Accounting for Optionality in Nonnative Grammars: Parametric Change in Diachrony and L2 Development as Instances of Internalized Diglossia

Helmut Zobl and Juana M. Liceras
Carleton University and University of Ottawa

1. Introduction

This paper introduces the Competing Grammars Hypothesis (CGH) and considers its potential as a framework for explaining central aspects of L2 development, in particular the occurrence of optionality in the acquisition of new parametric options. The CGH was formulated in the context of work in diachronic linguistics, which was similarly faced with the task of explaining optionality during periods of parametric change (Kroch 1994, 2001; Pintzuk 1999; Pintzuk, Taylor and Warner 2002). The essence of the CHG, or internalized diglossia, is that in periods of important linguistic changes when input to learners is ambiguous, children internalize more than one grammatical representation to achieve coverage of the primary data.

We suggest that the CGH offers a framework which is able to account successfully for a significant number of characteristics of L2 development identified by research, such as the occurrence of optionality (Sorace 2000). The hypothesis is advanced as an alternative paradigm to the currently dominant one, which seeks to advance understanding of L2A in terms of how it is different from L1A (e.g., Hawkins 2001, Meisel 1997, 1998).

The following presentation of the CGH tries to demonstrate that parametric change in diachrony exhibits some of the same features found in L2A, features which a prominent L2 research direction interprets as reflecting some form of impairment to the adult language faculty due to critical period effects (e.g., Beck 1998, Hawkins & Chan 1997, Tsimpli 2003, Tsimpli and Roussou 1991). These common characteristics of change, we maintain, undermine the case for impairment. The body of the paper is devoted to illustrating parallels in the loss of OV, head-final I position, verb-second (V2) and verb-raising in the history of English and in the English L2 acquisition of Dutch, German and French speakers.

2. Parameter setting in L1A and L2A

2.1 Claims regarding L1/L2 differences

Research carried out in the UG framework has proposed a number of characteristics of parameter (re)setting in L2A which are thought to distinguish it from its counterpart in L1A. The one that has attracted the most theoretical interest is the prolonged period of optionality that occurs, for example, in acquiring/losing V2 and verb-raising (Beck 1998; Sorace 2000; White 1992). Its significance resides in it being at odds with the accepted view in generative research that grammars do not allow optionality (Chomsky 1995). L2 aspects other than optionality that have been cited are 1) gradualness of change versus its abruptness in L1A (Meisel 1998); 2) the incremental nature of change, such that different surface forms relating to a parametric option are acquired in a construction-specific manner, rather than in clustered fashion (Herschensohn 2000); 3) the lack of coherence in the changes (Liceras, Diaz and...
Maxwell 1996), or their unscheduled nature (Herschensohn 2000); and 4) the lack of convergence on the target language. We argue that the fact of diachronic change undermines (4) as a characteristic unique to L2A, while evidence to be presented from changes in Old/Middle English concurs with (1) and (2) and refutes the claim in (3).

Two broad types of explanations have emerged for these characteristics. The first type encompasses versions of what may be called the 'Impairment Hypothesis'. They have in common the view that some less than fully functioning language faculty underpins L2A in post-critical period learners, which makes it difficult to acquire formal features or values of the L2 different from the L1, and whose successful internalization is necessary for parameter (re)setting (e.g., Beck 1998; Clahsen and Muysken 1989, Hawkins and Chan 1997; Tsimpli 2003, Tsimpli and Roussou 1991). The second type of explanation views these characteristics as resulting from the restructuring that is required to move from one parameter option to another. Underspecification of the content of functional projections (either the feature content or their strength values) is one specific proposal about a necessary stage in restructuring (Eubank 1994, Herschensohn 2000), although it should be noted that this notion has also been employed in L1A to explain the occurrence of optionality (e.g., Deprez 1994). Underspecification more properly accounts for optionality while the incremental and slow occurrence of the change can be attributed to the nature of the restructuring process itself.

3. L1A and diachrony as paradigms for understanding L2A

3.1 The L1/L2 difference paradigm

A succinct statement in favor of the position that identification of the differences between L2A and L1A will lead to a better understanding of the former is due to Hawkins (2001). He compares the Difference Paradigm with the Poverty of the Stimulus Paradigm of recent L2A research. The latter views the two types of acquisition as fundamentally similar, with both drawing on UG inasmuch as properties of the target are acquired which are underdetermined by the primary data (and in L2A by knowledge of the L1). Hawkins' claim is that only the difference paradigm will lead to progress; the other paradigm will not advance our insights beyond what already appears firmly established, namely, that L2A does not lead to grammars incompatible with UG.

Yet Hawkins' claim in support of the Difference Paradigm can be challenged on the grounds that parametric change in L2A is by its very nature different from parameter setting in L1A. In the latter, the initial state is UG, and importantly, the child does not make a transition from parameter value P(a) to parameter value P(b). By current accounts, children do not misset parametric options (but see Verrips 1994 further on for a differing view). In L2A, the adult ideally should abandon the parameter value ‘transferred’ from the L1 when acquiring a new one. This consideration suggests that a better appreciation of L2A may be gained if we compare it with another context where change similarly implicates a transition from P(a) to P(b). Diachronic change represents this context.

3.2 The competing grammars hypothesis and parametric change
3.2.1 Similarities of change in diachrony and L2A

In section two we mentioned characteristics which are claimed to distinguish parametric change in L2A. In this connection it is instructive to glance at Lightfoot's (1979) account of important syntactic changes in Middle English surrounding the modal verbs. Perhaps not coincidentally, he identifies the very same properties for historical change which research attributes to parameter-setting in L1A, namely that it occurs abruptly and in clustered fashion. Most importantly, Lightfoot argues that English speakers at that time tolerated a prolonged period of input ambiguity by adding exceptions to their grammars, until, finally, their cumulative weight forced a restructuring.
Lightfoot's view of syntactic change has been challenged by a number of historical linguists on two counts (e.g., Allen 1995; Kroch 1994; Pintzuk, Tsoulas and Warner 2002, Warner 1997). First, did speakers really cope with the ambiguity simply by adding exceptions to their grammar? Second, rather than being abrupt and clustered, parametric change is often slow and protracted, spanning several centuries. It is incremental and characterized by optionality between mutually exclusive options. Since the CGH adheres to the axiom that grammars do not tolerate optionality, it accounts for its presence in the historical data (e.g., in the writings of one and the same scribe or author), by proposing that speakers must have internalized more than one grammar: "Syntactic change proceeds via competition between grammatically incompatible options which substitute for one another in usage." (Kroch 1994: 180) The grammars compete against each other until one becomes preferred, due to factors such as greater success in parsing the primary data (cf. Clark & Roberts 1993) or easier processability (Sprouse and Vance 1999). For Kroch (1994), the implementation of the innovative grammar in its several surface manifestations is conditioned by performance factors, although the rate of progress should be constant across the several contexts.

3.2.2 The CGH, its antecedents and congeneres

It should be noted that linguists and researchers in L1A and in L2A have put forward proposals to account for variability which bear strong resemblance to the CGH. Hankamer (1977) argues that "...in the face of confusing and conflicting data... [the child] constructs conflicting analyses, and does not necessarily ever choose between them." (p.584) For L1A, Verrips (1994) introduces an acquisition procedure which she terms 'Maximize", suggesting that children operate with several analyses of the input data until they eventually settle on one. The variability in English children's suppliance of subject and object pronouns is explained as resulting from the temporary coexistence of three grammars -- a Chinese style topic-drop grammar, an Italian style [+pro-drop] grammar and the English [-pro-drop] grammar. In L2A, Montrul (1998), Robertson and Sorace (1999) and Trahey and White (1992), for example, have similarly suggested that parametric variability may arise from coexisting conflicting grammatical representations.

In summary, we argue that the CGH should be taken seriously as a framework for L2A research. It provides an alternative to the Difference Paradigm, in particular to the various versions of the Impairment Hypothesis. The common characteristics of parametric change in diachrony and in L2A, in particular optionality, incrementalism and slowness, to mention but three, throw into question the Impairment Hypothesis. The CGH also accounts for a characteristic of L2A which Selinker (1972) thought was definitional for it: backsliding. Under the CGH, the recessive grammar need never be completely erased and may still be accessed under conditions not entirely clear long after it appears to have been supplanted.

4. Some instances of parametric change in diachrony and L2A

In section four we consider four parametric changes in L2A and in the history of English, namely OV to VO, I-final to I-medial, the loss of V2 and the loss of verb-raising. The emphasis will be on demonstrating that L2A displays very similar changes both in terms of sequencing and the constellation of structural factors.

Sentences 1-4 illustrate the relevant parametric contrasts we deal with. We have recourse to ungrammatical Modern English sentences to illustrate the parametric option that English lost in earlier historical periods.

1. OV/VO
   *the strawberries ate/ate the strawberries
2. I-final/I-medial
   *...that John the strawberries eaten has/...that John has eaten the strawberries
3. [+/-V2]
   *Yesterday ate John the strawberries/ Yesterday John ate the strawberries
4. [+/- verb raising]
   * John eats not the strawberries/ John doesn't eat the strawberries
Lastly, in Figure 1 we provide a simplified tree structure containing the functional projections that are pertinent for the parametric changes.

Figure 1. Functional projections implicated in the parametric options

![Tree Structure Diagram]

4.1 *OV to VO, I-final to I-medial, and the loss of V2*

On one view, Old English was in many respects like contemporary German and Dutch, having OV word order, I-final and V2 (van Kemenade 1987). Pintzuk (1993, 1999) disputes this, claiming that OE already had an I-medial and a VO grammar in competition with I-final and OV. By around 1250, word order was VO, I was medial and V2 was beginning to wane (Kroch 2001). Thus word order and I-position changed first; in fact, Pintzuk (1999) argues for a constraint which bars the combination of VO and I-final, although the converse, OV and I-medial, is possible. This being the case, the innovative grammar with VO implicated the presence of I-medial.

As for the loss of V2, differences between northern and southern dialects must be acknowledged, with the northern dialects (influenced by Scandinavian languages) being almost categorically V2. For these it is assumed that the verb moved to C and preposed elements to Spec,CP. In the southern dialects the picture is more complex. Besides Pintzuk (1993), several researchers suggest that V2 in Middle English was to the I position, but that this IP-V2 grammar was vulnerable (Kroch and Taylor 1997). Adverbs in particular lent themselves to an adjunction reanalysis and triggered V2 less and less often; but even fronted objects failed to trigger V2 in about 20% of cases (Haeberli 2002). Yet movement to C also existed, triggered by operator elements in C, e.g., Q, NEG, and Focus. The latter feature, Focus, would explain why argument preposing more consistently triggered V2 while adverbs did not (Kroch 2001); of course, even in contemporary English, (residual) V2 (or I to C) (Rizzi 1996) occurs with Wh-elements and the presence of a NEG feature.

While this cursory exposition does not do justice to the complexity of V2 in Middle English (we also ignore here the absence of V2 with subject pronouns in the southern dialects), it appears that one can speak of a CP-V2 grammar and an IP-V2 grammar, with many speakers having internalized both.

We turn now to two L2A studies which are able to provide a comparison, beginning with Hulk (1991). She investigated the L2 acquisition of French by Dutch speakers in a cross-sectional study employing acceptability judgements. Table 1 lists seven sentences which tested subjects’ acceptance of VO (#25,7) versus OV (#1, 4,), and V2, both with thematic (#3) and auxiliary verbs (#4,5). The Dutch learners need to internalize the French parameter values VO, I-medial and [-V2].

Table 1. Percentage of acceptance (based on Hulk 1991)

<table>
<thead>
<tr>
<th>Sentences</th>
<th>Years of French Instruction</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>Jean a les fraises mangé</em></td>
<td>73</td>
<td>40</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2. Jean a mangé les fraises</td>
<td>42</td>
<td>86</td>
<td>100</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>3. <em>Hier mangeait Jean les fraises</em></td>
<td>92</td>
<td>50</td>
<td>32</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>4. <em>Hier a Jean les fraises mangé</em></td>
<td>92</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
5. *Hier a Jean mangé les fraises 38 64 8 6
6. Hier Jean mangeait les fraises 38 80 100 100
7. Hier Jean a mangé les fraises 19 85 100 100

Looking at Year 1, we see that subjects predominantly have a [+V2] grammar (92% for #’s 3 and 4), but a sizeable number of subjects also already accept [-V2] (38% and 19% for #’s 6 and 7, respectively). Also, in Year 1, OV (73% for #1) and VO (42% for #2) coexist. By Year 2, the VO grammar is supplanting OV in the judgements (86% versus 40%) and the [+V2] grammar is less preferred than the [-V2], although sentence #5 with 64% acceptance of V2 and VO is the strongest indication that both options coexist in the L2 grammars of an indeterminate but sizeable number of learners. By year 3, their grammar is essentially targetlike, except for the 32% and 8% acceptance of [+V2] in #’s 3 and 5, respectively. V2 remains marginally acceptable into the university year(s).

Sentence (5) is important for another reason. The test sentences did not employ embedded clauses, thus there is no way of directly comparing judgements on the I-medial versus I-final contrast. However, since acceptance of VO implicates a constraint that excludes I-final, we suggest that acceptance of this sentence rests on a grammar having V2, VO and head-medial. This grammar essentially mirrors an Old/ Early Middle English grammar of the southern dialects. Since (5) has a temporal adverb, ‘hier’ (yesterday), which is triggering inversion, whereas this class of adverbs typically failed to trigger V2 in southern dialects of Early Middle English, it is possible that acceptance of (5) rests on a Dutch-type CP-V2 grammar in which such adverbs more reliably attract the verb to C. Alternatively, the Dutch speakers have formulated an IP-V2 grammar on their way to losing V2 (although this type of V2 apparently occurs in root and embedded sentences (Kroch and Taylor 1997). Historically, the IP-V2 grammar exhibited considerable variability with respect to V2, much like the judgements of the Dutch speakers in the first two years.

In summary, the progression we witness here shows that parametric change in L2A is characterized by competition between options, particularly [+/-V2] and VO/OV. The sequencing of the changes appears to be the same as the one from Old English to Middle English, namely, loss of I-final and OV followed by loss of V2. This speaks against the claim that parameter setting in L2A is unscheduled (see Herschensohn 2000).

Next, we consider an investigation of German speakers acquiring English (Robertson & Sorace 1999). The findings of this investigation provide some suggestive evidence that German learners innovate an IP-V2 grammar on their way to losing V2.

The study employs preference ratings. In Table 2, the learner groups from left to right range from 3rd year of English exposure/instruction (G 08) to 12th year (UN4). In contrast to Hulk (1991), who used only the adverb ‘hier’, the sentences employed here varied the preposed constituent to include PP’s, object DP’s and adverbials.

Table 2. Preference ratings for [+/-V2] by German learners of English (based on Robertson and Sorace, 1999: 333)

<table>
<thead>
<tr>
<th>Years of English exposure instruction</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*V2 &gt; V3 n</td>
<td>12</td>
<td>8</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>%</td>
<td>18.2</td>
<td>11.6</td>
<td>11.1</td>
<td>4.8</td>
<td>4.1</td>
</tr>
<tr>
<td>* V2 ~ V3 n</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>%</td>
<td>9.1</td>
<td>8.7</td>
<td>3.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*V2 &lt; V3 n</td>
<td>48</td>
<td>55</td>
<td>69</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>%</td>
<td>72.7</td>
<td>80.0</td>
<td>85.2</td>
<td>95.2</td>
<td>96.0</td>
</tr>
</tbody>
</table>
Considering the preference ratings for [+V2] versus V3 (i.e. [-V2]), we see that [+V2] is clearly recessive by the third year, with 18.2% of subjects preferring V2 over V3, in contrast with 72.7% of subjects who prefer V3 over V2. Yet Robertson and Sorace point out that the majority of their subjects do not express judgements consistently in conformity with one grammar or the other; in fact, totally consistent judgements are the exception. As for the acceptance of V2 over V3 in relation to the type of preposed constituent, they report that after the early stage of exposure, preposed arguments do not trigger V2. The greatest number of acceptances of V2 over V3 occurred with what they refer to as ‘discourse adverbials’ (e.g., ‘because of this…’). As in the Hulk study, we can only speculate at this stage of our research as to the possibility that the German CP-V2 grammar is crowded out by an IP-V2 grammar early on, which then coexists in recessive fashion with an English type grammar not permitting inversion. This possibility acquires support from studies by DuPlessis et al. (1987) and Schwartz and Sprouse (1995) which, although they looked at the acquisition of V2, similarly suggest that an IP-V2 grammar may develop in which adverbs adjoin to IP or CP. It is noteworthy that not only do fronted objects not trigger V2, but even the presence of the feature NEG with negative adverbs fails to attract the finite verb to C in the preference judgements up to advanced levels. This may indicate that one of the L2 grammars is very much an English type residual V2 grammar, but in which only the feature Q has the ability to trigger V to C.

To summarize, this study, too, provides overwhelming evidence for coexisting representations, with judgements consistent with internalized diglossia forming the dominant pattern. Historically, V2 persisted longer and more reliably with preposed arguments, whereas in this study, topicalized objects very quickly lost their V2 triggering ability. This difference calls out for future investigation.

4.2 Loss of verb-raising

The final comparison concerns the loss of verb-raising with thematic verbs. Specifically, we will focus on White's (1990/91, 1992) work with francophone learners acquiring English. Her research focusses on whether French speakers 'transfer' thematic verb raising from French to English and produce sentences in which the verb moves to I in declaratives, and on to C in questions. If they do, sentences such as ‘eats she?’, ‘she likes not this sport’, and ‘she likes very much fresh strawberries’ will be judged acceptable.

Before turning to White's results, a brief exposition of the historical change is in order. Ellegard's (1953) figures concerning the emergence and spread of auxiliary DO in several contexts provide information about the loss of verb-raising. With the emergence of DO and the parallel reanalysis of modals as I-elements (Roberts 1993; Warner 1997), thematic verb-raising to I began to wane. Ellegard's figures indicate that the spread of DO was uneven, being progressive in question contexts and lagging behind in negative contexts. DO in negative imperatives was only beginning to emerge at a time when it was approaching 50% frequency in questions and negation (cf. Table 3).

Table 3. The spread of DO in Middle and Early Modern English (%)
(based on Ellegard 1953)

<table>
<thead>
<tr>
<th></th>
<th>Year 1400</th>
<th>1450</th>
<th>1500</th>
<th>1550</th>
<th>1600</th>
<th>1650</th>
<th>1700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg question</td>
<td>10</td>
<td>35</td>
<td>80</td>
<td>75</td>
<td>90</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Affirmative question</td>
<td>5</td>
<td>10</td>
<td>45</td>
<td>65</td>
<td>80</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Neg declarative</td>
<td>2</td>
<td>5</td>
<td>30</td>
<td>30</td>
<td>40</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Neg imperative</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>20</td>
<td>35</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Two recent studies have suggested that the differing frequencies reflect competition between a verb-raising and a non-verb raising grammar (Han 2000; Warner 1997). Han's proposal goes a step further, claiming that the loss of verb-raising actually implicated three distinct but overlapping subgrammars which differed from one another in terms of the distance or height, in the tree, of the functional projections that the verb was able to move to. The subgrammar permitting thematic verbs to raise from Tense or I to C was
the first to become recessive, reflected in both the high and early frequencies of DO in question contexts. Next, the grammar permitting movement from V to Tense/I became recessive, so that the verb was no longer able to move past negation. This is reflected in the spread of DO in negative declarative sentences in advance of negative imperatives. The late appearance of DO in the latter context is accounted for by assuming that imperative sentences have no Tense projection, but that the verb was still able to move to an Aspect projection, located above VP and above a second, lower, NEG position (refer to Figure 1). In this way, negative imperatives of the type ‘fear you not’ remained possible. The subgrammar permitting movement to the Aspect projection was the last to disappear. Thus, the loss of verb movement across the several contexts is seen as a stepwise reduction in the ability of the verb to move to higher functional projections.

The preference judgements of the francophone speakers in White's (1992) investigation can be summarized verbally instead of in tabular form. There is overwhelming preference for no verb raising in questions and in negative contexts. The puzzling feature of her results concerned the continued high level of acceptance of sentences such as ‘she likes very much fresh strawberries’, in which the verb has moved out of the VP past an adjoined Adverb Phrase (V Adv. O) to a functional projection, which White assumed was Agreement Object Phrase (AGROP) (i.e., her 'short movement'). While current generative theory no longer adopts an AGROP for English, an Aspect Phrase appears well motivated. Given the close connection between Aspect and the lexical semantics of the verb, an Aspect projection should be low in the tree, above the VP and the adjunction site for many types of adverb phrases.

This proposal for short movement links up with Han's (2000) historical analysis. It also links up with a recent investigation by Chu and Schwartz (2004), which finds that Chinese-speaking learners also accept thematic verb movement past adverb phrases. V Adv O order was accepted at a mean rate of 29%, lower nonetheless than the acceptance rate of White's francophone subjects.

In summary, the historical data and the L2A data from francophone and Chinese speakers lend support to White's (1992) proposal for ‘short verb-raising’. Again, as earlier, we find that the sequencing of the changes in L2A is scheduled, mirroring that of the historical data. Grammars appear to lose verb-raising in a top-down manner along the hierarchy of functional projections. This pattern for verb movement to be reduced, or shortened, and lost is observed by Rivero (1996, 1997) to have occurred historically in typologically diverse languages. She relates the loss to the principle that movement represents a last resort, a view echoed by Roberts (2001) for whom movement always results in more complex representations.

5. Conclusion

The evidence reviewed here from diachrony and L2A speaks strongly for the existence of internalized diglossia. The CGH accommodates optionality without abandoning the generative axiom that grammars do not tolerate optionality. It puts into question the repeated claim that optionality arises from some form of impairment to the language faculty.

In the loss of V2 and verb-raising we have evidence that the change is incremental, not clustered and abrupt. Also, the sequencing of parametric changes very closely matches those in diachrony. This speaks against the claim that parametric change in L2A is unscheduled. In connection with verb-raising we have suggested that its loss implicates coexisting subgrammars that differ in the height of the functional projection to which the verb may raise. The loss, both in diachrony and in L2A, proceeds in a downward fashion. With regard to the loss of V2 by German and Dutch speakers, we have some - albeit only suggestive - evidence that, as in Middle English, some type of IP-V2 grammar may be in competition with, first, a CP-V2 and, subsequently, a [-V2] grammar in the course of the parametric change. This being the case, the loss of V2 also proceeds down the hierarchy of functional projections.

The CGH has other broad ramifications for our understanding of various phenomena associated with L2A. For one, the notion of transfer will need to be reconceptualized, as will the role that restructuring is made to shoulder in current descriptions of how parametric change occurs. Lastly, the CGH will hopefully inspire a reinterpretation of many other findings in L2 research so that a comparison with the L1/L2 Difference Paradigm can be carried out on a broad basis.

References
