal by Roeper (1999), entitled *Universal Bilingualism*, stated that every monolingual speaker had in fact a multilingual grammar with a number of sub-grammars that reflected prominent properties of different human languages. We believe that the idea that any human grammar carries properties belonging to a vast array of languages is not only appealing to a UG audience eager to show the universality of the human language faculty, but also to bilingual and L2 researchers looking for a generative theory of language acquisition whose central theoretical construct highlights the fact that humans are essentially biologically programmed to be bilingual.

Acknowledgements

There are several people we would like to thank: (i) The members of the UMass Language Acquisition Research Center (LARC) for their comments on different stages of this work, in special: Jill deVilliers, Lisa Green, Magda Oiry and Barbara Pearson; (ii) Ana Pérez-Leroux, Rosemarie Tracy and the audience at Mannheim for their suggestions; (iii) Meghann Peace for her help during the experiment design, data collection and coding, Maria Turrero for helping us reviewing the situations, and all other graduate students in our courses who contributed to several interesting debates on MG and L2; and (iv) the UMass College of Humanities and Fine Arts for sponsoring LARC activities.

We would like to acknowledge the importance of this volume's contributors and thank all of them for engaging in a fruitful debate about Multiple Grammars Theory. Because of their indispensable contribution, we were able to provide further clarification to several MG concepts in our response (*Why Minimal Multiple Rules provide a unique window into UG and L2*). We encourage the reader to study their careful considerations and to read our follow-up paper where we clarify the minimalist nature of MG, detail our take on processing and symbolic parsing, and reanalyze the case of V2 from a MG perspective.

Last but not least, we want to express our gratitude to the Second Language Research editors, Silvina Montrul and Michael Sharwood Smith, for their support and advice, and to thank the anonymous reviewers who contributed to improve our paper with their detailed comments. All errors remain our own.

Notes

1. When we use the term 'rule' in this paper we are primarily thinking about morphosyntactic rules. However, as far as we can see, there is no incompatibility in extending the notions presented here to any system of rules applying to phonological, semantic or pragmatic levels.
2. For a more discussion of the full complexity of quotative inversion see Collins and Branigan (1997). Our approach is consistent with theirs but reanalyzed in terms of a heavy demand on simplicity, which leads to a re-analysis into two versions of the rule.
3. Another example comes from recent history. President Obama had to retake the oath of office because at first they asked him to say: "*And I will faithfully execute the laws*" and it had to be re-administered as: "*And I will execute the laws faithfully.*" The difference seems non-existent, yet the final adverb prefers an internal reading "faithful to the law" while the initial adverb is open, it could mean "faithful to God". These interpretive differences are enormous, and the re-administration was appropriate.
4. White (1991) shows that the word order in (5b) is actually quite common among French speakers of English. Notice that we do not have any evidence that this change is motivated by variations in meaning. Our observations are just speculative at this point, and an empirical study would need to be done in order to verify it.
5. This claim has important implications for grammatical theory as a whole which we will not explore here.
6. Further lexical variation is more idiomatic: (i) '*a certain someone has arrived*'; (ii) '*you must do it by a date certain*'.
7. We use the word *stable* instead of *final* to describe the ultimate attainment of a language because we believe it best describes the cognitive status of those grammars. A stable grammar for L_X can be modified by the acquisition of a second or third language. This is not only true for L2 and L3 (see De Angelis, 2007), but also for L1 in the case of attrition (see section *Evidence from processing* for details).

8. This is a simplified discussion of the complex interaction between context and syntax. For further discussion refer to Allen (2007) and Hyams (2011).
9. The MDG nature is far from obvious and an important challenge to the theory.
10. The question of independent or unified storage for L1 and L2 grammars in adult speakers is an empirical one, and we will not address it in this paper. As mentioned above the MG model would be compatible with both types of representation if we assume the *full transfer hypothesis*.
11. It is interesting to see the choice of words here, since they directly express the intuition behind the MG theory – “as if the grammar allowed optionality”. The core question we want to address in this paper is exactly how a grammar can allow for contradictory rules to co-exist.
12. See Truscott (2006) for several examples of optionality on current SLA research, and Sorace (2003) for a debate on near native grammars and optionality.
13. It is out of the scope of this paper to provide a full account of the MOGUL proposal. We limit ourselves to highlight its characteristics that are relevant to our analysis. We recommend that readers refer to Truscott and Sharwood Smith (2004) for a more detailed account.
14. Given the two complementary goals of MOGUL and MG, we believe it would be interesting to explore possible connections between them to see if one could inform the other in an attempt to create a more comprehensive theory of representation, language development, and processing.
15. We could even create a learning principle stipulating that the more ridiculous the speaker feels about his mistake, the less frequently he needs to be exposed to that language property.
16. For an explanation of why this is an interface phenomenon, and the latest view of the authors on the dichotomy between ‘interfaces’ and ‘pure syntax’, please see Sorace (2011).
17. There were also sentences with backward anaphora in the experiment, which are not relevant to the current discussion. See Sorace and Filiaci (2006) for full description.
18. See Odlin (2003) for a review.
19. As soon as we have articulated what two sub-grammars look like, as we have above, then it is obvious that speakers either use them both or do not. Of course, the argument does depend upon the claim that the theory requires the rules to be distinct rather than a complex formulation of a single rule with optionality built in.
20. We could call it *consistent optionality*.
21. Montrul (1997) shows that this construction is generalized by adult L2 speakers. The differences in ways L1 and L2 learners deal with the productivity of certain rules deserve to be studied in detail. While the L1 speaker may be building a rule verb by verb, the L2 speaker could be seeking to find the larger generalization too quickly. The conditions for double-object constructions are linked to phonology (monosyllables), verb class (Anglo-Saxon), and meaning. If only one were chosen we could easily imagine that the L2 speaker will make the rule too abstract. A similar prediction holds for other similarly limited classes. So-called perception verbs allow bare verb infinitives, but the class is surprisingly small. Thus we have ‘see John go’, but not ‘*perceive John go’ but rather ‘perceive John to go’. Since the more general form ultimately dominates one might expect an overgeneralization of forms like ‘see John to go’ or ‘*I made John to sing’ instead of ‘I made John sing’. Thus L2 overgeneralizations could have very natural explanations. Although we are not going to discuss this issue in greater length in this paper, we believe it is a central topic for future work in MG.
22. One could even imagine that this lack of conservatism in comprehension grammars is also motivated by pragmatic factors, along the lines of a *cooperative principle* (Grice, 1989).
23. In this section we limit ourselves to comment on the way the work done by Kroll and colleagues presents some evidence that support a theory of MG. It is not our goal to review the L2 processing literature at this point. Although there are other important contributions to the

L2 processing research, such as Clahsen and Felser (2006), we do not fully see in which way the proposal that “the syntactic representations adult L2 learners compute during comprehension are shallower and less detailed than those of native speakers” (p.3) would support the idea that there are multiple sub-grammars operating in the interlanguage. Unless, of course, one assumes that shallow parsing is governed by rules from yet another sub-grammar. In this case, an interesting future project using the MG framework would be to find evidence that L2 speakers can actually parse similar sentences in two different ways.

References

- Allen S (2007) Interacting pragmatic influences on children’s argument realization. In: Bowerman M and Brown P (Eds.), *Cross-linguistics perspectives on argument structure: Implications for learnability*. Erlbaum, Mahwah, NJ, pp. 191–210.
- Ameel E, Storms G, Malt BC and Sloman S (2005) How bilinguals solve the naming problem. *Journal of Memory and Language* 53, 60–80.
- Bauke L and Roeper T (2011) How phase-based interpretations dictate the typology of nominalizations. In: Grewendorf G and Zimmermann T. E (Eds.), *Grammar and Discourse*. From Sentence Types to Lexical Categories. De Gruyter, pp. 289–320.
- Carroll S (2009) Re-assembling formal features in second language acquisition: beyond minimalism. *Second Language Research* 25, 245–253.
- Chomsky N (1995) *The Minimalist Program*. MIT Press, Cambridge, MA.
- Clahsen H and Felser C (2006) Grammatical processing in language learners. *Applied Psycholinguistics* 27, 3–42.
- Collins C and Branigan P (1997) Quotative inversion. *Natural Language and Linguistic Theory* 15, 1–41.
- Costa A (2005) Lexical access in bilingual production. In: Kroll J and de Groot A (Eds.), *Handbook of Bilingualism: Psycholinguistics Approaches*. Oxford University Press, New York, NY, pp. 308–325.
- De Angelis G (2007) *Third or Additional Language Acquisition*. Multilingual Matters, Clevedon, UK.
- De Houwer A (2005) Early bilingual acquisition: Focus on morphosyntax and the separate development hypothesis. In: Kroll J and DeGroot A (Eds.), *Handbook of Bilingualism: Psycholinguistics Approaches*. Oxford University Press, New York, NY, pp. 30–48.
- Dekydspotter L, Sprouse R and Anderson B (1998) Interlanguage a-bar dependencies: binding construals, null prepositions and Universal Grammar. *Second Language Research* 14, 341–358.
- Dijkstra T (2005) Bilingual word recognition and lexical access. In: Kroll J and de Groot A (Eds.), *Handbook of Bilingualism: Psycholinguistics Approaches*. Oxford University Press, New York, NY, pp. 179–201.
- Dornyei Z (2009) *The Psychology of second language acquisition*. Oxford University Press, Oxford, UK.
- Doughty C and Williams J (1998) *Focus on Form in classroom second language acquisition*. Cambridge University Press, New York, NY.
- Dussias P (2003) Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition* 25, 529–557.
- Flynn S, Foley C and Vinnitskaya I (2004) The cumulative-enhancement model for language acquisition. *International Journal of Multilingualism* 1 (1), 3–16.
- Foote R (2009) Transfer in L3 acquisition: the role of typology. In: Kit In-grid Leung Y (Ed.), *Third Language Acquisition and Universal Grammar*. Multilingual Matters, Bristol, UK, pp. 89–114.

- Galasso J (2002) Interference in second language acquisition: A review of the fundamental difference hypothesis.
- Galasso J (2003) *The Acquisition of Functional Categories*. IULC Publications, Bloomington, IN.
- Grice HP (1989) *Studies in the Way of Words*. Harvard University Press, Cambridge, MA.
- Grosjean F (2008) *Studying Bilinguals*. Oxford University Press, Oxford, UK.
- Hendriks P and Koster C (2010) Production/comprehension asymmetries in language acquisition. *Lingua* 120 (8), 1887–1897.
- Holmberg A (2010) Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics* 36 (1), 1–48.
- Hyams N (1983) The pro drop parameter in child grammars. In: *Proceedings of the West Coast Conference on Formal Linguistics*. Vol. 2. pp. 126–139.
- Hyams N (1986) *Language Acquisition and the Theory of Parameters*. Reidel, Dordrecht, Holland.
- Hyams N (2011) Missing subjects in child language. In: de Villiers J and Roeper T (Eds.), *Handbook of Generative Approaches to Language Acquisition. Studies in Theoretical Psycholinguistics*. Springer, New York, pp. 13–52.
- Jurafsky D and Martin J (2000) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, Upper Saddle River, NJ.
- Kanno K (1997) The acquisition of null and overt pronominals in Japanese vs English speakers. *Second Language Research* 13, 265–287.
- Klein E (1995) Second versus third language acquisition: is there a difference? *Language Learning* 45 (3), 419–465.
- Kroch A and Taylor A (1997) Verb movement in Old and Middle English: Dialect variation and language contact. In: van Kemenade A and Vincent N (Eds.), *Parameters of Morphological Change*. Cambridge University Press, Cambridge, UK, pp. 297–235.
- Kroll J, Bobb SC and Wodniecka Z (2006) Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition* 9, 119–135.
- Kroll J, Bogulski C and McClain R (2012) Psycholinguistic perspectives on second language learning and bilingualism. *Linguistic Approaches to Bilingualism* 2 (1), 1–24.
- Lardiere D (2009) Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25, 173–227.
- Lawall R (2012) A alternância causativa/incoativa em espanhol como L1 e L2. *Revel* 10 (18), 1–27.
- Liceras J (2009) On parameters, functional categories and features ... and why the trees shouldn't prevent us from seeing the forest. *Second Language Research* 25, 279–289.
- Liceras J and Diaz L (1999) Topic-drop versus pro-drop: null subjects and pronominal subjects in the Spanish L2 of Chinese, English, French, German and Japanese speakers. *Second Language Research* 15 (1), 1–40.
- Linck J, Kroll J and Sanderman G (2009) Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second language learning. *Psychological Science* 20, 1507–1515.
- Long M (2000) Focus on form in task-based language teaching. In: Lambert R and Shohamy E (Eds.), *Language Policy and Pedagogy: Essays in honor of A. Ronald Walton*. John Benjamins Publishing Company, Philadelphia, pp. 179–192.
- MacSwan J (2005) Codeswitching and generative grammar: A critique of the MLF model and some remarks on “modified minimalism”. *Bilingualism: Language and Cognition* 8 (1), 1–22.
- Montalbetti M (1984) After binding; on the interpretation of pronouns, ph.D. Dissertation. Massachusetts Institute of Technology.

- Montrul S (1997) On the parallels between diachronic change and inter-language grammars: the L2 acquisition of the Spanish dative case system. *Spanish Applied Linguistics* 1, 87–113.
- Montrul S (2008) *Incomplete Acquisition in Bilingualism*. John Benjamins, Amsterdam.
- Montrul S (2011) Multiple interfaces and incomplete acquisition. *Lingua* 121, 591–604.
- Montrul S and Yoon J (2009) Putting parameters in their proper place. *Second Language Research* 25, 291–311.
- Odlin T (1989) *Language Transfer*. Cambridge University Press, Cambridge, UK.
- Odlin T (2003) Cross-linguistic influence. In: Doughty C and Long M (Eds.), *Handbook on Second Language Acquisition*. Blackwell, Oxford, pp. 436–486.
- Pavlenko A and Jarvis S (2002) Bidirectional transfer. *Applied Linguistics* 23, 190–214.
- Perez-Leroux A and Glass W (1997) Op e cts on the L2 acquisition of Spanish. In: Perez-Leroux A and Glass W (Eds.), *Contemporary Perspectives on the Acquisition of Spanish*. Cascadilla Press, Somerville, MA, pp. 149–165.
- Pintzuk S (1999) Phrase Structures in Competition: Variation and Change in Old English Word Order. Garland, New York, NY.
- Reuland E (2011) Syntax and interpretation systems: how is their labour divided? In: Boeckx C (Ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford University Press, Cambridge, UK, pp. 377–395.
- Roeper T (1999) Universal bilingualism. *Bilingualism, Language and Cognition* 2, 169–186.
- Romaine S (2003) Variation. In: Doughty C and Long M (Eds.), *The Handbook of Second Language Acquisition*. Blackwell Publishing, Malden, MA, Ch. 14, pp. 409–435.
- Schwartz B and Sprouse R (1996) L2 cognitive states and the full transfer/full access model. *Second Language Research* 12, 40–72.
- Serratrice L (2005) The role of discourse pragmatic in the acquisition of subjects in Italian. *Applied Psycholinguistics* 26, 437–462.
- Serratrice L, Sorace A and Paoli S (2006) Crosslinguistic influence at the syntax-pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition. *Bilingualism, Language and Cognition* 7 (3), 183–205.
- Sharwood Smith M and Truscott J (2005) Stages or continua in second language acquisition: A mogul solution. *Applied Linguistics* 22 (2), 219–240.
- Snyder W (2007) *Child Language: The Parametric Approach*. Oxford University Press, Oxford, UK.
- Sorace A (2003) Near-nativeness. In: Long M and Doughty C (Eds.), *Handbook of Second Language Acquisition*. Blackwell, Oxford, pp. 130–152.
- Sorace A (2011) Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism* 1 (1), 1–33.
- Sorace A and Filiaci F (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22 (3), 339–368.
- Truscott J (2006) Optionality in second language acquisition: A generative, processing-oriented account. *International Review of Applied Linguistics in Language Teaching (IRAL)* 44 (4), 311–330.
- Truscott J and Sharwood Smith M (2004) Acquisition by processing: A modular perspective to language development. *Bilingualism: Language and Cognition* 7, 1–20.
- Vainikka A and Young-Scholten M (1994) Direct access to X'-theory: evidence from Korean and Turkish adult learning German. In: Hoekstra T and Schwartz B (Eds.), *Language Acquisition Studies in Generative Grammar*. John Benjamins Publishing Co., Amsterdam, The Netherlands, pp. 256–316.
- Valian V (1990) A problem for parameter-setting models of language acquisition. *Cognition* 35 (2), 105–122.

- Valian V and Eisenberg Z (1996) The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language* 23, 103–128.
- van Hout A, Harrigan K and de Villiers J (2010) Asymmetries in the acquisition of definite and indefinite NPs. *Lingua* 120 (8), 1973–1990.
- White L (1991) Adverb placement in second language acquisition: some effects of positive and negative evidence in the classroom. *Second Language Research* 7, 133–161.
- White L (2003a) On the nature of interlanguage representation. In: Doughty C and Long M (Eds.), *The Handbook of Second Language Acquisition*. Blackwell Publishing, Malden, MA, Ch. 2, pp. 19–42.
- White L (2003b) *Second Language Acquisition and the Universal Grammar*. Cambridge University Press, Cambridge, UK.
- White L (2009) Grammatical theory: Interfaces and L2 knowledge. In: Ritchie W and Bathia T (Eds.), *The New Handbook of Second Language Acquisition*. Emerald, Bingley - UK, pp. 49–68.
- Yang C (2003) *Knowledge and Learning in Natural Language*. Oxford University Press, Oxford, UK.
- Zobl H (1992) Prior linguistic knowledge and the conservatism of the learning procedure: grammaticality judgements of unilingual and multilingual learners. In: Gass S and Selinker L (Eds.), *Language Transfer in Language Learning*. Newbury House, Rowley, MA, pp. 176–196.

Copyright of Second Language Research is the property of Sage Publications, Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.