Near-nativeness, Optionality and L1 Attrition

Antonella Sorace
University of Edinburgh

1. Introduction
One of the central questions in contemporary adult second language acquisition (L2A) theory is whether, and to what extent, Universal Grammar (UG) still constrains the acquisition process: given that adults can rely on general cognitive abilities, it is at least conceivable that they may use them, instead of UG, in the task of learning a second language, particularly if UG, for neurological reasons, ceases to operate after a certain age. Indeed, this may appear intuitively plausible, given two obvious differences between first and second language acquisition: first, the adult learner already knows (at least) one other language; the initial state of the child and of the adult are not the same (e.g. Schwartz and Eubank 1996, Schwartz and Sprouse 1994); second, unlike children, who reach perfect mastery of whatever language they are exposed to, many adults after long periods of exposure to a second language display varying degrees of “imperfection” (by monolingual native standards), and even those who are capable of native-like performance often have knowledge representations that differ considerably from those of native speakers (Sorace, 1993). So not only the initial state, also the final state of the child and of the adult learner are different.

But how different can the final state be? Research specifically focused on ultimate attainment can tell us what kind of ‘steady state’ can be reached in non-primary language acquisition, whether such state is quantitatively/qualitatively different from the monolingual steady state, and whether it is UG-constrained. The characteristics of final states are, in a sense, even more revealing of UG constraints on L2 acquisition than those of other stages. If it is found, for example, that intermediate grammars appear to violate UG, the argument can always be made that, given more input, or more time, or a better learning environment, the non-native grammar may in due course converge on the target. However, if adult learners have become virtually undistinguishable from native speakers, and continue to benefit from full exposure to the L2, they can be assumed to have progressed to the furthest attainable competence level: any differences between their grammar and the L1 grammar may be considered as final, and any difference that embodies a UG violation may likewise be regarded as a permanent feature of the final state.

It is important to realize that ‘near-nativeness’ is not (only) a slightly imperfect version of ‘nativeness’. This point is often overlooked in existing definition of the phenomenon. One definition, for example, reads as follows: “A late bilingual who has a wide vocabulary comparable to that of an educated native speaker, and whose proficiency is
potentially native, i.e. according to informed native-speaker opinion sufficient to maintain a fully native-like level of production for some length of time at least under certain favourable conditions. A slight foreign accent in pronunciation and intonation will be allowed" (Ringbom 1993: 302). Like most existing definitions, this one presupposes: (a) an accepted definition of the native speaker as a fully competent language user in most contexts; (b) an "additive" view of bilingualism, according to which a bilingual is the sum of two monolinguals; (c) the deficit hypothesis: the near-native comes very close to native-like competence, but falls short of it.

These presuppositions are in turn due to the fact that most research on adult bilingualism adopts as points of reference the monolingual native speaker and L1 acquisition by monolinguals. In addition, knowledge of the L1 is regarded as a factor that can have a determinant influence on L2 attainment but is itself unaffected by it. The evidence from second language acquisition research is rather contradictory. On the one hand, within the dominant additive view adult second language acquisition is regarded as (a) incomplete, (b) variable, and (c) non-equipotential in comparison with L1 acquisition (see Bley-Vroman 1990). We know that there are some age-related effects: generally speaking, starting young seems to confer some advantages, although it is by no means clear what the reasons are (Johnson and Newport 1989; Long 1990). On the other hand, we also know that there is a 'logical problem' of second language acquisition; that properties that are not instantiated in the L1 and are not explicitly taught may be successfully acquired; that interlanguages are, by and large, natural languages (i.e. they present no violations of Universal Grammar), although they may also be shaped by other cognitive principles. It will be suggested in this paper that the additive view at best accounts for a small minority of non-native speakers, those who succeed in acquiring native competence and at the same time maintain their original native competence in the L1. The majority of non-native speakers, on the other hand, develop a competence - in both the L2 and the L1 - which diverges, often in non-obvious ways, from the monolingual native's.

2. Perspectives on near-nativeness and empirical evidence
Empirical studies of near-nativeness to date have focused on the 'completeness vs incompleteness' issue. The reasoning guiding this research has been that if near-natives have the same knowledge as that exhibited by natives, the existence of UG constraints on L2 acquisition is confirmed; if, on the other hand, near-natives do not possess this knowledge, their competence is missing particular properties because it is not UG-constrained. This line of argument is fundamentally flawed because it does not consider the possibility of other UG-constrained final outcomes. Furthermore, the assessment of near-nativeness has targeted isolated grammatical properties, rather than clusters of properties related to particular parameters. As Neeleman & Weerman (1997) maintain,
success in the acquisition of construction-specific parameters is compatible both with a general learning strategies model and with a UG model, thus making it difficult to distinguish between the two. Only success in the acquisition of clusters of properties related to parameters would be unambiguously due to UG constraints. More generally, these studies support a variety of (often contradictory) conclusions, reflecting a lack of consensus over what counts as 'UG-constrained behaviour' and ultimately about what counts as 'near-native'. The following brief summary of the main studies carried out to date will highlight this muti-plicity of views in the field.

The pioneering study by Coppierter (1987) tested near-native speakers of French from a variety of language backgrounds. The variables investigated included both semantic contrasts (e.g. tense/aspect distinctions) and syntactic conditions (cliticization and raising). The method used was a grammaticality judgment test. The results indicate significant differences between native and non-native norms, less marked for syntactic than for semantic variables. Inspiring as it was, Coppierter's study was criticized on methodological grounds, particularly because of the impressionistic criteria he employed in the selection of near-native speakers.

Birdsone's (1992) study is a methodological more rigorous replication of Coppierter's study. It also tested near-native speakers of French on various syntactic and semantic properties, including some of those investigated by Coppierter. Methods used ranged from grammaticality judgments to think-aloud procedures. Birdsone's results point to the opposite conclusion from Coppierter: there are no significant differences between native and non-native norms in a number of individual near-natives, although such differences are there when natives and near-natives are compared as groups.

White and Geneseed (1996) looked at 42 near-native speakers of English from various language backgrounds. To counterbalance the frequent criticism of arbitrariness and subjectiveness in the criteria used to identify near-native speakers, they used more rigorous selection procedures (interviews, psychological tests, ratings by two native speakers). The area of grammar tested was subjacency. The tests employed included grammaticality judgments and question formation, and reaction times were also recorded. Their results support Birdsone's conclusion: no differences were found between the performance of natives and that of near-natives, but reaction times were shorter for native speakers. White and Geneseed's suggestion is that the near-native competence is (or can be) the same as the native competence.

An orthodox experimental psycholinguistic approach is taken by Kilborn (1992), who investigated German near-native speakers of English. His experiments focused on the on-line integration of different types of grammatical information in comprehension. Word-monitoring tasks were used in normal and noise conditions. Kilborn's results show that the performance of near-natives in normal listening conditions is
similar to that of natives in noise conditions, indicating a failure to integrate syntactic and semantic information as rapidly as monolingual L1 speakers.

Johnson et al. (1996) questioned the intuition by Sorace (1988) that replications of the same test with L2 learners is unreliable because of possible interim competence developments. They tested Chinese very advanced speakers of English in experiments aimed to assess the degree of consistency between two successive administrations of the same (aural) acceptability judgment test. The results suggest that the natives are consistent, but the non-natives are not; their knowledge is indeterminate. However, these findings may be been affected by the uneven proficiency level of the learners tested (whose average length of residence in the US ranged from 5 to 12 years), and by the choice of an aural acceptability judgment test which, because of its demands on on-line comprehension, may have imposed an additional burden on the subjects, creating a confounding effect.

In contrast with the other studies, Sorace (1993) demonstrates that final states may not only be incomplete, but also complete and systematically different from the target. Her study targeted English and French near-native speakers of Italian, testing knowledge of (a) auxiliary choice with different semantic types of unaccusative verbs, and (b) auxiliary choice in syntactic phenomena related to restructuring (i.e. change of auxiliary, clitic-climbing). Grammaticality judgments were elicited by means of magnitude-estimation techniques (Sorace 1996; Bard, Robertson and Sorace 1996). The results point not only to significant differences between natives and near-natives, but also to significant differences between English and French near-natives. Specifically, the English near-natives display incompleteness, that is, the absence of properties required by the L2, whereas the French near-natives exhibit divergence, namely representations of L2 properties that are consistently different from native representations. Both incomplete and divergent representations are affected by the L1 grammar, and both are UG-constrained (see White 1996).

The interim conclusion to be drawn from this brief overview of research on near-nativeness is that an overall state of competence identical to that of monolingual speakers is difficult to attain in adult second language acquisition. However, what looks like incompleteness may at closer scrutiny turn out to be systematic divergence. The empirical question facing L2 research is what exactly constitutes divergence; what forms divergence can take; and which of these forms can or cannot be part of the make-up of a natural language grammar.

3. A different perspective on near-nativeness: optionality

One type of divergence that has emerged from recent research on L2 final states is optionality. Optionality can be defined as the co-existence within an individual grammar of two or more variants (otherwise identical
at LF, or making use of the same lexical resources) of a given construction. It is well-known that optionality exists both in mature and in developing child grammars. The question is whether 'stable' and 'developmental' optionality are the same phenomena; a related question, of central concern to this paper, is whether second language developmental optionality is a phenomenon of a different nature from that of developmental optionality in a first language. In what follows I will discuss some of the empirical and conceptual issues raised by optionality. First examining stable optionality and then L1 developmental optionality. I will then turn to an examination of second language optionality.

3.1 Optionality in mature grammars

The existence of (stable) optionality in native grammars has been amply documented. Examples analysed in the literature are scrambling in West Germanic; multiple WH-movement in Hungarian; WH-questions in French; auxiliary alternations under Restructuring in Italian; singular concord in Belfast English; phonologically overt agreement with object shift in French, among others. Optionality poses a conceptual challenge to formal grammatical theory. Contemporary formal models of grammars (e.g. Minimalism, Optimality Theory) are in fact essentially comparative: they assume a set of candidates competing for well-formedness and an evaluation metric based on economy principles. Only one candidate in each set emerges as the 'optimal' winner: the others are assigned no grammatical status. Optionality, however, involves the co-existence of two (or more) optimal forms, one of which is usually 'more grammatical' than the other, in a sense to be made precise. The question is whether optionality can be internal to an individual grammar, or whether it is entirely dependent on external factors. Most of the solutions proposed so far either aim to demonstrate that 'real' optionality does not exist; or that optionality is to be relegated to domains other than the grammar proper, or that there is limited scope for optionality in terms of 'ties' within the grammar, with any differences in frequency, acceptability etc. to be ascribed to factors external to the grammar.

Accommodating optionality within the grammar is feasible, but only in a restricted set of circumstances. All theories of generative grammar make some restricted allowance for optionality when it is costless for the grammar, i.e. when the alternatives are equal in terms of economy of derivation. This is the way Fukui (1993), for example, analyses certain types of movement, such as rightward movement in English and scrambling in Japanese.

Within Optimality Theory (see section 4 for more details), constraints may be either ranked freely, in which case each variant is the outcome of fixing the free ranking in a particular way; or they can be tied, in which case violations on each constraint count equally in candidate evaluation. It is worth noticing that the OT solutions, like the orthodox generative grammar solution, provide no insight into the relative frequency of the
variants. One possible way out of this problem would be to introduce probabilities into the grammar, which would obviously affect some of the current basic foundations of formal grammatical theories: recently, some modifications of Optimality Theory have been proposed which essentially adopt this solution (e.g. Boersma, 1997). There are recent proposals within a Minimalist framework (e.g. Pettitward 1996) that allow for optionality within the grammar, by associating it not with the possibility of occurrence, but rather with the timing of movement.

At the present stage of the theoretical debate, it would seem to be easier to have 'optionality-free' grammars. Given that optionality is dependent on whether two or more alternatives are perfectly equivalent in terms of meaning, it is possible to show that optionality is more apparent than real (Grimshaw and Samek-Lodovici 1995; Adger 1996) if there are constraints on the distribution of optional forms, or if optional forms involve different semantic representations, and thus are not in competition with one another. A well-known alternative solution, which still expunges optionality from within the grammar, is the 'double base' hypothesis (Kroch 1989), according to which more than one grammar may underlie a single language. Under this scenario, optional forms belong to different grammars, therefore optionality, as a visible manifestation of a state of diglossia, is not internal to the grammar.

3.2 Developmental optionality

Let us now turn to developmental optionality, which is widely attested in both child L1 grammars and adult L2 interlanguage grammars.

In first language acquisition, children go through stages of temporary optionality which allow for the co-existence of forms that are mutually incompatible in adult grammars. Optionality is attributed to different causes, depending on the theory of grammatical development adopted. Within the 'structure-building' account proposed by e.g. Radford (1996), optionality arises because of maturational constraints that delay the appearance of functional categories in child grammars until the third year of age: during periods of transition between stages (i.e. from a VP grammar to an IP grammar, or from an IP grammar to a CP grammar), the child will often exhibit features of both the old and the new grammar, alternating between the two until the new grammar becomes categorical.

A child at the age of 2.9 may produce both the sentence types in (1):

(1) a. May I hang this up?
    b. Where candles go?
    (examples from Radford 1995)

An alternative view of development (often named the 'Continuity approach') assumes that whole phrase structure characterizes child grammars from the beginning, but features may be temporarily underspecified: it is this underspecification that generates optionality (Wexler 1994; Hyams 1996). However, the term "underspecification"
has not received a uniform interpretation in the literature. For Wexler (1994; to appear), it means the optional absence of a particular functional head (Tense or Agr), which may itself be caused by the presence of a developmental constraint in child grammars. Underspecification in this sense is viewed as responsible for the alternation between finite and non-finite verbal forms that characterizes child grammars in a well-represented set of languages, including English, Dutch, German and French. In a similar vein, Rizzi's (1994) "Truncation Hypothesis" assumes that child grammars may lack the principle "CP = root," so that the starting point of the child's projection is sometimes VP, sometimes IP, and sometimes CP. Hyams (1996), on the other hand, argues that functional nodes, both in the clausal and nominal domains, may be underspecified in the sense of "unindexed"—not part of syntactic chains that anchor the event or the referent. In recent work, Hoekstra and Hyams (1996, to appear) develop this analysis further, arguing that this lack of anchoring can be traced back specifically to the underspecification of the Number feature. Under these underspecification scenarios, the child can, for example, optionally project either Agr or Tense; optionally choose CP as the root node (Rizzi, 1994); or optionally project Number. What is left unexplained, however, is the observed pattern of systematic changes in the preferences for one option over the other in the course of development. Whatever its cause, children in due course abandon optionality and retain the option allowed by the target grammar (unless the optionality is in the target grammar, in which case children acquire both options and the same pattern of distribution: see Henry 1997). This process involves a gradual decrease in the use of the non-target option and a gradual increase in the use of the target option. For example, the abandonment of optionality of Tense or Agr involves the gradual increase in the proportion of finite sentences. Logically, the initial stage of the process should present a predominance (or even the exclusive use) of non-finite forms: this is problematic to detect in very young infants, although it may become possible to detect by recently developed techniques for the study of infants in the prelinguistic stage (see Jusczyk 1997).

In second language acquisition, learners also go through stages characterized by optionality (Eubank 1994; Robertson and Sorace, to appear). However, there are two main differences between L1 and L2 optionality. The first is that compared to L1 acquirers, L2 learners have the L1 as an additional source of optionality. The second is that, unlike L1 optionality, L2 optionality tends to persist until advanced competence levels. Most examples of optionality discussed in the literature to date are related to optional verb movement. A substantial body of research has tried to provide an explanation for the long-lived alternations in adverb placement that characterize the second language production of French learners of English (White 1990/91, 1992; Eubank 1996; Schwartz 1996; Schwartz and Gubala-Ryzak 1992, etc.).
(2) a. Mary speaks very well English
   b. Mary speaks English very well

Robertson and Sorace (to appear) show that advanced German learners of English retain a residual V2 constraint, occasionally producing sentences like (3a).

(3) a. For many kids is living with their parents a nightmare
   b. For many kids living with their parents is a nightmare

Eubank (1994, 1996) discusses the optionality in the placement of negation in German-English interlanguage. More recently, Eubank and Grace (1996) analyse optionality of verb movement in the interlanguages of Chinese learners of English: this is a different case from the ones previously discussed in that neither the L1 nor the L2 instantiate verb raising, and poses the question of what causes raising in these interlanguage grammars. Eubank's account of optionality for both adverb placement and negation placement involves the underspecification of feature values associated with functional heads: his 'valueless features' hypothesis assumes that while functional categories themselves transfer from the L1, such features are 'inert' until the learner, on the basis of exposure to the L2 morphological paradigm, acquires the value instantiated by the target language. The problem with this approach is that it predicts a sudden demise of optionality once the inflectional paradigm is in place, contrary to fact.

Beck (1997) explains the persistence of optionality by assuming a 'deficit' in the grammatical representations of L2 learners, whereby the connection between morphology and syntactic movement is severed. Lardiere (to appear), Schwartz (1996), and Sprouse (to appear), however, question the theoretical presupposition of a developmental link between morphology and syntactic verb raising, and provide evidence for a dissociation between mastery of inflection, which can be poor, and mastery of adverb placement, which can be target-like. The 'deficit hypothesis' is also inaccurate in that it predicts unconstrained (unsystematic) optionality at all stages of L2 development. What seems to happen instead (see Robertson and Sorace, to appear, for detailed examples and discussion) is that, as in L1 acquisition, the pattern of preferences for one option over the other changes over time, until a potentially permanent stage is reached at which the target option is strongly, but not categorically, preferred, and the dispreferred non-target option is never completely expunged, but still surfaces in some circumstances. This stage may be difficult to capture without appropriate elicitation techniques, since advanced non-native speakers' metalinguistic knowledge would in most cases successfully prevent the expression of the non-target option. The debate is thus far from being settled. The only uncontroversial conclusion is that protracted optionality in L2 grammars constitutes further evidence that pre-emption is not a consistently
operative mechanism in L2 acquisition (see Rutherford 1989; Trahey and White 1993): L2 grammar clearly have a much greater degree of tolerance for synonymy than native grammars.

On the basis of the evidence available so far, the following two generalizations about optionality can be proposed:

- 'Optionality' refers to a state of grammatical competence. It is therefore not the same as variation. Variation is not necessarily a manifestation of optionality; optionality at the level of underlying knowledge is neither a necessary nor a sufficient condition for variable performance. Sorace (1993) discusses a case of optionality in auxiliary choice that does not give rise to any variation in performance, probably because speakers have well-established routines that lead to the access of only one option in production.

- Optionality is gradient. Alternating forms are almost never in free variation, but are acceptable/determinate to different degrees (i.e. the strength of preference for one variant over the other changes over time).

The challenge for a theoretical account of optionality is to explain these generalizations, and to account only for the optionality that can possibly arise in natural language grammars, thus avoiding the overprediction of optionality. As always, the aim is to have a theory which is both descriptively accurate and explanatory. In the following section I will present a theoretical account in terms of Optimality Theory that seems to point in the right direction.

4. An optimality-theoretic account of optionality: fringes

OT is a formal theory of UG that assumes that (a) universal constraints are ranked in language-specific ways, and (b) constraints can be violated: more specifically, lower-ranked constraints may be violated in order to satisfy higher-ranked constraints (Prince and Smolensky 1993). Within this approach, the child is assumed to be equipped with knowledge of the abstract form of an OT grammar, as well as with an interpretive parser, which is essentially the same mechanism used by the competent adult speaker in assigning a structural description to an overt form in the input (see Smolensky, 1996; Tesar and Smolensky, 1998), and a constraint demotion mechanism, which is a procedure for learning constraint rankings. Learners use this knowledge of the abstract form of an OT grammar from the start to refine their ranking hypotheses and converge on the ranking required by the target language. In so doing, they rely on an iterative algorithm, which involves repeated alternations between structure-assignment (by means of interpretive parsing) and grammar learning (by means of constraint demotion). Tesar and Smolensky show that the iterative algorithm can be proved to converge very fast on the basis of limited evidence, hence confirming the effectiveness of the constraints imposed by OT on the acquisition space. Constraint demotion obtains the correct rankings, however, only at the price of
ignoring any optionality or variation in the input, since the crucial assumption is that if an underlying form \([A]\) is acquired as well-formed, any other competing form \([A']\) must be ill-formed; the learning algorithm would not know how to deal with two equally optimal candidates.

In an attempt to accommodate intermediate degrees of well-formedness within the grammar, Hayes (1997) proposes a modification of OT such that constraints are not only mutually ranked, but also have a value for strictness along a continuous scale. Variable rankings can be obtained by assigning \textit{strictness bands} to some constraints. In production (or comprehension) speakers exercise choice for the strictness of each constraint that has a strictness band. So for example, the hypothetical constraints in Figure 1 vary in their strictness through particular ranges: Constraint \(A\) does not allow any variation at all, whereas Constraints \(B\) and \(C\) have overlapping strictness bands, and the strictness band of Constraint \(E\) partly overlaps with that of Constraint \(D\).

\begin{center}
\begin{tabular}{l}
CONSTRAINT A & \begin{center} \textbf{\textbullet} \end{center} \\
CONSTRAINT B & \begin{center} \textbf{\textbullet} \end{center} \\
CONSTRAINT C & \begin{center} \textbf{\textbullet} \end{center} \\
CONSTRAINT D & \begin{center} \textbf{\textbullet} \end{center} \\
CONSTRAINT E & \begin{center} \textbf{\textbullet} \end{center} \\
\end{tabular}
\end{center}

Figure 1: Hypothetical constraints and their strictness bands

In accessing the grammar in production the speaker exercises free choice for the strictness of each constraint along its band. In the case of Constraints \(B\) and \(C\), this choice completely determines the output form. In Figure 1, for example, Constraint \(C\) dominates Constraint \(B\); but in other cases the ranking order might be the reverse: the options are in free variation. For Constraints \(D\) and \(E\), the speaker will have a strong preference for ranking Constraint \(E\) over Constraint \(D\), but in some cases the ranking order will be reversed, as determined by the area of overlap. In general, the occurrence of optional variants, and their relative degree of well-formedness, are a function of the degree of overlap between their strictness bands, and range from complete overlap (free variation) to minimal overlap (very strong preference for one variant over the other).

In some cases the range of variation of constraints is not firmly delimited, and their strictness bands will have \textit{fringes}. Fringes arise in the acquisition process when the input data are not sufficient, or sufficiently robust, to establish the upper or lower bounds of a constraint. Selection points may occur within a fringe but only at a grammaticality cost:
forms generated in this way will have a certain degree of ill-formedness. In Figure 2, for example, Constraint B is normally dominated by Constraint A. However, it is sometimes possible for Constraint B to outrank Constraint A if the speaker's selection point for B occurs close to the left edge of the shaded area of its strictness band, and the selection point for A occurs at the right edge of its band.

more strict --------------- > less strict

CONSTRANT A

CONSTRANT B

Figure 2: Hypothetical constraints with fringes.

According to Hayes, forms generated by using the fringe of a strictness band are "intuited to be mildly ill-formed" (p. 18): speakers will 'know' that the variant produced by that particular ranking has a marginal acceptability status. That form will be the dispreferred optional variant. Hayes's proposal, unlike canonical OT, allows the theory to model gradients of acceptability in natural (mature) grammars, and therefore the relative order of preferences in optionality.

6. Optionality in L2 ultimate attainment
What is the relevance of Hayes's model for L2 acquisition? Let us assume that optionality arises from the learner's trying to establish the bounds of two conflicting constraints, one of which is L1-influenced and the other is the target one. Suppose further that the learner receives partly inconsistent evidence in the input. A case in point is the V2 constraint in German-English interlanguage. Descriptively, this case is as follows. The L1-influenced constraint imposes a V2 order, leading the learner to produce sentences such as (3a), repeated here as (4):

(4) For many kids is living with their parents a nightmare

The target constraint, on the other hand, is non-V2 and leads to the production of (3b), repeated here as (5).

(5) For many kids living with their parents is a nightmare

The English input obviously abounds with evidence of non-V2 constructions. However, the learner also gets conflicting evidence in the form of 'residual V2' sentences such as

(6) Under no circumstances should luggage be left unattended
Because of the combination of the L1 constraint and the (partially) ambiguous L2 evidence, the learners may be unable to establish the upper bound of the target constraint, and thus this constraint will have a fringe. While the speaker will have a strong preference for (5), which will appear more frequently in production, the use of the fringe in production/comprehension will occasionally result in sentences like (4). This phenomenon, in all likelihood, is of much wider import. While fringes are relatively sparse in mature native grammars, they are likely to be more common in non-native grammars, since (a) L2 learners often do not get quantitatively and qualitatively robust input, and (b) they do not always 'notice' available input (see Sorace 1993). Learners may therefore introduce optionality where the L2 has a categorical choice. As a further example consider the case of inversion in subordinate clause in English, which is usually late-acquired in both L1 and L2 acquisition, and which is occasionally missing even in the speech of advanced non-native speakers. Grimshaw (1997) posits the following constraints to account for this phenomenon:

\[
\text{OPERATOR IN SPECIFIER (OP-SPEC):} \quad \text{syntactic operators must be in a specifier position}
\]

\[
\text{OBLIGATORY HEADS (OB-HEAD):} \quad \text{a projection has a head that must be filled}
\]

\[
\text{PURITY OF EXTENDED PROJECTION (PURE-EP):} \quad \text{no movement is possible into the highest head of a subordinate projection.}
\]

The (simplified) ranking for English is the following:

\[
\text{OP-SPEC} \gg \text{PURE-EP} \gg \text{OB-HEAD} \gg \text{STAY}
\]

What is crucial with respect to the lack of inversion in subordinate interrogatives is the position of PURE-EP above OB-HEAD. The final state of a near-native speaker of English may be represented as one in which the strictness bands of the two constraints overlap only by a thin fringe. This representation will occasionally generate sentences such as

(7) I don't know what time is it

in particular circumstances when the speaker's selection point for OB-HEAD involves the use of the fringe and thus the reverse ranking OB-HEAD \(\gg\) PURE-EP.

To recapitulate the argument so far: the L2 learner's task is largely one of differentiating the rankings required by the L2 from those required by the L1, and the variation bands of the L1 constraints from those of the L2 constraints, by 'pushing them apart' so that their overlap is minimal. In the vast majority of non-native speakers, the overlap is substantial; in
the minority who reach a near-native level, the overlap is minimal; in the exceptional ones, there is no overlap at all. Permanent optionality (when it does not correspond to a state of stable optionality in L2) may be regarded as a type of divergence. Divergent grammars are UG-constrained, since divergence involves a greater use of a mechanism that is allowed in native grammars.

7. Emerging optionality in L1

This brings us to the remaining question: what happens to the L1 syntax of the highly proficient L2 speaker? My current hypothesis is that prolonged exposure to / use of an L2 causes fringes to appear in what once were firmly bounded L1 constraints, which may in turn determine partial overlaps and (uncharacteristic) situations in which 'normal' rankings may be relaxed. Two examples (among many) from native Italian speakers who are also highly competent in English are the over-use of subject pronouns where a (monolingual) native would have a null subject, as in (8b) and (9), and the over-use of the SVO order where a (monolingual) native would have the subject in postverbal position, as in (10b):

(8) A: Non so perché ma Paola è uscita
   I don't know why but Paola has gone out
B: Si, lei voleva prendere un po' d'aria fresca
   Yes, she wanted to get some fresh air.
   (correct: "Si, voleva prendere un po' d'aria fresca")
(9) L'anno scorso io sono andata a Cipro
   Last year I went to Cyprus
   (correct: "L'anno scorso sono andata a Cipro")
(10) A. Perché la segreteria telefonica lampeggia?
    Why is the answering machine flashing?
   B. Perché Mario ha telefonato
    Because Mario called
   (correct: "Perché ha telefonato Mario")

Interestingly, sentences of the same type as in (8b), (9), and (10b) are also produced by advanced English non-native speakers of Italian. Grimshaw and Samek-Lodovici (1995, 1998) explain the cross-linguistic variation in the distribution of post-verbal subjects by varying the ranking of the following constraints:

ALIGN-FOCUS: Aling the left edge of focus constituents with the right edge of a maximal projection
SUBJECT: The highest A-specifier in a maximal projection must be filled.

The (simplified) rankings for Italian and English would be the following:
A. Sorace

Italian: \textbf{ALIGN-FOCUS >> SUBJECT}

English: \textbf{SUBJECT >> ALIGN-FOCUS}

With respect to referential null subjects, the relevant constraints are:

\textbf{SUBJECT:} The highest A-specifier in an extended projection must be filled.

\textbf{DROP-TOPIC:} Argument co-referential with the topic are structurally unrealized.

And the (simplified) rankings are:

Italian: \textbf{DROP-TOPIC >> SUBJECT}

English: \textbf{SUBJECT >> DROP-TOPIC}

In both cases, the grammar of the Italian near-native speaker of English might be characterized as in Figure 3:

\begin{center}
\begin{tabular}{c|c}
\hline
\textbf{DROP-TOP / ALIGN-FOCUS} & \textbf{SUBJECT} \\
\hline
\textbf{+strict} & \textbf{-strict} \\
\end{tabular}
\end{center}

\textit{Figure 3: Constraint ranking for null and postverbal subjects in some Italian-English speakers.}

These speakers will still have a preference for the L1 ranking, and will thus place a focused constituent in postverbal position, but will in some circumstances reverse this ranking by using fringes, therefore positioning focused constituents in preverbal position. The cause for emerging L1 optionality is the same as that for L2 optionality: insufficient input and conflicting evidence. What this suggests is that all grammars, native or non-native, need continued exposure to robust input in order to be not only acquired, but also maintained. This account also raises the issue of whether the dispreferred ranking obtained by the use of a fringe can be predicted to occur in a precise set of circumstances, or whether the variation is unsystematic. I will return to this question.

Are L1 and L2 optionality related? This is ultimately an empirical question. There is at least preliminary evidence from experimental phonology that divergence in both the L1 and the L2 characterizes the grammatical competence of most - but not all - very advanced non-native speakers. In a series of experiments on the acquisition of intonation, Mennen (1998) discovered two possible types of ultimate attainment in Dutch near-native speakers of Greek: most of the subjects
in this group do not establish the target Greek category for peak alignment, and also exhibit different Dutch alignment categories from Dutch monolinguals (although not an truly intermediate system as e.g. Flege’s (1995) model would predict). Only two subjects exhibit evidence of target-like attainment of Greek peak alignment, and at the same time evidence of essentially unaffected native Dutch alignment.

While it remains to be ascertained (beyond anecdotal evidence) whether this dual pattern of ultimate attainment also occurs in the acquisition of L2 syntax, we could hypothesize that the truly successful L2 learners are the minority who manage to maximally differentiate the L1 rankings from the L2 rankings, so that there are no fringes and no overlaps between them: these learners acquire native L2 competence and at the same time preserve their L1 intact. This outcome, while rare in adult language acquisition, is normal in bilingual first language acquisition (cf. Paradis and Genesee 1996, 1997).

The model that has been discussed so far has an obvious caveat: it is unable to predict precisely when optionality will become permanent in the L2, and precisely when optionality will emerge in the L1: in other words, it includes no constraints on optionality. That there are constraints on optionality is obvious from the fact that many aspects of the L2 grammar can be learned categorically, and many aspects of the L1 grammar seem to be impervious to the influence of the L2. When is optionality most likely to arise?

7. Constraints on optionality

While only further empirical data specifically bearing on this question will provide an answer, some recent studies are suggestive. Research on normal and impaired L1 acquisition points to a (temporary or potentially permanent) misunderstanding of interpretive conditions as the cause of divergences between child grammars, on the one hand, and mature grammars, on the other. Tsimpli and Stavvakaki (to appear) demonstrate the existence of a specific language impairment in the Determiner system which affects functional categories differentially, depending on whether they include LF-interpretable or non-interpretable formal features: only the latter (e.g. object clitics and the definite article) are impaired, but the former (e.g. strong pronouns and the indefinite article) are spared. For normal L1 acquisition, Wexler (to appear) explains the optionality of finite vs non-finite forms by assuming a developmental constraint which prohibits more than one checking of the DP; he hypothesizes that such a constraint is ultimately due to the child’s temporary misunderstanding of the interpretive conditions in D on the DP, whereby D can be not only [+interpretable], as in the mature grammar, but also [-interpretable].

For L2 acquisition, it has been proposed that a testbed for the existence of UG constraints on the learner’s hypothesis space is knowledge of the interpretive conditions that operate at the LF interface. Many of such conditions are undetermined by the input, and not amenable to classroom instruction: their presence in interlanguage grammars would
therefore constitute evidence for UG. Dekydtspotter, Sprouse and Anderson (to appear), for example, argue that English learners of French are sensitive to the semantic distinctions governing the licensing of multiple postnominal genitives. It is possible, however, that knowledge of interface conditions is a primary candidate for advanced or emerging optionality. The case represented in Figure 3, for example, shows that optionality affects precisely the interpretive conditions that govern the distribution of null vs overt subjects in Italian, so that both the advanced non-native speaker of L1 Italian and the native Italian speaker who has near-native knowledge English, in a minority of cases, allow both types of subjects to appear in contexts that require null subjects. This account, however, raises a number of questions that need to be addressed. It is unclear, for example, whether optionality would also affect the context which requires overt subjects to the same degree. A related issue is whether the same degree of optionality is present in the advanced non-native grammar of the Italian speaker of English, giving rise to an underproduction of overt subjects. If it were the case that there is an asymmetry in optionality, such that only overproduction but not underproduction is found, this might be evidence, under Grimshaw and Samek-Lodovici’s theory of subject universals, for markedness effects in both L1 and L2 grammars of the advanced late bilingual speaker. But an answer to these question must await further experimental research on near-native grammars.

References


Hoekstra, T. and N. Hyams, to appear. Aspects of root infinitives. In Sorace et al. (eds.).


