

Putting parameters in their proper place

Silvina Montrul and **James Yoon** *University of Illinois at Urbana-Champaign*

Received June 2008; revised September 2008; accepted October 2008

Seeing the logical problem of second language acquisition as that of primarily selecting and re-assembling bundles of features anew, Lardiere proposes to dispense with the deductive learning approach and its broad range of consequences subsumed under the concept of parameters. While we agree that feature assembly captures more precisely the complexity of the form–meaning mapping task in second language acquisition, we disagree with the dismissal of a parametric approach. We argue instead that the notion of parameter is not incompatible with feature assembly, if parameters and features are understood in a particular way.

Keywords: parameters, features, syncretic categories, genericity, perfectivity

I Introduction

Throughout its evolution from the 1980s until today, the generative approach to second language acquisition has always defined, and continues to define, the learning problem in the following terms. Just like a child learning her first language, the adult second language (L2) learner must discover and build a mental representation for particular form–meaning mappings in the second language from exposure to L2 input. While in many ways input overdetermines the grammar – in the sense that some particular structures of the language may be very frequent and obvious – there are many other ways in which input underdetermines the complexity of the system, since many abstract properties of language are not readily observable from the input and cannot therefore

Address for correspondence: Silvina Montrul, Department of Spanish, Italian and Portuguese, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building, MC-176, 707 South Mathews Avenue, Urbana, IL 61801, USA; e-mail: montrul@illinois.edu

© The Author(s), 2009. Reprints and permissions:

<http://www.sagepub.co.uk/journalspermissions.nav>

10.1177/0267658308100290

be acquired inductively. These are the now famous Abundance of the Stimulus and Poverty of the Stimulus problems, germane to both first language (L1) and L2 acquisition.

Under the generative-nativist view of language acquisition, the initial state of L1-acquiring children, S_0 , is hypothesized to be Universal Grammar (UG). UG is assumed to consist of an invariant part ('principles') and a part that permits a finite number of options ('parameters'). Guided by input, the child successfully fixes the parametric options in UG to construct the grammar of L1. The difference between L1 and L2 acquisition is that in the latter the learner already comes to the task with a fully developed L1 so that the S_0 in L2 acquisition is S_{L1} , the end-state grammar of L1. Through exposure to L2 input, the L2 learner must gradually reconfigure her interlanguage grammar to fit the parameters of the target language grammar, in part by accessing the principles and parameters of UG, as under the Full Transfer/Full Access Hypothesis (Schwartz and Sprouse, 1996).

In her keynote article, Lardiere summarizes how the concept and empirical coverage of the construct 'parameter' that is central to nativist accounts of language acquisition has evolved in recent years, and argues that under the current view of parameters where parameters equal features, the notion might no longer be theoretically suited to predict and explain either the complexity in the cross-linguistic realizations of a given grammatical phenomenon (e.g. plural marking in English, Mandarin Chinese and Korean) or how such features are acquired in a second language. Concluding that the notion of parameter is 'distracting' as it stands (p. 180), Lardiere proposes to change the way we conceptualize the L2 learning task from one of trying to understand how parameters are 'reset' to that of investigating how learners acquire and re-assemble features in particular lexical items through a 'contrastive analysis' with L1 (p. 219), a task that she deems is not within the purview of parameters.¹ More generally, while seemingly acknowledging that UG constrains L2 acquisition, she wants to banish parameters as currently understood, at least from L2 acquisition, if not from the theory of UG altogether.

We agree with Lardiere that learning how features are assembled into lexical items is something that learners must engage in when structuring (or restructuring, in the case of L2 learners) their grammars. But we

¹ In her emphasis on the role of features in L2 acquisition, she echoes the sentiments of Travis (2008) and Lardiere (2008) herself.

deeply disagree with Lardiere's dismissal of the role of parameters and with her suggestion that the notion of parameters is no longer apt in L2 acquisition. We also disagree with the view that feature assembly is not constrained by UG or parameters and suggest ways in which they can be integrally involved in constraining feature assembly.

II Parameters: are they still useful?

As Lardiere's argument in her article elsewhere (Lardiere, 2008) against the theoretical utility of parameters is grounded in a particular understanding of UG and parameters, we shall first make explicit what we take to be the assumptions that she is making about these central concepts.

The backdrop of her discussion is the proposal in Chomsky (1998) that the grammar of a language is defined by a selection of features from a universal inventory and their assembly into a set of lexical items (LIs), which are accessed by the Computational System to construct syntactic expressions that are interpreted at the interfaces. Aside from defining a lexicon of LIs specific to a given language,² Chomsky takes the selection and assembly of features into LIs to define the parameter setting of a given language (Chomsky, 1998: 13, quoted by Lardiere). This is so for the following reasons.

The leading idea in the Minimalist Program is that parameters should be lexically grounded (Lexical Parameterization Hypothesis). This means that LIs must contain the ingredients necessary to characterize the parametric profile of a given grammar. What could these features be? A moment's reflection shows that these ingredients cannot be found in phonetic features, which are irrelevant to syntax, or in semantic–conceptual features (such as theta-roles and ingredients that differentiate lexical classes), which are universal. Nor can they be found in the language-specific pairings of sound and meaning.³ For reasons such as these, a second hypothesis – Functional Parameterization Hypothesis – is adopted, according to which the ingredients necessary to characterize parametric differences between languages are restricted

² Chomsky (2001: 4) proposes a constraint on the selection and assembly of features that it be 'one time'. We take this to mean that once the features have been selected and assembled into LIs, the lexicon of a language is fixed so that LIs are not created 'online' (from the component features) every time they are accessed by the Computational System.

³ For example, the fact that French associates *chien* with the concept of DOG while English associates the concept with *dog* cannot be recruited to explain the parametric differences between the two grammars.

to features of functional categories. This means that the features that determine the parametric profile of a language come from the class of syntactic features associated with functional categories, which we call 'formal features'. Thus, the absence or presence of certain formal features and how they are packaged into functional categories determines the parametric profile of a given language. It is when things are understood this way that we can speak of (formal) features being equivalent to parameters.⁴

Lardiere expresses alarm over the ramifications of this view of parameters and finds such parameters of little use, especially in the context of L2 acquisition. Specifically, she is concerned about the fact that the number of such small-scale 'micro' parameters will be quite large and that microparameters lack a 'cluster of deductive consequences' (Lardiere, this issue: 177), whereas the older, or 'macro' parameters were quite small in number and each parameter was associated with a range of deductive consequences. The concerns expressed are based on considerations of learnability restrictiveness. However, there is no reason to think that the relatively large (but still manageably finite) number of microparameters will constitute a learnability obstacle if they are lexicalized, as lexical learning is evidently possible. On the contrary, learning large-scale macroparameters that are not grounded in lexical properties might prove to be a bigger challenge. What exactly are the 'triggering data' for a particular parameter setting, if parameters are not lexically grounded? It is unclear.

With regard to the argument that microparameters lack a 'cluster of deductive consequences', it is true that the reach of a microparameter is more limited than that of a macroparameter, but then we do not know, outside of the Head-Initial/Final Parameter, any other macroparameter whose presumed 'cluster of deductive consequences' has held up under scrutiny, the classic Pro-drop Parameter (Rizzi, 1982) or the Configurationality Parameter (Hale, 1983) being cases in point. The putative cluster of effects associated with a given parameter have all been decomposed into smaller (hence, less interesting for Lardiere) microparameters.⁵

⁴ We will make specific proposals about the type of formal features that can be 'selected and assembled' to yield the parametric setting of a language.

⁵ The same point holds for the Nominal Mapping Parameter (NMP), which figures in Lardiere's discussion against the utility of parameters. We find her dismissal of this parameter odd, because the NMP is not a microparameter, but a macroparameter, with a presumed 'cluster of deductive

While Lardiere only sees potential problems with the new perspective on parameters, and we are sensitive to her concerns, in fact, the thesis of functional parameterization was adopted because it promises many theoretical dividends. For example, if parameters are lexicalized, the Computational System (CS) can be invariant. All languages have the same CS, though a particular operation from among the universal repertoire of syntactic operations (say, Move/Internal Merge) may be triggered by a feature of a functional category that is selected in one language but not another, giving rise to a ‘parametric’ difference between the two. This is the reason why Chomsky can speak of *the* (i.e. unique) Computational System of Human Language (=C_{HL}), or make claims to the effect that there is only one human language. This perspective on grammar variation is not novel, but shared with other lexically based frameworks, notably, Categorical Grammar.

By the same token, the fact that 100 functional categories with binary parameters can yield an astronomic number of potential grammars (a point Lardiere highlights; Lardiere, this issue: footnote 6) is not a learnability problem. Even if all the 100 parameters are independent, as long as they are lexicalized and the CS is unique, there is no learnability problem, as Kayne (2005) acknowledges. Besides, it is highly likely that the parameters will be inter-dependent, and not independent, so that setting the value of one has implications for the parameter setting of others. This will help reduce the number of potential grammars. By contrast, classical parameters were essentially unrestricted and could be defined directly on the CS, yielding different kinds of CSs (called Core Grammars).⁶ Therefore, functional parameterization yields a more restrictive theory of parameterization than the classic theory, allowing us to answer the question, ‘What is a possible parameter?’, which was largely unanswerable in the older approach.

In addition, old-style macroparameters characterized Core Grammars, and left much of the detailed analyses of individual phenomena to the ill-defined ‘periphery’. The newer, microparameters allow for

consequences’. Since her overall position can be read as endorsing macroparameters and dismissing microparameters, the fact that the predictions of the NMP are falsified cannot be used as evidence that microparameters are of little use theoretically.

⁶ For example, the Subjacency Parameter applies to the DS–SS mapping, but not the SS–LF mapping. Similarly, the head-initial/final parameter of X-bar theory is defined for a particular CS (Core Grammar) as a whole. The provenance of other parameters, such as the Noun Compounding Parameter, is even more mysterious. It seems that the presence of a particular structure (or rule) is the parameter.

a greater range of empirical coverage, precisely because they are more numerous than classical parameters, as Kayne has demonstrated in his recent work.⁷ In saying this we are not denying the possibility that an adequate theory of UG might need macroparameters, in addition to microparameters. We are simply attempting to set the record straight on the theoretical rationale behind the proposal for microparameters.

Given this perspective of parameters, we do not believe that the notion of parameter should be discarded from the theory, even if we restrict parameter setting to the selection of formal features and their assembly into (functional) LIs in a given language. Instead, we should investigate the questions that are raised by this overall approach; for example:

- Just what sorts of formal features can be selected?
- What are the formal features that UG makes available to begin with?
- Since features are assembled into LIs, what constraints, if any, exist on the assemblage?
- Is there a logical order in which features are acquired? That is, does the selection of a feature entail that of another?
- Are parameters independent? That is, does choosing a value for a parameter for one class of LI have consequences for other classes?

Now, it is true that the answers to these and other questions may not have the dramatic flair associated with old-style parameters which sought to tie together seemingly unrelated clusters of properties (but then remember that the promised drama never materialized), or with investigation of the design properties of the Computational System (which

⁷ We do not believe, however, that this approach entails that every point of difference between languages is accounted for by a microparameter, as Lardiere suggests. First of all, since parameters draw from formal features, only syntactic differences should fall under parameters. Second, parameters cannot explain differences between languages that violate syntactic rules. Neither can parameters account for non-syntactic effects that may be associated with a given parameter. For example, if plural marking preferentially goes with animates but not inanimates (as in Korean), this is not something that should be accounted for by the parameter which chose the feature [Plural] for the language, since there is no reason to believe that animacy is a syntactically relevant feature in Korean.

Kayne's point that the exceptional distribution of *enough* (*rich enough* vs. **enough rich*) might be accounted for by a parameter should be understood in light of these restrictions. The distribution of *enough* can be accounted for by adding an uninterpretable EPP feature on this item (Deg head). This is clearly within the purview of what parameters can do. On the other hand, the fact that English has a phrase *by and large*, which violates syntactic restrictions (Law of Coordination of Likes), cannot be considered a 'parameter' of English grammar.

has dealt in lofty concepts like ‘economy’, ‘locality’, ‘perfection’, etc.). However, to the extent that parameter setting is feature selection and assembly, exploration of answers to these questions now becomes the bread and butter of parametric work.

III Parameter resetting in L2 acquisition

Suppose we understand parameters in the sense described above, i.e. as lexicalized and limited to the formal features that can be selected and assembled differently on functional categories. If L2 acquisition is UG-constrained and involves parameter resetting, we need to spell out specific assumptions about what it means for parameters to be ‘reset’. Given what we have said, the task reduces to one of the following scenarios. If a feature does not exist in L1 but exists in L2, parameter resetting involves adding (and assembling) a formal feature to the relevant functional category in L2. If a feature exists in L1 and L2, but is packaged differently, parameter resetting would entail figuring out the correct assemblage of the feature in L2.⁸

Regarding the first scenario – selection of formal features – there has been a line of thinking that holds that L2-ers have difficulty with certain formal features that are missing in L1. This is the idea of transfer, cashed out in the vocabulary of features; i.e. the features of L1 affect those of L2. The idea goes under various rubrics; e.g. Valueless Features Hypothesis (Eubank, 1996), the Failed Functional Features Hypothesis (Hawkins and Chan, 1997), the Local Impairment Hypothesis (Beck, 1998) and, most recently, the Interpretability Hypothesis (Tsimplici and Dimitrakopoulou, 2007), which all claim that failure to achieve native-like representations in the L2 is somehow related to malfunctions in the system of formal features, specifically the uninterpretable formal features, either because uninterpretable formal features may be selected but cannot be assigned values in the L2, or because uninterpretable formal features that are not selected by the L1 are hypothesized to be unacquirable due to some sort of critical period effect.

But here is another problem for second language acquisition. If the notion of parameters is only linked to formal uninterpretable features,

⁸ In the current theory, there are two types of uninterpretable features: uninterpretable phi/agreement features and the EPP feature. These are uninterpretable because even though they are lodged on a functional category, the information they express is not inherent to these categories, as they signal the contextual dependencies between the functional category on which they occur and a dependent.

what constrains the acquisition of interpretable formal features, which are based on universally available semantic–conceptual features? We have in mind the acquisition of features like genericity, definiteness and perfectivity when they are encoded syntactically as interpretable features.⁹ As we discuss shortly, there is cross-linguistic variation as to how these interpretable features are grammaticized, and the mapping of these features to lexical forms in particular languages is also a stumbling block in L2 acquisition, at least initially. A strict interpretation of the previous ‘failed features’ proposals entails that interpretable formal features are not subject to such a constraint and that an interpretable feature that is absent in L1 and present in L2 can still be acquired. The question is whether this is the case.

In addition to the issue of the acquisition of formal features – interpretable or otherwise – our reading of the literature is that there has not been much work on the topic of how features are re-assembled into LIs in L2 acquisition. Lardiere, who has worked extensively on this topic, thinks that while feature selection is equivalent to parameter setting, feature assembly is not (Lardiere, 2008, makes this point more clearly). From this premise, she goes on to argue that since feature selection by itself is insufficient to model the complexity of the remapping of form and meaning in L2 acquisition, which requires knowledge of how features are assembled in the LIs in the target language, the role of parameters in L2 acquisition is superfluous at best and perhaps ‘distracting’. In place of parameters, she advocates a greater role for L1, suggesting that, through a ‘contrastive analysis’ of L1 and L2 LIs, speakers manage to construct – or not – the relevant feature assemblies for L2 LIs, perhaps on an inductive basis. According to our understanding, Lardiere’s feature assembly approach does not clearly distinguish between interpretable and uninterpretable features.

We agree with Lardiere that the difference between languages cannot simply reside in whether or not a particular formal feature is chosen, but depends on how features – both interpretable and uninterpretable – are packaged. However, we do not think that feature re-assembly happens inductively through ‘contrastive analysis’. We shall

⁹ We make a distinction between semantic–conceptual features that are universally available and the grammatical encoding of such features. We restrict the term interpretable formal features to refer to the latter. For example, all humans can distinguish definite and indefinite reference, but the distinction is not grammaticized in all languages. When it is and is marked on the relevant functional category, Definiteness becomes an interpretable formal feature.

sketch a way in which not only feature selection but assembly can be conceived of in parametric terms. Before doing that, we first emphasize our common ground: that the (non-)existence of formal features is not sufficient to characterize the differences between the two languages and, consequently, that L2 acquisition must involve re-assembly and remapping of features into LIs, without prejudicing the answer to the question as to whether such a process falls within the purview of UG and parameters.

IV Feature selection and feature (re-)assembly in L2 acquisition

With the example of how the feature [Plural] is instantiated in English, Mandarin Chinese and Korean, Lardiere argues that just selecting a feature or failing to do so does not aptly capture the amount and nature of variability typically observed in L2 grammars. She points out that what feature selection fails to account for is how features are realized by different pieces of morphology and syntax in different languages. And here is where she claims that the L1 appears to play a very prominent role.

Another example of this problem – that feature selection is not enough for successful L2 acquisition – is the acquisition of gender. Languages like Spanish, French, German, Dutch, Russian, etc. all mark gender on nouns and have gender agreement in the syntax with other elements of the DP, while English nouns do not instantiate a formal feature for gender. While Spanish and French has two genders (feminine, masculine), Dutch, German and Russian have a three-way gender system (masculine, feminine, neuter). There are a plethora of L2 studies documenting that English-speaking L2 learners of languages with gender find it very difficult to master gender agreement to a level of accuracy (above 90%) comparable to a native speaker, a result which has been used to support the Failed Functional Features Hypothesis that restricts parameterization to feature selection (Franceschina, 2001; 2005; Hawkins and Franceschina, 2003). Yet, it is not uncommon to find that speakers of gender-marking languages also make errors when acquiring another gender-marking language (Bruhn de Garavito and White, 2003), although it is also true that speakers of gender-marking languages make fewer errors than speakers of gender-less languages (Sabourin *et al.*, 2006). The problem here is not just one of feature selection from the inventory of features provided by

Universal Grammar but also one of lexical mapping.¹⁰ And the magnitude of the learning problem for L2 acquisition might be different.

Here is another example, identified in Ionin and Montrul's current project on definite articles and genericity in English and Spanish as second languages (see also Ionin *et al.*, 2004). English allows only specific reference for definite plurals (*The hungry dogs are dangerous*), instantiating genericity in bare plurals (*Hungry dogs are dangerous*). By contrast, Spanish allows both generic and specific reference of nominals marked with the definite plural article. *Los perros hambrientos son peligrosos* ('the hungry dogs are dangerous') refers both to hungry dogs in general (the kind reading) and to a specific pack of hungry dogs being dangerous.

In one study (Ionin and Montrul, in preparation), Korean-speaking learners of English, whose L1 does not have Determiners, and Spanish-speaking learners of English, whose L1 has Determiners like English, were tested. For the Korean speakers, the learning problem entails selecting and mapping features onto a specific morphological form. Since their L1 has no Determiners and no features bundled to them, Koreans have to learn the existence of Determiners in English and their interpretations, as expressed by the features [D, +Def, ±PI] associated with definite Determiners.¹¹ But for the Spanish speakers, it is a question of feature re-assembly. Spanish has definite Determiners that contain the feature bundle [D, +Def, ±PI, ±Generic]. When learning English Determiners, the Spanish-speaking learners need to subtract the interpretable feature [+Generic] from the bundle and learn that this feature is expressed by bare plurals (which are not allowed in Spanish in preverbal subject position).¹² In another study, Ionin and Montrul (in preparation), English-speakers learning Spanish were tested. English-speaking learners of Spanish have to realize that bare plurals with

¹⁰ Clearly, L2 learners have to learn how gender is morphologically instantiated at the lexical-morphemic level in each language, since words that may be masculine in one language (*nose* in Portuguese) are feminine in another (*nose* in Spanish).

¹¹ Tania Ionin (personal communication) questions our assumption that Korean does not have Determiners, since it has demonstratives and definiteness is one of the properties expressed by demonstratives in Korean.

Are demonstratives determiners in Korean? Perhaps. If so, then Korean has Determiners optionally and Determiners-Demonstratives package features like spatial deixis, anaphoricity, etc. in addition to definiteness and specificity. Alternatively, we may take Demonstratives to be modifiers (adjectives) in Korean.

¹² When genericity is expressed by bare plurals, it is not clear how and where the feature is packaged. Tania Ionin (personal communication) points out that definite plural Determiners and bare plurals are

generic reference are not allowed in Spanish, and must now append the feature [+Generic] to Spanish definite articles. Comparing the learning problem in the two languages, while English learners of Spanish have to add the feature [+Generic] of bare plurals to definite articles (i.e. learn a new interpretation for definite plurals), Spanish-speaking learners of English have to do the opposite: namely, they have to subtract a feature from the definite plural (i.e. unlearn the generic interpretation of definite plurals) and remap it to bare plurals. The overall findings of these studies can be summarized as follows.

While almost 60% of the Korean learners of English acquired English articles and their features in English (i.e. learned that definite plurals have specific but not generic reference), the English- and Spanish-speaking learners presented a different profile. It proved easier for English-speaking learners of Spanish to add [+Generic] to Spanish articles than for Spanish-speaking learners of English to fail to attribute [+Generic] to English articles. This study suggests that feature re-assembly is a difficult task even when the L1 and the L2 share the same features; however, it is not impossible, since near-native participants achieved target-like performance. In general, it seems that when two or more features are bundled and mapped onto a single morphological form (yielding a syncretic category, according to Giorgi and Pianesi, 1997) it is difficult for the learners to break up this bundle and scatter the features onto new and different forms.

Other examples of this problem are the acquisition of preterite and imperfect Romance forms by English-speaking learners (Montrul and Slabakova, 2002; Slabakova and Montrul, 2003) and the acquisition of the two Spanish copulas *ser* and *estar* by English-speaking learners (Bruhn de Garavito and Valenzuela, 2008; Montrul, 2008). According to the aspectual analysis adopted by Bruhn de Garavito and Valenzuela, *ser* and *estar* are allomorphs: the overt morphological expression of the aspectual distinction [\pm Perfective]. The feature [Perfective] is available in both Spanish and English, but again the feature is packaged in different ways in the two languages. In Spanish, *ser* contains the feature [-Perfective] and *estar* the feature [+Perfective]. The problem for English speakers learning Spanish is that copula *be* does not encode

the two cross-linguistically common ways in which genericity is expressed. One option is to analyse the bare plural as headed by a null Determiner, to which the feature is added. However, Ionin points out that this analysis might be problematic, especially given the different distributions of bare plurals and DPs in the relevant languages.

the perfective/imperfective distinction, as do English stative verbs in general, according to Giorgi and Pianesi (1997). In English, perfective/imperfective are not features of the copula, but expressed by aspectual morphology (grammatical aspect) or computed in conjunction with other elements in the sentence. Compare the following examples:

- | | | |
|----|-------------------------------|----------------------------------|
| 1) | English: | Spanish: |
| a. | Robert <i>is</i> drunk. | Roberto <i>está</i> borracho. |
| b. | Robert <i>is a</i> drunk. | Roberto <i>es</i> (un) borracho. |
| c. | Robert <i>is</i> silly. | Roberto <i>es</i> tonto. |
| d. | Robert <i>is being</i> silly. | Roberto <i>está</i> tonto. |

Sentence (1a) describes the state of being drunk, which typically occurs within one delimited time period, whose beginning or end (or both) is implied. The English sentence is interpreted as [+Perfective] and the Spanish equivalent is expressed with *estar*. Sentence (1b) describes the event of being drunk as a quality of Robert, and is thus interpreted as [–Perfective]. Note that in English the DP takes an indefinite article, which contributes to the [–Perfective] interpretation of the sentence. In Spanish, the presence of the indefinite article is optional, while the natural copula to use is *ser*. The adjective *silly* in sentence (1c) is sufficient to guarantee a [–Perfective] interpretation in English because the copula is unmarked for perfectivity, and the Spanish equivalent also takes *ser*. But in (1d), progressive *-ing* is needed to turn a [–Perfective] state in English into a [+Progressive] and [+Perfective] event with the same adjective.¹³ Once again, Spanish expresses this same interpretation with [+Perfective] lexicalized in the copula *estar*. These sentences illustrate how the potential aspectual ambiguity of the English copula – which may be said not to be specified for the feature [±Perfective] – is resolved in English via inflectional morphology or in conjunction with other elements in the sentence. By contrast, to interpret aspect in Spanish copular constructions, English-speaking L2 learners must associate the features, already available from their L1, with the copulas, and lexically re-assemble [–Perfective] to *ser* and [+Perfective] to *estar*. As Bruhn de Garavito and Valenzuela show, even advanced L2 learners of Spanish have tremendous difficulty with this task.

¹³ Perfective in the sense described here denotes delimitedness of an event. Present progressive in conjunction with a state-denoting predicate in English is a way of signalling that the event has an internal duration (hence, delimited/perfective). This is how [+progressive] comes to signal [+perfective]. It might be better to use [delimited] or [telic] as the relevant features. We are keeping the feature name in part because the name has been used by earlier researchers.

To summarize, we have illustrated how the problem for L2 acquisition is greater than merely selecting from UG features not already selected in the L1, and on this we agree with Lardiere's conceptualization of the problem. The question that arises is how feature (re-)assembly is constrained in L2 acquisition. In her haste to get rid of parameters, it is not clear to us whether Lardiere conceives of feature assembly as a deductive operation constrained by Universal Grammar or sees it as an inductive process as envisaged under Contrastive Analysis, where a biologically determined UG is not invoked at all. She writes 'The magnitude of the task of inductively mapping morpholexical items that occur in the input in a particular linear order to functional feature matrices should not be underestimated' (Lardiere, this issue: 183). Building on the observations introduced in this section, we sketch a system where both selection and assembly of features are constrained parametrically.

V Constraining feature assembly

The points of contention between Lardiere and our position concern the nature of feature assembly into LIs in grammars: Is it constrained by UG and parameters? Or does this depend on the type of feature, i.e. interpretable vs. uninterpretable? A separate but related question is whether feature dis- and re-assembly in L2 acquisition is constrained similarly, or whether it takes place inductively through a contrastive analysis with corresponding L1 LIs, as Lardiere assumes.¹⁴ Lardiere does not state her position on the first issue, but it should be clear that the answer to the second depends on one's perspective on the first. Therefore, we start by sketching a plausible picture of how feature selection and assembly in L1 acquisition might be understood.

Cinque (1999) showed that although no language has all the attested functional categories associated with a given lexical category (verbs, in his case), when languages do possess them, their hierarchical order is cross-linguistically fixed.¹⁵ From this fact, he develops the strongest

¹⁴ We take it that Lardiere assumes that the features expressed in functional categories are based on a universal inventory provided by UG and that the selection of features from this set, at least in the case of L1 acquisition, is parametric, though she sees little role for this sense of parameter, especially in the case of L2 acquisition. Again, it is unclear to us whether she thinks that lexicalized parameters are useless in L1 acquisition as well.

¹⁵ This order is constant regardless of whether the heads are expressed as affixes or periphrastically, as Cinque (1999) shows.

possible hypothesis; namely, that the 60 or so verbal functional categories are universally available in all languages, but languages may differ in whether they possess an overt exponent of a given functional category. This is because the functional categories that are not overtly realized can be detected by the corresponding hierarchy of adverbs, which he assumes are licensed as Specifiers of various functional heads. It follows that when the head is unrealized, it is either filled by a null form, a moved lower head, or unfilled but contributes to the (default) interpretation.

There are other implementations of his basic insight that do not necessitate positing a plethora of null, or unrealized, heads. For instance, Williams (2003) posits an algorithm by which all languages start with the Cinque Hierarchy, but a contiguous sub-sequence of the hierarchy is mapped to the vocabulary items, yielding syncretic categories. Williams calls such LIs the ‘spanning vocabulary’ for the sub-sequence of the hierarchy. In another implementation, Grimshaw (2005) assumes that the hierarchy is universally available, but that languages may choose certain heads from the hierarchy, and not necessarily from contiguous regions of the hierarchy. However, the chosen heads are realized in the order determined by the Cinque Hierarchy. For example, Cinque proposes 66 possible verbal functional heads. But a language may choose to implement only three: F60, F22 and F8. If it does, it must respect the order of the universal categories for c-command and complement selection, such that F8 cannot select for F60. Grimshaw also allows for syncretic forms that realize multiple heads. Again, they are constrained in a similar way: the form that syncretically realizes the two heads must be realized in a position that respects the hierarchy. For example, following Bhatt and Yoon (1992), Grimshaw (2005) takes the English Complementizer *that* to conflate Mood and Subordination, while Korean (and Kashmiri) realize them separately. Korean *ko* introduces subordinate clauses and Korean *ta* introduces declarative mood. Mood and Subordination can be conflated in English because they are adjacent in the functional hierarchy. The form (*that, if, etc.*) that realizes the two syncretically is positioned above the exponents of lower functional heads (such as Tense), respecting the Cinque Hierarchy.

The features that are strictly ordered in the Cinque Hierarchy express functional features that are inherent to a given lexical category (or, those that form Extended Projections of the given lexical category, in the

parlance of Grimshaw, 2005). Other features may be interspersed in variable positions within this hierarchy. For example, optional features such as Negation can be realized in several places in the hierarchy, even within the same language, with predictable interpretive differences.¹⁶ And Agreement, though it is usually obligatory, can also be interspersed in various places in this hierarchy. One way to think of the various positions of Agreement is to adopt the idea in the Minimalist Program that it is a realization of uninterpretable (ϕ)-features that are lodged on another functional category. For example, if Tense and Aspect have ϕ -features, we can assume that they would be realized in the same hierarchical region where these heads are realized.

Though the work on the functional hierarchy associated with other lexical categories, such as Nouns, is not nearly as exhaustive as those on verbal functional categories, there is an emerging consensus that there is a fixed, universal, functional hierarchy for the other lexical categories as well. For instance, the work by Li (1999) that is cited by Lardiere is couched within this framework of assumptions. The (partial) hierarchy of (inherent) nominal functional categories may be Determiner–Number–Classifier–Noun, and languages may differ ‘parametrically’ in how vocabulary items are mapped to this hierarchy (Williams, 2003), or what heads are selected – and perhaps conflated with other selected heads – to receive expression (Grimshaw, 2005). Concretely, Shin (2007, following Krifka, 1995) proposes that the difference between Korean and English is not that all nouns in the former are mass (as claimed by Chierchia, a position that is problematic, as Lardiere shows), but that while English numerals conflate Number and Classifier (‘Object Unit’, in the words of Krifka, 1995), Korean numerals lexicalize just Number. Languages may also choose to seed certain heads that form the Extended Projection of nominals with uninterpretable features. Li’s (1999) account of the different loci of [+Plural] in the extended DP in Chinese and English is a specific implementation of this idea.

The point so far should be clear. There is a universal pool of features (perhaps better thought of in terms of already ‘pre-bundled’ functional heads that are arrayed into a strict hierarchy) that are available to all languages, as part of UG. Constructing a grammar of a language involves figuring out the form–function mappings of LIs specific to that

¹⁶ The so-called Short Form and Long Form Negations in Korean are cases in point. For languages that distinguish verbal and nominal negation, we may need to treat negation as an inherent, albeit, optional, verbal functional projection.

language. This process starts with the universal hierarchy and consists of the following ‘parametric’ choices:

- learning how the hierarchy is lexicalized, i.e. which interpretable features are chosen and how they are packaged into syncretic forms; and
- figuring out what uninterpretable formal features, if any, are added to the lexicalized hierarchy.

This process is not random, but heavily constrained, and it is more than reasonable to assume that the constraints are derivative of UG.

Having said this, we turn to the next question: Is L2 acquisition fundamentally different? Of course, one difference is the initial state, which is S_{L1} , rather than S_0 . We thus expect the knowledge of available features and their assembly into LIs in L1 to affect L2 acquisition (Full/Partial Transfer). But is this all there is to it? That is, does UG shut off and remain inaccessible or continue to guide and constrain the process (Full/Partial Access)? We have reasons to believe that UG is available and continues to guide and restrict the ‘resetting’ of parameters in L2 acquisition.

First of all, L2 learners do acquire features and functional categories that are not present in their L1. In Section IV, we noted that Korean learners of English whose L1 does not have the category of Determiners were able to learn not only that English has Determiners, but also the formal and semantic properties associated with English definite Determiners.¹⁷ The question is how do speakers whose L1 does not have the category Determiner:

- posit a novel category in the L2; and
- learn the features that typically cluster in this category, i.e. Quantificational properties (Definiteness, Specificity, Genericity, etc.), Number, Gender/Noun Class, etc.?

Lardiere does not address the first question but suggests that the answer to the second might come from the fact that the features in question may exist in the L1, though not clustered in the Determiner category, so that learners can pick the relevant features and assemble them inductively, item by item. We think that the inductive learning hypothesis predicts a far greater amount of variability than is actually found.

¹⁷ However, recall that it is not clear whether Korean has Determiners.

For example, while learners may have trouble attributing the right set of features to Determiners, they can learn whether or not a language has Determiners relatively easily. In fact, even when they misattribute certain features, they seem to be constrained by UG, as the work by Ionin and her colleagues has shown (Ionin *et al.*, 2004). For instance, instead of Definiteness, L2 Korean speakers use Specificity as the property distinguishing definite and indefinite Determiners. But this is exactly the dimension along which Determiners are split in languages (e.g. Samoan vs. English as described by Ionin *et al.*, 2004). This strongly suggests UG involvement in the L2 acquisition/assembly of features that are missing in L1.

What about when features exist in both L1 and L2 but are packaged differently? Lardiere's work, as well as the work by Ionin and Montrul (in preparation) mentioned earlier, suggests that this poses a greater challenge to the learner, implying a greater degree of inductive learning and L1 transfer. And yet, even here, we see that the process is guided and constrained. For example, languages regularly mark genericity on definite or bare plurals and do not use other options. English learners of Spanish generics were relatively successful in learning that plural definite Determiners can be interpreted generically. Suppose this is modelled as the addition of [+Generic] to the Determiner. How did English learners learn to associate this feature with the Determiner rather than something else in the sentence, as genericity is not overtly marked in the Determiner? Presumably learners arrive at this solution on the basis of their knowledge that quantificational properties are associated with Determiners or the noun, which is putatively UG-based, and this knowledge is what must have 'guided' them to this solution.

By contrast, we noted that Spanish learners of English faced greater challenges in learning the fact that English Determiners with plural Nouns do not admit a generic interpretation. This seems to suggest that the L1 assemblage of features poses a stumbling block to the learners and is consistent with L1-based inductive learning, a position advocated by Lardiere; see also Slabakova's (2006) Bottleneck Hypothesis. However, we believe that there is an additional factor involved in this situation which accounts for why it is easier to learn or add a feature than to unlearn or subtract a feature from a syncretic complex (a form in which there is no overt exponent of the added/subtracted feature).

The reason that adding a feature to a syncretic complex is easier than subtracting a feature from such a complex is that from a learnability perspective – and as several studies invoking the Subset Principle (Wexler and Manzini, 1987) have shown – unlearning is more difficult than learning, because if learning proceeds on the basis of positive evidence, it is easier to detect the existence of something in the input and infer its availability and correctness than to notice the absence of something and thus infer impossibility of a given interpretation.

In sum, evidence from the L2 acquisition of the existence and features of functional categories that are lacking in L1 suggests UG is involved in the selection and assembly of features of the acquired functional categories. The mis-attribution of features in L2 acquisition provides confirmation that UG is available even when learners fail to attribute the right features to the acquired category. Finally, though adding and removing features from syncretic forms is difficult, it is nevertheless possible, and the difference between feature addition and feature removal can be explained independently. We take facts such as these to be plausible first considerations in favour of a UG-constrained process of feature selection and (re-)assembly in L2 acquisition.

VI Conclusions

Lardiere's feature assembly approach to L2 acquisition is certainly provocative and controversial. Is this the end of Universal Grammar in second language acquisition? We do not think so. Although we have sketched several ways in which feature assembly can still be conceived as constrained by parameters, we realize that such a move requires more specific characterizations of how different types of features and their combinations relate to functional structure in different languages than those currently available, or at least a more in-depth consideration of how the available models can be applied to the L2 acquisition situation than in the sketchy exposition we have been able to lay out here. We leave this for further research.

Acknowledgements

We thank Tania Ionin and Roumyana Slabakova for insightful and stimulating discussion. All other errors and misinterpretations are our own.

VII References

- Beck, M.-L.** 1998: L2 acquisition and obligatory head movement: English-speaking learners of German and the local impairment hypothesis. *Studies in Second Language Acquisition* 20, 311–48.
- Bhatt, R. and Yoon, J.** 1992: On the composition of COMP and parameters of V2. In Bates, D., editor, *Proceedings of the Tenth West Coast Conference on Linguistics*. Stanford, CA: CSLI, 41–52.
- Bruhn de Garavito, J. and Valenzuela, E.** 2008: The L2 acquisition of *ser* and *estar*: a matter of aspect. *Bilingualism: Language and Cognition* 11, 323–36.
- Bruhn de Garavito, J. and White, L.** 2003: The L2 acquisition of Spanish DPs. The status of grammatical features. In Pérez-Leroux, A.T. and Licerias, J., editors, *The acquisition of Spanish morphosyntax: the L1/L2 connection*. Dordrecht: Kluwer, 151–76.
- Cinque, G.** 1999: *Adverbs and functional heads*. Oxford: Oxford University Press.
- Chomsky, N.** 1998: Minimalist inquiries: the framework. *MIT working papers in linguistics* 15, 1–56.
- 2001: Derivation by phase. In Kenstowicz, M., editor, *Ken Hale: a life in language*. Cambridge, MA: MIT Press, 1–52.
- Eubank, L.** 1996: Negation in early German–English interlanguage: more valueless features in the L2 initial state. *Second Language Research* 12, 73–106.
- Franceschina, F.** 2001: Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research* 17, 213–47.
- 2005: *Fossilized second language grammars: the acquisition of grammatical gender*. Amsterdam: John Benjamins.
- Giorgi, A. and Pianesi, F.** 1997: *Tense and aspect: from semantics to morphosyntax*. Oxford: Oxford University Press.
- Grimshaw, J.** 2005: Extended projection. In Grimshaw, J., *Words and structure*. Stanford, CA: CSLI.
- Hale, K.** 1983: Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1, 5–47.
- Hawkins, R. and Chan, C.** 1997: The partial accessibility of universal grammar in second language acquisition: the failed formal features hypothesis. *Second Language Research* 13, 187–226.
- Hawkins, R. and Franceschina, F.** 2004: Explaining the acquisition and non-acquisition of determiner–noun gender concord in French and Spanish. In Prévost, P. and Paradis, J., editors, *The acquisition of French in different contexts*. Amsterdam: John Benjamins, 175–205.

- Ionin, T., Ko, H. and Wexler, K.** 2004: Article semantics in L2-acquisition: the role of specificity. *Language Acquisition* 12: 3–69.
- Kayne, R.S.** 2005: Some notes on comparative syntax, with special reference to English and French. In Cinque, G. and Kayne, R.S., editors, *The Oxford handbook of comparative syntax*. New York: Oxford University Press, 3–69.
- Krifka, M.** 1995: Common nouns: a contrastive analysis of Chinese and English. In Carlson, G. and Pelletier, F., editors, *The generic book*. Chicago, IL: University of Chicago Press, 398–411.
- Lardiere, D.** 2008: Feature assembly in second language acquisition. In Licerias, J.M., Zobl, H. and Goodluck, H., editors, *The role of formal features in second language acquisition*. New York: Lawrence Erlbaum, 106–40.
- Li, Y.-H.A.** 1999: Plurality in a classifier language. *Journal of East Asian Linguistics* 8, 75–99.
- Montrul, S.** 2008: Form–meaning mappings in the aspectual domain: What about the L1? Commentary on Bruhn de Garavito and Valenzuela. *Bilingualism: Language and Cognition* 11, 337–39.
- Montrul, S. and Slabakova, R.** 2002: Acquiring morphosyntactic and semantic properties of aspectual tenses in L2 Spanish. In Pérez-Léroux, A.-T. and Licerias, J.M., editors, *The acquisition of Spanish morphosyntax: the L1/L2 connection*. Dordrecht: Kluwer, 131–49.
- Rizzi, L.** 1982: *Issues in Italian syntax*. Dordrecht: Foris.
- Sabourin, L., Stowe, L. and de Haan, G.** 2006: Transfer effects in learning a second language grammatical gender system. *Second Language Research* 22, 1–29.
- Schwartz, B. and Sprouse, R.** 1996: L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research* 12, 40–72.
- Shin, K.-Y.** 2007: Partitive structures. Unpublished PhD dissertation, University of Illinois, Urbana-Champaign, IL.
- Slabakova, R.** 2006: Is there a critical period for semantics? *Second Language Research* 22, 1–37.
- Slabakova, R. and Montrul, S.** 2003: Genericity and aspect in L2 acquisition. *Language Acquisition* 11, 165–96.
- Tsimplici, I.-M. and Dimitrakopoulou, M.** 2007: The interpretability hypothesis: evidence from *wh*-interrogatives in second language acquisition. *Second Language Research* 23, 215–42.
- Travis, L.** 2008: The role of features in syntactic theory and language variation. In Licerias, J.M., Zobl, H. and Goodluck, H., editors, *The role of formal features in second language acquisition*. New York: Lawrence Erlbaum, 22–47.

Wexler, K. and Manzini, R. 1987: Parameters and learnability in binding theory. In Roeper, T. and Williams, E., editors, *Parameter setting*. Dordrecht: Reidel, 41–76.

Williams, E. 2003: *Representation theory*. Cambridge, MA: MIT Press.