Pinning down the concept of “interface” in bilingualism

Antonella Sorace
University of Edinburgh and University of Tromsø.

The Interface Hypothesis (IH) was put forward by Sorace and colleagues as an attempt to account for patterns of non-convergence and residual optionality found at very advanced stages of adult second (L2) acquisition. The IH originally proposed that language structures involving an interface between syntax and other cognitive domains are less likely to be acquired completely than structures that do not involve this interface. At the same time, the IH was extended to bilingual first language (L1) acquisition and to the very early stages of L1 attrition, which exhibit optionality in precisely the same structures: this provides a unifying framework for the study of bilingual language development. This paper selectively reviews the research on the IH, addressing some common misinterpretations and outlining the most recent interdisciplinary developments.

Keywords: bilingualism, interfaces, syntax, pragmatics, processing, anaphora resolution

1. Introduction

The Interface Hypothesis (IH) was put forward by Sorace and colleagues as an attempt to account for patterns of non-convergence and residual optionality found at very advanced stages of adult second (L2) acquisition. The IH originally proposed that language structures involving an interface between syntax and other cognitive domains are less likely to be acquired completely than structures that do not involve this interface. At the same time, the IH was extended to bilingual first language (L1) acquisition and to the very early stages of L1 attrition, which exhibit optionality in precisely the same structures: this provides a unifying framework for the study of bilingual language development. The IH has spurred much research since its original formulation, particularly on the syntax–pragmatics interface; it has also raised fundamental issues concerned with the architecture
of the language faculty and the interaction of linguistic, processing, and general cognitive abilities in bilinguals. While the IH has evolved over time as new results have become available, a number of misinterpretations have arisen, primarily due to the imprecision of the original proposal but also to unwarranted extensions of the IH to domains other than the ones for which it was proposed. This paper does not aim to provide an exhaustive summary of research on the IH or directly inspired by it (for fuller discussion see Montrul, 2011; Sorace & Serratrice, 2009; Rothman, 2008): rather, it traces the progress from the original hypothesis and addresses some common misinterpretations. It concludes by offering a reading of recent developments of the IH as a productive move towards interdisciplinarity and methodological cross-fertilization in the field of bilingual language development.

2. Background

Although the term ‘Interface Hypothesis’ first appeared in Sorace and Filiaci’s (2006) study of near-native L2 speakers, phenomena that are dependent on pragmatic and contextual variables had already been at the centre of attention in research on both early and late bilingual speakers. Testing the interpretation of Italian pronominal subjects in near-native speakers of L2 Italian, Sorace and Filiaci discovered that these highly proficient speakers gave significantly different responses from monolingual native Italian speakers only with respect to overt subject pronouns; in contrast, their interpretation of null subject pronouns was entirely native-like. The pattern emerging from this study, replicated in Belletti, Bennati and Sorace (2007), showed a clear over-extension of the scope of overt subject pronouns, which led to the production and acceptance of these pronouns in the presence of a topical antecedent, as in (1b) and (2b), respectively.

(1)  
a. Perché Giovanna non è venuta?  
   “Why didn’t Giovanna come?”

b. Perché lei non ha trovato un taxi.

c. Perché ____ non ha trovato un taxi.  
   “Because she couldn’t find a taxi.”

(2)  
a. La vecchietta, saluta la ragazza quando pro____ attraversa la strada.

b. La vecchietta, saluta la ragazza, quando lei, attraversa la strada.  
   “The old woman greets the girl when ø/she crosses the road.”

In contrast, errors involving misuse and misunderstanding of null subject pronouns are not attested: both native and near-native speakers of Italian have a clear and determinate preference for the subject of the matrix clause as the antecedent of the null subject pronoun.
A similar asymmetry had been found in Tsimpli, Sorace, Heycock and Filiaci’s (2004) study of L1 attrition, which focused on the incipient changes in the native language of L1 Italian speakers who had been exposed to L2 English for a protracted period of time and had reached near-native competence in it. Similarly to the findings for L2 near-natives, this study also found an inappropriate extension of the scope of overt subject pronouns, although to a lesser degree than in L2 speakers of Italian. Both bilingual groups, however, are significantly different from monolingual Italian speakers. This suggests that anaphora resolution involving pronominal forms is an unstable domain that presents both residual optionality in advanced stages of L2 development and emerging optionality in first-generation speakers in language contact situations.

The same pattern of convergence between L2 acquisition and L1 attrition in the domain of anaphora resolution has emerged more recently from studies of a different language combination. Wilson (2009) and Wilson, Keller and Sorace (2009) looked at anaphoric dependencies with demonstratives and personal pronouns in German. In this language, both personal (er, sie, es) and demonstrative pronouns (der, die, das) can be used anaphorically. The division of labour between these anaphoric forms is similar to that between null and overt subject pronouns in null subject languages. As shown in (3), the pronoun identifies the subject of the matrix clause (NP1), whereas the demonstrative picks the complement (NP2):

(3) Der Kellner_{i} erkennt den Detektiv_{j} als das Bier umgekippt wird. Er_{i}/Der_{j} ist offensichtlich sehr fleißig.

“The waiter recognizes the detective while the beer gets spilled. He-PRON/ he-DEM is obviously hard working.”

Using a visual world eye-tracking method, Wilson tested native German speakers, advanced L2 German speakers, and German speakers who were residents in the UK. She found that the dependencies of personal pronouns are acquired; L2 learners’ preferences for the antecedents of pronouns are similar (though not identical) as those of L1 speakers. In contrast, the dependencies of demonstratives are more indeterminate; L2 speakers have either no clear preference or a weak preference for NP1. Furthermore, demonstratives are significantly more affected by attrition than pronouns. Attrited German speakers’ preferences for the antecedents of demonstratives show variability and divergence compared to monolingual Germans; either no clear preference, or a preference for NP1, depending on the length of residence in the UK. This pattern of convergence between L2 acquisition and L1 attrition is the exact parallel to the pattern obtained for Italian subject pronouns.

A third bilingual developmental domain in which a similar pattern was attested is that of bilingual L1 acquisition. Research in this area has already independently indicated the syntax–pragmatics interface as a prime locus of protracted delays
in young bilingual children who have not fully acquired the C-domain (Müller & Hulk, 2000). Anaphoric dependencies involving pronominal subjects, however, were not one of the foci of this early research; it was Serratrice, Sorace and Paoli’s (2004) longitudinal study of an Italian-English bilingual child that brought this aspect to the fore. Despite the possible prediction of a protracted early stage of null subjects in English, due to the influence of Italian, this child was shown to over-produce overt subjects from the very earliest to the last stages of data collection, while at the same time never using more null subjects than monolinguals. The same pattern has since been attested in other case studies of young bilingual children acquiring a pro-drop and a non-pro-drop language (Paradis & Navarro, 2003; Pinto, 2007; Haznedar, 2007).

The developmental difficulty associated with the use of pronominal subjects has been confirmed in a large-scale study, (Sorace, Serratrice, Filiaci, & Baldo, 2009), which compared two groups of school-age bilingual children acquiring two different combinations of languages: Italian-English (in which only one language allows null subjects) and Italian-Spanish (in which both languages allow null subjects). Elicited acceptability judgment experiments showed that both child bilingual groups accepted significantly more overt subjects referring to topic antecedents than monolingual children, although the younger monolinguals also did this significantly more often than the adult controls. These experiments therefore yield results on subject pronouns that are consistent with those obtained for other bilingual groups. The common finding is that bilinguals tend to produce and accept overt subject pronouns referring to a pragmatically inappropriate topic subject antecedent significantly more often than monolinguals.

The same children’s intuitions on other linguistic structures, however, turned out to be very different. One of these structures was the encoding of specificity and genericity in subject NPs, which distinguishes English, on the one hand, from Italian, and Spanish on the other. Only English, not the Romance languages, allows bare nominals in subject position to express generic meaning

\[\begin{align*}
\text{a. Pigs } & \text{"the pigs don’t fly.}\nonumber \\
\text{b. *Maiali / i maiali non volano} & \quad \text{“Pigs don’t fly.”}\nonumber
\end{align*}\]

The other structure involved the expression of Focus with object pronouns, which again differentiates English from both Italian and Spanish. These Romance languages express Focus by means of postverbal pronouns, while unfocused objects are encoded by preverbal clitic pronouns. In English, all object pronouns occupy a postverbal position and the difference between +/- Focus is given prosodically by stress.
Italian-English bilingual children unidirectionally accept null determiners with generic noun phrases in Italian (*In genere fragole sono rosse 'Generally strawberries are red'), but Italian-Spanish children do not (Serratrice, Sorace, Filiaci, & Baldo, 2009). The same group of children also accept a significantly higher proportion of postverbal pronouns in unfocused contexts in Italian (*Che cosa ha fatto Minnie a Paperina? Ha abbracciato lei; Serratrice, Sorace, Filiace, & Baldo, in press).

The discrepancy among the results indicates that the structures investigated are of a different nature: two of these structures (the expression of genericity and specificity, and Focusing with object pronouns) require conditions that involve formal semantic features internal to grammatical representations, the other (the anaphoric dependencies of subject pronouns) require conditions that involve contextual information external to the grammar (see Sorace & Serratrice, 2009 for more discussion). Crosslinguistic influence from English to Italian affects only structures interfacing with formal features but not structures interfacing with contextual features. Moreover, developmental effects in monolingual children are found only for the latter structures but not the former (see also Lillo-Martin & Müller de Quadros, 2011 for the late acquisition of pragmatic conditions on markers of point of view in children acquiring American Sign Language and Brazilian Sign Language).

In order to appreciate this crucial distinction between interface conditions, we need to examine in more detail the concept of ‘interface’. The strong version of the IH, in the three bilingual domains in which it was proposed, predicts that structures involving an interface between syntax and other cognitive domains present residual optionality (in L2 acquisition), emerging optionality (in L1 attrition), and protracted indeterminacy (in bilingual L1 acquisition), but structures that require only syntactic computations are completely acquired in L2, remain stable in L1, and are acquired early in bilingual L1 acquisition. This version is too broad, not only because there are developmental differences among interfaces, but also because it is unclear whether there are any structures that do not require interfaces. However, it has proved to be a fruitful starting point for research and for successive refinements of the hypothesis.
3. What’s an interface?

The term ‘interface’ refers to syntactic structures that are sensitive to conditions of varying nature: the meaning of the term therefore denotes the fact that these conditions have to be satisfied in order for the structure to be grammatical and/or felicitous. Thus, the interface between the structure and the domain that defines the conditions on its grammaticality and/or felicity is critical for its appropriate use. The original IH suggested a distinction between structures that involve an interface with other cognitive domains and structures that do not. Subsequent research has raised four issues:

a. how can interfaces be defined within a formal model of grammar?

b. is there a principled difference between structures involving interfaces with different cognitive domains? Are they developmentally different?

c. are there structures that require only syntactic computations but no interface conditions?

d. what are the reasons for the developmental instability of structures involving interfaces?

Let us consider each of these issues in turn. Both the term and the concept of interface long pre-date their use in recent studies on bilingualism. As Ramchand and Reiss (2007) suggest, the term interface can refer to both (a) the components that link sub-modules of language and (b) the link between language and non-linguistic cognitive systems. The second meaning of ‘interface’ has been explored in cognitive psychology and psycholinguistics (Friederici & Thierry, 2008; Bornkessel-Schlesewsky & Schlesewsky, 2009; Frazier & Clifton, 1996; Townsend & Bever, 2001; Keller & Asudeh, 2001; Steinhauser, White, & Drury, 2009, among many others) rather than in theoretical linguistics. With regard to the first meaning of interface, Jackendoff’s (2002, etc.) work has developed a fine-grained theory of how the modules of language (syntax, semantics, phonology) are connected with each other in a parallel architecture model. Generative grammar, since its inception, regards the interface with articulatory-perceptual systems and with conceptional-intentional systems as the level at which representations generated by the syntactic computational system converge, i.e., are legible to these external systems (Chomsky, 1995, 2000; Reinhart, 2006). The concept of interface also enjoys wide currency in linguistic frameworks of different theoretical orientations, where linguists have dealt with the interface between syntax and information structure (see Erteschik-Shir, 2007 for a review), the semantics–pragmatics interface (Kamp & Reyle, 1993), and the syntax–semantics–pragmatics interface (see e.g., Van Valin, 2006; Van Valin & La Polla, 1997).
3.1 Not all interfaces are created equal

Research on bilingual development points in particular to the interface between syntax and pragmatics as a locus of optionality and instability. Other interfaces, however, do not appear to be equally unstable, as demonstrated in Sorace et al.’s (2009) and Serratrice et al.’s (2009) study of older bilingual children. Evidence pointing in the same direction is presented by Tsimpi and Sorace (2006), who distinguish descriptively between “internal” and “external” interfaces (also see Ramchand & Reiss, 2007). Their study of advanced learners of L2 Greek shows a split between phenomena that involve ‘internal’ formal features and operations within syntax and LF, such as Focus, and phenomena that involve ‘external’ pragmatic conditions of contextual appropriateness, such as pronominal subjects. Tsimpi & Sorace argue the following:

“The distinction between the two interfaces is based on the assumption that the syntax–discourse interface is a ‘higher’ level of language use, integrating properties of language and pragmatic processing, whereas syntax–semantics involves formal properties of the language system alone.”

In Greek, object-fronting is allowed both in Focus (7a) and Topicalisation (7b) structures (Alexiadou, 1999; Tsimpi, 1995, 1998, among others).

(7)  
\[ \text{T ON \ PETRO }_1 \text{ sinandise } e_1 \text{ i adhelfi mu } \]  
\[ \text{the-ACC Petro met-3s the-NOM sister my } \]  
\[ \text{“It was Petro that my sister met.”} \]

b.  
\[ \text{T on } \text{ Petro} \_1 \text{ ton }_1 \text{ -sinandise i adhelfi mu } \]  
\[ \text{the-ACC Petro him-met-3s the-NOM sister my } \]  
\[ \text{“Petro, my sister met him.”} \]

However, Focus involves different syntactic properties from Topicalisation. Only the latter requires Clitic-Left Dislocation); furthermore, only focusing involves an operator-variable dependency, a unique (identificational) focus operator in the left-periphery, it is subject to island constraints and it requires verb-raising to the left periphery (Kiss, 1998). Tsimpi and Sorace found that advanced speakers of L2 Greek do not have problems with respect to Focusing, but they show optionality with respect to pronominal subjects involving the over-extension of the overt subject option to null subject contexts. In fact, the same split between Topic and Focus can be observed in Sorace’s & Serratrice’s studies, in which bilingual children exhibit crosslinguistic influence in the case of object pronouns but not in the case of subject pronouns. Conditions of contextual and pragmatic appropriateness regulate the choice of overt vs null subject pronouns in null subject languages; the concept of Topic, which is crucial for pronominal anaphoric dependencies,
is referential in nature and involves consideration of external conditions such as familiarity and prominence in context and interlocutor’s perspective (Gundel & Fretheim, 2006). Conditions on object pronouns, however, are of a different nature. Focus is a relational feature that refers to the informational content of a sentence and identifies new information with respect to the topic (see also Rizzi, 1997, 2004) for structural differences between Focus and Topic. Belletti, Bennati and Sorace (2007) also report a difference in near-native Italian speakers between the use of pronominal overt subjects and in the use of postverbal subjects in narrow focus contexts, but in the opposite direction from Tsimpli and Sorace (2006); their participants show optionality for both structures but are less target-like in the use of postverbal subjects than in the interpretation of overt subjects (see also Lozano, 2006). This lack of correlation between properties that have long been considered as two faces of the same syntactic parameter might be due to the different tasks employed in the two conditions (elicited production vs. comprehension in a picture verification task). It may also suggest more subtle differences in the grammatical implementation and development of Focus in Greek and Italian, as well as the need for a more fine-grained differentiation among interface conditions.

The relationship between syntax and the lexicon can also be seen as having an internal, computational side that regulates the links between semantic roles and syntactic expression (Levin & Rappaport Hovav, 1995, 2005), and an external side that looks outwards to encyclopedic knowledge and extralinguistic factors. This gap has been investigated with respect to split intransitivity, where the syntactic distinction between unaccusative and unergative verbs has been found to be conditioned by gradient aspectual distinctions encoded by the verb’s argument structure (Sorace, 2000, 2004, forthcoming). Developmentally, the syntactic distinction appears to be acquirable in L2 and stable in L1 attrition (see Sorace, 1993; & Kraš, 2008a, 2008b on auxiliary selection in Italian and Montrul, 2005a, 2005b on unaccusativity in heritage Spanish speakers and L2 Spanish speakers, among others). Sensitivity to the subtle aspectual and semantic distinctions impacting on the syntactic behavior of verbs is acquirable but does not reach the same level of determinacy as in monolingual native speakers. Moreover, it is affected by attrition, at least in second generation heritage speakers (there are no data on first generation speakers). However, whether split intransitivity can be formally regarded as a syntax–lexicon interface phenomenon depends on theoretical assumptions on the nature of the lexicon and on whether one wants to have the lexicon or the syntax do the job of encoding aspectual distinction and argument structure alternations (see Ramchand, 2008 for fuller discussion of these issues).

Finally, a recent study on simultaneous and early consecutive bilingual acquisition in Greek-English bilingual children, Argyri, Sorace and Tsimpli (2010), found a split between the acquisition of Greek gender, which is early and relatively
unproblematic for both bilingual and monolingual children, and the acquisition of voice morphology alternations, which are conditioned by lexical factors and are therefore more difficult to acquire for all children. In the examples in (8b, c), non-active (NACT) voice morphology marked on the verb signals transitivity alternations:

(8)  
   a. plen-o (wash-act.1s) /  
   b. plen-ome (wash-nact.1s)  
   c. To agori plithike  
      the boy   washed-nact.3s  
      “The boy washed himself / The boy was washed.”

Non-active morphology is multiply ambiguous in Greek because it makes available different readings — reflexive, middle, passive and anticausative — which are restricted in interpretation on the basis of lexical factors. In contrast, gender is unambiguously marked as a formal feature of the noun, which participates in the formal syntactic computation through agreement.

To sum up, there is sufficient evidence for important developmental differences between linguistic structures that require conditions of a formal nature within the grammar, and structures that require the integration of contextual factors. Calling these conditions ‘internal’ and ‘external’ is a descriptive convenience that does not imply that all formal or all contextual factors should be treated in the same way.

3.2 What is “narrow syntax” (or is there such a thing)?

The IH originally stated that structures requiring an interface between syntax and other cognitive domains may present optionality at particular stages of bilingual development (L2 advanced endstates, L1 attrition initial states, bilingual L1 intermediate states) but structures requiring only syntactic computations do not. The first part of this formulation is too broad, as discussed above, since a finer differentiation among different types of interface are necessary. Are there structures that involve only syntactic computations? The testability of the hypothesis depends on demonstrating that such structures are immune from optionality effects and on the definition of ‘narrow syntax’.

Recall that the IH was proposed for the highest possible level of ultimate attainment in L2. At first sight, the ‘interface vs narrow syntax’ distinction seems to capture some characteristics of L2 ultimate attainment in near-native speakers that had been reported in previous research. Coppieters’s (1987) seminal study of near-nativeness, for example, described larger native/near-native differences with respect to structures that involved semantic interpretation, but less noticeable
differences with respect to ‘purely syntactic’ structures. Sorace’s (1993) study, more ambiguously, reported the absence of any differences between L1 French near-native speakers of Italian and monolingual controls in their judgments on auxiliary selection in ‘clitic-climbing’ constructions (as in (9)), but found that the L1 English near-native speakers were different from the controls.

(9) Maria doveva andare dal medico, ma poi non ci è voluta /*ha voluto andare. 
    Maria was supposed to go to the doctor’s but then she there-cl not is wanted/has wanted to go
    “Maria was supposed to go to the doctor’s but then she didn’t want to go there.”

White & Genesee (1996) found complete convergence between native and near-native speakers of English with respect to subjacency principles, which syntactically constrain movement of constituents, allowing (10a) but not (10b).

(10) b. Which car did the police claim Ann had stolen?
    a. *Who does Tom love the woman who married?

In contrast, Robertson and Sorace (1999) reported that advanced L1 German speakers of L2 English optionally over-extended the V2 constraints, ignoring the lexical and/or pragmatic conditions required by residual V2 in English and producing sentences such as (11).

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

As for L1 attrition, Tsimpli et al. (2004) tested syntactic properties of null subject languages such as the grammaticality of subject extractions out of Wh-islands (Chi si chiede Maria se ha invitato Paolo? “Who does Maria wonder whether has invited Paolo?”), in addition to pronominal subjects, and found that these properties, unlike the use of overt subject pronouns, had not changed as a result of attrition.

However, two types of criticism have been raised: one is that it is difficult to identify structures that are sensitive exclusively to syntactic constraints, the other is that many structures are sensitive to multiple types of conditions and are therefore unlikely to be classifiable under only one type of interface. Both criticisms make valid, if not particularly original points. While all structures “ultimately involve discourse and must be read off at all linguistic interfaces” (Montrul, 2011), models of language that assume the existence of a syntactic module have repeatedly emphasized syntactic principles and dependencies as having a different status from non-syntactic ones in terms of acquisition and processing (Guasti, 2002; Burkhardt, 2005). Theoretical treatments of ‘multiple interfaces’ can be found in the linguistic and psycholinguistic literature (see e.g., Van Valin, 2006). In research on bilingualism, however, there have so far been few attempts to investigate the development of structures that ‘belong to’ different interfaces. One important
exception is Hopp (2007, 2009, 2010), who carried out a detailed investigation of the late stages of acquisition of scrambling in embedded clauses in L2 German by L1 speakers of Dutch, English and Russian at advanced and near-native levels (more details in Section 4.3). Hopp argues that scrambling is at the crossroads of four interfaces: syntax–morphology, syntax–discourse, syntax–lexicon, and syntax–semantics. The syntax–morphology interface is involved because scrambled constituents are identifiable by case marking (all examples from Hopp, 2007):

(12) a. Maria glaubt dass der Vater den Wagen kauft. (SO)
    b. Maria glaubt, dass den Wagen der Vater kauft. (OS)
    “Maria believes that the father buys the car.”

Scrambling requires conditions at the syntax–lexicon interface since dative experiencer verbs (e.g. *gefallen*, *nützen*, *fehlen*) are unaccusative and so have an underlying OS order. SO and OS should therefore be equally acceptable as unmarked word orders with these verbs.

(13) a. ….dass der Vater dem Onkel gefällt.
    b. ….dass dem Onkel der Vater gefällt.
    “….that the father the uncleDAT  pleases.”

The syntax–pragmatics interface is involved because scrambling changes the focus structure of a sentence by moving some constituents out of sentential focus and allowing other constituents to bear focus, as shown in the examples below where focus is capitalized:

(14) a. What is the boy reading?
    Ich glaube, dass den ROMAN der Junge liest.
    b. Who is reading the novel?
    Ich glaube, dass den Roman DER JUNGE liest.
    I think that the novel the boy reads
    “I think that the boy is reading the novel.”

Finally, scrambling indefinites in German and Dutch lead to interpretive differences, specifically, it imposes a quantificational reading, as in (15b), whereas the non-scrambled indefinite in (15a) has both a specific and an existential reading:

(15) a. … dass ich gestern einen Hausbesetzer gesprochen habe.
    … that I yesterday a squatter spoken have
    “….that I talked to a squatter yesterday.”
    b. *?* … dass ich einen Hausbesetzer gestern gesprochen habe.

Other structures that are arguably sensitive to multiple interface conditions have been mentioned in recent research: for example, the syntax–semantics–pragmatics,
syntax–morphology–lexicon, prosody–syntax–semantics, etc. These (rather unmemorable) terms just list the conditions that affect syntactic realization; by the same token, scrambling could be defined a syntax–pragmatics–lexicon–semantics interface phenomenon. Multiple-interface structures should be researched with the same breadth, depth and methodological soundness of Hopp’s study in order to see which of these conditions poses problems in bilingual development, and for what reasons.

4. What causes the optionality at interfaces in bilingual speakers?

The main challenge raised by the IH is the identification of the sources of optionality and the instability found in bilingual speakers’ use of structures that are sensitive to interface conditions. It is easier to appreciate the problem if one considers that speakers need to acquire the following:

a. knowledge of the structure and of the mapping conditions that operate within interface components, and
b. the processing principles that apply in the real-time integration of information from different domains.

It is probably fair to say that much generative research on bilingual development has targeted knowledge representations of structures and interface conditions, rather than the real-time processing operations involved in their production and comprehension. Indeed, early work on the IH had taken this restricted perspective on the nature of interfaces. The accounts that have since been proposed mainly fall within two camps: the first (which will be labeled the representational account) assumes that there are differences between bilinguals and monolinguals at the level of knowledge representations, in most cases because one of the grammatical systems affects the other. The other (which will be termed the processing resources account) looks at differences between bilinguals and monolinguals at the level of processing strategies required in the use of interface structures in real time. Let us first review the arguments underlying the representational account.

4.1 Underspecification of interpretable features

Recall that different groups of bilinguals who speak a null subject language and a language, like English, that does not allow null subjects, tend to overgeneralize overt pronouns to contexts which would call for a null pronoun. The representational account holds that because of the knowledge of English, where subjects are typically always overtly expressed, bilingual individuals’ representation
of the pragmatic constraints regulating null and overt pronouns are weakened in the null-subject language. English, the language with the most economical syntax–pragmatics interface system for subject pronouns (null subjects are typically not allowed), is assumed to influence the language with a more complex interface system (Italian, Spanish); null subjects are allowed alongside overt pronominal subjects and where their distribution is regulated by pragmatic constraints.

This approach was in fact proposed by Tsimpli et al. (2004) for L1 attrition in Italian and Greek. The reason argued to be at the root of the extension of overt pronouns is underspecification of the interpretable feature [+Topic Shift], which in the monolingual grammar is mapped onto the overt pronoun. Thus, the monolingual Italian grammar before the onset of attrition involves a one-to-one mapping for each pronominal form:

(16) Monolingual grammar:

\[
\text{OVERT} \leftrightarrow [+\text{TS}]
\]

\[
\text{NULL} \leftrightarrow [−\text{TS}]
\]

In contrast, the attrited L1 Italian grammar has the same mapping for the null pronoun but two mappings for the overt pronoun:

(17) Attrited grammar:

\[
\text{OVERT} \leftrightarrow [+\text{TS}]
\]

\[
\text{OVERT} \leftrightarrow [−\text{TS}]
\]

\[
\text{NULL} \leftrightarrow [−\text{TS}]
\]

The same underspecification account can be extended to residual optionality in L2 acquisition. Optionality in L2 grammars/L1 attrition involves the residual/emerging underspecification of discourse interface conditions linked to a parametric choice that differs between the L1 and the L2 (see also Belletti, et al, 2007). The state of underspecification is due to the absence of a similar condition in L1 (or L2) English in the same syntactic context. This account explains the directionality of crosslinguistic effects: it is always the language that instantiates the less restrictive option that affects the other, but not vice versa. Hence, it is English that affects Italian regardless of whether it is the L1 or L2. Endstate grammars in L2 acquisition present neutralization of target L2 distinctions towards the less restrictive L1 option. L1 individual attrition similarly involves neutralization of native distinctions towards the less restrictive L2 option.

The limitation of the underspecification account is that it applies only to bilinguals that are speaking language combinations in which one of the languages has a complex setting dependent on syntax–pragmatics interface conditions and the other does not.
However, overt pronouns are also overextended by bilingual speakers of two null subject languages. The results of Sorace et al’s and Serratrice et al’s studies on older bilingual children are an example: both the Spanish-Italian bilinguals and the English-Italian bilinguals were significantly more likely than the Italian monolinguals to accept as pragmatically appropriate a null subject pronoun in contexts in which the antecedent was not a topical subject. These findings suggest that, regardless of whether the bilinguals’ other language is a null subject language (Spanish) or not (English), the process of anaphora resolution in Italian is affected. In addition, the inappropriate use of overt subject pronouns has been attested in late adult bilingual speakers of two null subject languages: Bini (1993: L1 Spanish → L2 Italian), Margaza and Bel (2006: L1 Greek → L2 Spanish), Lozano (2006: L1 Greek — L2 Spanish), Guido Mendes and Iribarren (2007: L1 Spanish — L2 Brazilian Portuguese). In all these studies, the null subject language being acquired appears to have been the participants’ second language. In contrast, the phenomenon was not observed in a few studies in which the null subject language was the participants’ third language: Kraš (2008a,b: L1 Croatian — L3 Italian); Montrul, Dias, and Thomé-Williams (2008: L1 Spanish — L3 Brazilian Portuguese); De Prada (2009: L1 Spanish— L3 Catalan). We will return to the possible reasons for this discrepancy between L2 and L3 acquisition. For the moment, it is important to notice that if speakers of two similar null subject languages overgeneralize the scope of the overt subject pronoun, the representational account and the crosslinguistic influence argument behind it are weakened. However, very recent experimental evidence on Spanish and Italian warns us that the two languages are very similar in their inventory of pronominal forms and in the broad universal principles governing anaphoric dependencies, but they are not identical with respect to the scope of the overt pronoun (Filiaci, 2010; Filiaci, Sorace & Carreiras 2010). They are very similar in their inventory of pronominal forms and in the broad universal principles governing anaphoric dependencies (Ariel, 1990).

The possibility exists that the attested patterns of optionality at the syntax–pragmatics interface may, in actual fact, reflect differences in processing between bilingual and monolingual speakers, rather than (only) crosslinguistic influence (see Sorace, 2006a; Sorace & Filiaci, 2006; and Sorace & Serratrice, 2009) for fuller discussion).

Under this account, the hypothesis is that bilinguals are less efficient than monolinguals in the integration of multiple sources of information and that bilingualism itself, rather than (only) the particular language combination spoken, may be the underlying cause of the observed differences with monolinguals.

What could make bilingual processing less efficient than monolingual processing? There are at least two routes to explore in this respect:
a. bilinguals are less efficient than monolinguals because their knowledge of or access to computational constraints within the language module is less detailed and/or less automatic than in monolinguals;
b. bilinguals are less efficient than monolinguals because they have fewer general cognitive resources to deploy on the integration of different types of information in online language comprehension and production.

4.2 Processing costs of interface structures

A vast psycholinguistic literature on anaphoric dependencies in monolingual speakers provides evidence that parsing based on syntactic constraints is faster and more automatic. Syntactic dependencies, such as binding constraints, are more economical and operate at the very early stages of processing. For example Sturt’s (2002) study of reflexive anaphors shows that early processing depends exclusively on Principle A of binding theory (which binds the reflexive to a local antecedent in the same clause), although it does not fully constrain subsequent stages of anaphora resolution. Much research in this area provides evidence that accessing and integrating two levels of representation (e.g., syntax and discourse, syntax and lexical semantics) is more costly than accessing only the syntactic level. Thus, referring pronouns (as in 18a) are more costly to process than bound variable pronouns in 18b) (Burkhardt, 2005; Piñango, Burkhardt, Brun, & Avrutin, 2001).

(18) a. The teacher thinks that the students like him.
    b. Everyone thinks that the students like him.

Similarly, logophoric reflexives (as in 19a) are more costly to process than co-argument reflexives, as in (19b) (Piñango & Burkhardt, 2005).

(19) a. The lawyer who was young defended himself.
    b. The daughter hid a present behind herself.

The same differential costs between syntactic and discourse dependencies can be seen in the division of labor between anaphoric forms in languages like Italian and German. Carminati (2002, 2005) for Italian (and later Alonso-Ovalle, Clifton, Frazier and Fernández-Solera, 2005 for Spanish) proposes the Position of Antecedent Strategy (PAS) as a processing constraint on anaphora resolution. The PAS states the following in intersentential anaphoric contexts:
- null pronouns are assigned to the antecedent in Spec, IP (normally, the subject).
- overt pronouns are assigned to an antecedent in a structurally lower position (normally, a non-subject).
Experimental results in these studies, however, show that the antecedent preferences of overt pronouns are more flexible than those of null subjects. Adult monolinguals sometimes disregard these preferences when the context is unambiguous, while they tend to respect the preferences when the context is ambiguous. Thus, an overt pronoun in (20), which is unambiguous because only one possible referent is present in the linguistic context, is both more acceptable/less costly in processing and more likely to be produced than in (21), where two equally plausible and grammatical antecedents are present.

(20) Paolo ha detto che pro / lui andrà al matrimonio di Maria.
   “Paolo has said that Ø / he will go to the wedding of Maria.”

(21) Marta scriveva spesso ad Anna quando pro / lei era in vacanza.
   “Marta wrote frequently to Anna when Ø / she was on holiday.”

For German, Kaiser and Trueswell (2008) propose a Form-Specific Multiple Constraint approach which explicitly acknowledges the difference between the anaphoric dependencies of personal pronouns and demonstratives. The proposal is that personal pronouns are sensitive to a syntactic dependency and identify subject antecedents, whereas demonstrative pronouns are sensitive to a discourse dependency and identify non-topic antecedents. The form-specific multiple-constraint approach may account for the more variable behavior of L2 German speakers and the L1 German speakers experiencing attrition, described in Wilson (2009). Corpus evidence suggests that the use of demonstratives is more flexible than the use of personal pronouns in native German speakers and that these forms are sometimes used to refer to topic antecedents. This example in (22) from the Frankfurter Rundschau (8 October 2007) proves the point:

(22) Interviewer: Die Boxerin Joe ist 19 Jahre alt, aber Sie verkörpern sie völlig glaubhaft. Andererseits spielen Sie oft Frauen, die sehr erwachsen wirken. Wie schafft man das?
   Interviewer: The [female] boxer Joe is 19 years old, but you portray her completely convincingly. Yet you often play women who come across as very mature. How does one do that?
   Actress: Die Frauen sahen früher viel älter aus, als sie waren, und ich glaube, das hat neben Kostüm und Make-up auch viel mit Körpersprache zu tun. Für Joe wiederum gilt das auch, denn die ist trotz ihrer 19 Jahre immer noch nicht in ihrem erwachsenen Körper angekommen. Im Gegensatz zu Joe fühlt sich die Frau aus dem “Wunder von Bern” in ihrem Körper und ihrer schicken Kleidung ausgesprochen wohl. Ich suche bei jeder Rolle nach so einem Ansatzpunkt: Wie lebt diese Frau in ihrer eigenen Hülle?
Actress: Women used to look a lot older than they were, and I think that in addition to dress and makeup that had a lot to do with body language. That’s true for Joe, too, because despite being 19 she (DEMONSTRATIVE) still doesn’t feel right in her adult body. Unlike Joe, the woman in “The Wonder of Berne” feels extremely comfortable in her body and her fancy clothes. With every role I’m looking for some kind of approach: how does this woman live in her own shell?

While Wilson (2009) and Wilson, et al (2009) focus on comprehension and therefore do not provide any direct evidence of overgeneralization of demonstratives in bilingual production, there is at least one study (Juvonen 1996) which shows that Finnish-Swedish bilinguals use demonstratives more extensively than either Finnish or Swedish monolinguals.

The Form-Specific Multiple Constraint approach may also be extended to Italian subject pronouns, which exhibit the same split between the stable and consistent preferences of null pronouns attested in both monolingual and bilingual speakers as well as the more variable and unstable preferences of overt pronouns. This ties in with Carminati’s intuition about variation among null subject languages. As she suggests, variation among these languages may be largely restricted to the scope of the overt subject pronoun (see Filiaci, et al., 2010 for experimental evidence on Spanish and Italian).

4.3 Bilingual processing

One reason why bilingual speakers may be less efficient at processing structures at the syntax–pragmatics interface is that syntactic processing is less automatic for them. This may be due to less developed knowledge representations or to less efficient access to these representations. Comparatively more research on these issues has focused on adult L2 acquisition than on L1 attrition or bilingual children.

The “Shallow processing” hypothesis proposed by Clahsen and Felser (2006a, 2006b) assumes that L2 speakers use ‘shallow’ processing strategies that privilege semantic, pragmatic and lexical information at the expense of syntactic information. This route is available to native speakers too but L2 speakers may have no option other than resorting to shallow processing because they do not have target-like grammatical representations. As Sorace (2006b) argues however, there is no evidence that the Shallow Processing account applies to near-native speakers, who have been shown to have native-like representations and syntactic knowledge but optionality at the syntax–pragmatics interface. At this very advanced competence level, then, it is more plausible to assume that the residual difficulties may be related to access and integration of syntactic knowledge.
An early study by Kilborn (1992) provided revealing evidence for this hypothesis. Kilborn asked native and advanced non-native speakers of English to perform in a word-monitoring task in which they saw three types of sentences. The task required speakers to detect a target word in the sentences: for example, the word ‘rules’ in (23) or ‘trip’ in (25):

Normal sentences:

(23) Playing hockey without observing the rules is very dangerous.

Syntactic structure but no meaning:

(24) Checking gravel without walking the train is perfectly yellow.

Neither syntactic structure nor meaning:

(25) Is ducks without securely the tired trip blocking illegal.

Native participants were tested in both a ‘normal’ condition and in a ‘noise’ condition in which the auditory stimuli were partly masked by background noise. L2 speakers were tested only in the normal condition. The results show that, for all participants, performance in word monitoring is best for the normal sentences and worst for the random sentences. However, both L2 speakers in the normal condition and L1 speakers in the noise condition are less able to integrate syntactic with semantic information as rapidly as L1 speakers under normal listening conditions.

The same reduced ability to integrate syntactic and contextual information is reported in a study by Roberts, Gullberg and Indefrey (2008), who investigated the interpretation of overt pronouns in L1 Turkish and L1 German speakers of L2 Dutch who were resident in the Netherlands and scored highly in language proficiency tests. The materials used in the experiments exemplified three conditions:

Local Resolution:

(26) De werknemers zitten in het kantoor. Terwijl Peter aan het werk is, eet hij een boterham. Het is een rustige dag. “The workers are in the office. While Peter is working, he is eating a sandwich. It is a quiet day.”

Disjoint Resolution:

(27) De werknemers zitten in het kantoor. Terwijl Peter aan het werk is, eten zij een boterham. Het is een rustige dag. “The workers are in the office. While Peter is working, they are eating a sandwich. It is a quiet day.”
Pinning down the concept of “interface” in bilingualism

Optional Resolution:

(28) Peter en Hans zitten in het kantoor. Terwijl Peter aan het werk is, eet hij een boterham. Het is een rustige dag.

“Peter and Hans are in the office. While Peter is working, he is eating a sandwich. It is a quiet day.”

Both online and off-line tasks were employed: eye-tracking (reading for meaning), a grammaticality judgment test, and a comprehension questionnaire.

The patterns of results are different across the three tasks. The three groups performed similarly in the grammaticality judgment test. However, L1 influence was observed in the Comprehension test: both the Dutch and the German participants chose the internal referent (i.e., the most prominent topic antecedent) as the antecedent for the pronoun, but the Turkish learners preferred the external referent half of the time. This betrays a possible influence of the Turkish null subject system, in which overt pronouns refer to non-topical antecedents. The Turkish learners’ optional preference for disjoint interpretations of pronouns may be related to the greater flexibility of the overt pronoun in null subject languages. Strikingly, the analysis of fixation duration patterns in the eye-tracking shows a native/non-native difference but no L1 influence; both L2 groups spent more time reading the pronoun in the Optional conditions than in the other two conditions. This disadvantage is not seen in the Dutch group, who have the shortest fixation times overall for this condition. Roberts et al.’s interpretation is that this processing disadvantage attested for the Optional condition in both non-native groups, regardless of their divergent performance in the off-line task, signals a greater effort in integrating syntactic information with the appropriate discourse conditions.

Finally, Hopp’s (2007, 2009) studies of German scrambling explicitly point to the possibility of complete convergence at the representational level for near-native speakers, and also to remaining problems in integrating information across modules, including morphological information, which is more difficult to use online (see Slabakova’s 2008 Bottleneck Hypothesis). Recall that Hopp tested both advanced and near-native L2 speakers; only the advanced participants have more problems with inflectional morphology and do not use morphological information for incremental reanalysis in L2 processing. The near-natives, however, pattern together with natives.

In Hopp’s words,

“L2 problems at the interfaces would be due to economy strategies that minimize processing resources rather than to problems specific to interfaces per se.... phenomena at the syntax–discourse interface may not present insurmountable difficulty in adult L2 acquisition, neither in off-line comprehension, nor in on-line
processing, and, in consequence, that this particular area of L2 acquisition is not constrained by representational deficits.”

Hopp’s conclusion is that there is no obstacle to convergence of knowledge representations between native and near-native speakers, but reduced processing efficiency in L2 speakers may prevent full convergence in all tasks. The problems are more visible in tasks that tax processing routines. Increased computational task demands affect native controls, indicating the absence of qualitative difference between native and non-native speakers.

Overall then, there is robust evidence that at least L2 speakers are less efficient than monolinguals at integrating information from different domains in real-time language use. Let us now revisit the possible sources of instability at the syntax–discourse interface for bilingual speakers. There is enough evidence that attributing these phenomena only to differences between monolinguals and bilinguals at the representational level would be inaccurate. The underspecification account, which is based on the effects of the language offering a more economical setting, does not reach very far in explaining why similar patterns are attested in speakers of different language combinations. It is possible that the bilinguals’ problem lies in their less-than optimal ability to consistently and efficiently integrate different types of information. This account fits with the fact that (a) bilingual-monolinguial differences appear to be more quantitative than qualitative, (b) performance is affected by the characteristics of the task, and (c) much variation is attested both within and across individual speakers. From a methodological point of view, it is important to dispel a frequent misconception that only online tasks can test the processing resources account. In fact, both offline and online tasks give insights about speakers’ processing abilities, although only online tasks (for example, eye-tracking or self-paced reading methods) provide information about the temporal resolution of processing. External load on both offline and online tasks (in the form of a concurrent task, for example, or time pressure, as in a speeded grammaticality judgment test) can provide information about the point at which the processor breaks down under load and thus about the processing resources available to the speaker. An untimed grammaticality judgment task, however, may be too close to the metalinguistic end to capture optionality at the syntax–pragmatics interface, and is not the best method to investigate this phenomenon.

The over-extension of overt pronouns, under the processing resources account, becomes a default strategy that compensates for occasional failure to compute the correct syntax–pragmatics mappings in real time. There is independent evidence from studies of language contact, pidginization, and diachronic change that overt subject pronouns enjoy this unmarked default status (see Bresnan, 2001, 2004; Sprouse and Vance, 1999). For bilinguals, overt pronouns may be regarded
as a ‘fall back’ option that has the advantage of being redundant rather than ambiguous; a type of ‘heuristic’ that even native speakers may employ when the use of syntactic computation is less available (see e.g., Ferreira & Patson, 2007). Among the possible reasons that may lead bilingual speakers to use overt pronouns as a default strategy, two non-linguistic factors are being explored in current research: input and executive control.

4.4 Input

The effect of the quantity and quality of input and exposure received by bilingual speakers is likely to be particularly relevant in an explanation of the reduced integration ability of bilingual speakers. Quantitative and qualitative properties of input are especially critical for individual L1 attrition; native speakers who are removed from their original language community, first of all, experience a decrease in the overall quantity of input in their language, which in turn determines a drop in the frequency of the appropriate usage of pronominal forms. Secondly, they are likely to be exposed to their language spoken by both other L1 attrited speakers and L2 speakers. These effects interact with those of exposure to the L2 in ways that are still not understood (see Sorace, 2005 for hypotheses about the differential impact of changes in input quantity and input quality).

Input effects have been shown to be relevant for bilingual L1 acquisition. Sorace et al and Serratrice et al tested English-Italian bilingual children both in Italy and in the UK; they found that the bilinguals living in the UK, who received more English input, perform in Italian less accurately than the bilinguals living in Italy, although the overall pattern of behavior is the same. In a similar vein, Argyri and Sorace (2007) in their study of older Greek-English bilingual children observed an overextension of over subject pronouns in Greek, but only in the English-dominant group. Thus, the issue of ‘dominance’ in bilingual children is far from settled. In some cases, linguistic effects seem to override dominance (as in the child in Serratrice, et al, 2004; see also discussion in Müller, 2007), but in other cases it interacts with language development in significant ways. Important input effects were recently found by Unsworth, et al, (2010) and Argyri Sorace, and Tsimipli (2010) in their study of simultaneous vs. early consecutive child bilingualism: these researchers found that the type and quantity of input interacts significantly with age of first exposure to a second language and it is therefore impossible to analyze age effects without taking into account exposure effects. These findings suggest that linguistic factors, age of exposure, and input interact with each other in defining the outcome of bilingual L1 acquisition and more research is needed to understand their interactions.
The importance of input also clearly emerges, from a different perspective, in research on priming and alignment mechanisms (Costa, Pickering & Sorace, 2008). A large body of psycholinguistic research indicates that speakers are susceptible to priming effects, which lead to the enhanced activation of a structure that has just been heard (Loebell & Bock, 2003; Hartsuiker, Pickering & Veltkamp, 2004). More controversially, research on antipriming (Marsolek, 2008) shows that priming effects involve the depressed activation of a related and partially overlapping structure. Both priming and antipriming have immediate effects, but also potential longer-term effects (Allport & Wylie, 2000). Furthermore, research on negative priming in monolinguals (Tipper, 1985, 2001) and in bilinguals (Treccani, Argyri, Sorace, & Della Sala, 2009) shows that inhibition of a competing stimulus leads to difficulty in accessing that stimulus in successive trials; this is known as the negative priming effect. These effects need to be better understood for three reasons. First, bilinguals who use both their languages, alternate between languages and different contexts of use for pronouns many times in the course of the day. Second, in experimental L2 studies it is common to use a sequence of switch trials that alternate sentences requiring null pronouns and sentences requiring overt pronouns. For example, there may be sequencing effects causing negative priming. Inhibition of overt pronoun → non-subject mapping may lead to occasional slower/more inaccurate access to this mapping when it is called for. Consider an experimental sequence of trials involving (29a) presented before (29b):

(29) a. Piero non vede più Guido da quando si è sposato.
    “Piero hasn’t seen Guido since he got married.”
    b. Maria ha scritto a Francesca quando lei era negli USA.
    “Maria wrote to Francesca when she was in the States.”

In order to perform the appropriate antecedent-assignment mapping operation pronoun → matrix subject in 29a, speakers have to inhibit the inappropriate mapping pronoun → matrix complement. If 29b is encountered after 29a, the pronoun → matrix complement mapping becomes the appropriate one, but there will be a cost in performing it because of the previous inhibition (= negative priming effect).

Third, mechanisms of priming and alignments may be responsible for the synchronic diffusion of forms produced by L2 speakers and L1 attrited speakers within the same community, which increases their frequency in the input.

In sum, more research is needed on the effects of input and exposure, beyond the simple concept of ‘dominance’, in order to establish the impact of variation in these two parameters on language acquisition and language maintenance in bilinguals.
4.5 Executive control in bilinguals

If the bilingual inefficiency at computing syntax–pragmatics mappings is a temporary, fluctuating problem, it may relate to bilingualism per se and to the allocation of general cognitive resources in bilingual processing. While there is a growing body of research on the effects of bilingualism on general cognition (Costa, Hernández, & Sebástian-Gallés, 2008; Bialystok, 2009, among many others), research has just begun to explore the effects in the opposite direction, that is, from general cognition to linguistic performance at the syntactic level.

4.5.1 Resource allocation

Wilson (2009) tested native German speakers in an extra processing load condition (a concurrent memory task) to determine whether the non-target behavior of bilingual speakers was simply a problem of cognitive resource limitations. The reasoning was that if both L2 German speakers and attrited L1 German speakers have a capacity limitations in processing resources, we would expect the two groups to be similar to the native speakers under processing pressure. However, this is not what was found; the results are that processing load affects both pronouns and demonstratives, leading to an inability to form both syntactic and discourse dependencies and preventing successful integration of information.

An alternative view on the nature of the bilingual problem is that it might be one of resource allocation in the calculation of syntax–discourse dependencies, rather than resource limitation (Wilson, Keller, & Sorace, 2010) Resource allocation has been defined as the ability to flexibly direct attentional resources as a function of the complexity of the incoming material (Titone, Prentice, & Wingfield, 2000). This ability is more likely to break down in less automatic processes, such as those dependent on the integration of syntactic and pragmatic information, leading to cognitive discoordination. The effect of resource misallocation is that bilinguals may occasionally direct attention to ‘the wrong referent’, which delays them and prevents successful integration of information, and ultimately successful interpretation/encoding of anaphoric dependencies. Interestingly, discoordination in pronominal reference is also emerging as a factor in other populations sensitive to variation in cognitive load, including ageing speakers (Titone, et al, 2000), schizophrenic patients (Phillips & Silverstein, 2003) and autistic children (Arnold, Bennetto, & Diehl, 2009).

4.5.2 Inhibition of the language not in use

There is a consensus in the cognitive psychology literature on bilingualism that both languages are always simultaneously active (Green, 1998). Bilinguals therefore need to exercise executive control to avoid interference from the unwanted
language; it is this experience of constantly inhibiting the unwanted language that is supposed to be at the root of the well-attested executive control bilingual advantage in non-linguistic tasks (see Bialystok, 2009 for an overview; Treccani, et al, 2009). The need to keep the two languages separate, however, may take attentional resources away from other tasks, including linguistic tasks. One hypothesis is that the computation of anaphoric dependencies partly draws on the same pool of attentional resources used to inhibit the language not in use. If this is the case, it should be possible to find a correlation between the amount of attention deployed on language inhibition and frequency of extension of the overt subject pronoun as a default strategy. This correlation would shed light on the differential magnitude of the effect in L2 near-native speakers and attrited L1 speakers. Recall that the over-extension of overt pronouns is more pronounced in L2 acquisition than in L1 attrition. This may be due to the fact that in L2 speakers of Italian, the unwanted language is their (still dominant) L1, which needs more resources to be inhibited. In L1 Italian speakers affected by attrition, on the other hand, the unwanted language is their (less dominant) L2, which needs fewer resources to be inhibited. In other words, Meuter and Allport’s (1999) asymmetric switching effects may be at work.

Furthermore, this hypothesis might account for the fact that the overextension of the overt pronoun has so far been attested in native speakers of a null subject language who acquire another null subject language as a second, but not as a third language. There may be less resource competition for L3 learners and for multilinguals in general. Costa and Santesteban (2004) propose that balanced bilinguals may develop a language selection mechanism that does not depend on inhibition and can be applied to subsequent languages. It is therefore possible that trilinguals, or polyglots in general, show less or no overgeneralization of overt pronouns because they do not need to apply as much inhibitory control as bilinguals.

It is important to realize that these cognitive explanations are not inconsistent with representational accounts. Different factors may interact in a cumulative way without necessarily cancelling out each other. Recent research suggests increasing awareness that no single factor can explain all of the data and that accounts in ‘either-or’ terms have limited explanatory power.

5. Criticisms of the IH

The IH has been influential in generating research and is evolving to accommodate new findings. The hypothesis has given rise to a number of often valid criticisms but also a few misunderstandings, generally arising from either an exclusive focus on representational aspects of bilingual development or from a lack of consideration of the stages of bilingual development for which the IH has been proposed.
From a generative grammar perspective, it may appear as if the IH confirms the ‘access to UG’ position and relegates problematic L2 behavior to the interfaces. In fact, the IH is agnostic on the ‘access to UG’ question; it points to the need for interdisciplinary work that goes beyond theories of linguistic representations.

The IH requires an adequate theory of linguistic representations, to the extent that generative grammar still offers the most detailed theory there is and provides one of the crucial elements that have to be considered in order to explain bilingual development. However, linguistic theory is not sufficient by itself to explain the interaction of linguistic and non-linguistic factors that appears to underlie optionality in linguistic behavior. For the same reason, the method typically used in generative linguistic research, grammaticality judgment tasks, is not the most suitable to capture states of optionality or inconsistent integration of information in processing (although some versions of this method, such as magnitude estimation or speeded grammaticality judgment tests may be more suitable for the first purpose; see Bard, Robertson & Sorace, 1996; Sorace, 2010; Bader & Häussler, 2009).

As has been explained in this paper, the initial dichotomy between ‘interfaces’ and ‘pure syntax’ was too broad. The vagueness of the original version of the IH has been the target of two kinds of criticism: on the one hand, that “all structures interface with something” and, on the other, that “not all interfaces are equally problematic.” As pointed out in Section 3.1, more recent research has revealed that not all interfaces present the same types of problems in bilingual development. Moreover, it is not necessary to neatly classify structures as ‘belonging’ to only one type of interface, as particular linguistic structures may respond to different types of interface conditions, which in turn may pose different kinds of learnability challenges to bilingual speakers. The researcher’s job is to identify these conditions and design studies that can capture bilingual speakers’ acquisition and real-time access to them; well controlled experiments can allow researchers to establish both whether bilingual speakers on average have the same kind of ability with respect to a given structure that monolingual speakers on average have, and whether fluctuations in individual production and comprehension of specific structures depend on factors external to their linguistic competence. Research on native language processing indicates the plausibility of distinguishing between syntactic conditions and dependencies and conditions that involve consideration of contextual factors. Instead of a rigid dichotomy, it may be more appropriate to differentiate structures on a gradient according to the type of conditions they need to satisfy and whether they are closer to the ‘strongly biasing’ syntactic end or to the ‘weakly biasing’ contextual end.

Some researchers (White, 2009; Montrul, 2011) question the conclusion that structures at the syntax–pragmatics or the syntax–semantics interface present protracted instability in bilingual acquisition. Focusing on interpretive properties
of syntactic constructions (such as adjective position, word order alternations, semantic vs. pragmatic implicatures, and aspectual distinctions) and mostly using data from untimed metalinguistic tasks, they conclude that these structures are acquired by L2 speakers (see e.g., Slabakova & Montrul, 2003; Rothman & Iverson, 2008; Rothman, 2008 for an overview). The incorrect assumption here is that the IH predicts the impossibility of ‘resetting parameters.” However, the knowledge representation of these structures is not in question; the IH predicts that both syntactic and pragmatic conditions are acquirable, but the integration of syntactic and pragmatic conditions remains less than optimally efficient and gives rise to optionality.

Criticisms of the IH sometimes ignore the fact that the hypothesis is not about intermediate stages of L2 development or inter-generational L1 attrition. For L2 acquisition, the IH does not predict that L2 learners of a null subject language who are still in the process of learning do not experience developmental problems with respect to null subject pronouns (see Montrul & Rodríguez-Louro, 2007): it predicts that the speakers at the highest level of L2 ultimate attainment will exhibit residual optionality only with respect to overt pronouns. This type of ‘problematicity’ seems quite distinct from the developmental problems encountered by learners at developmental stages. Positions such as the ‘Bottleneck Hypothesis’ (Slabakova, 2008) which assume that functional morphology is the main problem in L2 acquisition, both developmentally and at the near-native level, have so far not been tested with near-native speakers and with a wide range of experimental methods. Thus, it is unclear what kind of bottleneck is posed by functional morphology; future research will have to tease apart difficulties due to the acquisition of representations and difficulties due to processing (access and retrieval) of morphological exponents, and test speakers who have reached the highest stages of L2 acquisition. As for L1 attrition, the prediction of the IH is that emerging optionality in individual attrited L1 speakers is initially manifested only with overt pronouns (and possibly with other syntax–pragmatic structures), but at later stages of the process of generational transmission there may be changes that concern the whole pronominal system and/or structures that involve different interface conditions. Heritage speakers (who are the focus of most research on L1 attrition) are one stage ahead compared to first-generation individual attrited speakers; the input received by second generation speakers is typically different from the input received from their parents and their grammatical representations may also be affected by attrition, in addition to their integration abilities in processing. Heritage speakers may acquire a divergent grammar, if the input is only qualitatively different, or an incomplete grammar, if the input is also quantitatively impoverished (Sorace, 2005; Rothman, 2007; Pires & Rothman, 2009).
6. Conclusions (on a positive note)

The IH has opened up new directions of research on bilingualism and has encouraged reflection on the foundations of this field. First, work on the IH has indicated the need for a model of bilingual development that gives equal weight to formal definitions of linguistic phenomena and to their interactions with other cognitive systems. To use Sprouse, Wagers, and Phillips’s (2010) terms, what is needed to interpret bilingual development is a ‘grounded’ theory of grammar. Second, emerging results are pushing researchers towards serious interdisciplinary work that benefits from a wider range of methodologies and perspectives: there is no reason why ‘linguistic’, ‘psychological’ and ‘neurocognitive’ (but also ‘sociological’) research on bilingualism should be in separate sub-fields that ignore each other if, as Rizzi (2004: 325)) states, for linguistic theory “the final objective is a full integration of the different levels postulated in the study of the mind/brain.” Third, work on the IH has encouraged comparison and cross-fertilization across the sub-domains of bilingual L1 acquisition, L2 acquisition, and L1 attrition. Using other bilingual speakers as a term of comparison, rather than only the ‘classic’ monolingual speaker, is not only methodologically more sound (as has been known for a long time by researchers working on child L2 acquisition — see Schwartz, 1998) but helps to see the ‘forest’ of a general model of bilingualism beyond the individual ‘trees’ of bilingual types. This in turn is a concrete move away from the concept of bilingualism as the ‘sum of two monolinguals’, which Grosjean (1998, 2010) was the first to challenge. All in all, there is a long way to go to address many of the issues raised by the IH, but the road is wide open before us.

References


All rights reserved


Pinning down the concept of “interface” in bilingualism


Unknown article title. 8 October (2007). *Frankfurter Rundschau*.


**Author’s address**

Antonella Sorace
The University of Edinburgh
School of Philosophy, Psychology and Language Sciences
Dugald Stewart Building
3 Charles Street
Edinburgh EH8 9AD
United Kingdom
antonella@ling.ed.ac.uk