

1 The ‘Comparative Logic’ and why we need to explain interlanguage grammars

2 **Laura Domínguez^{1*}, María J Arche²**

3 ¹Modern Languages and Linguistics, University of Southampton, Southampton, United Kingdom

4 ²Centre for Research and Enterprise in Language, School of Humanities & Social Sciences,
5 University of Greenwich, United Kingdom

6 * Correspondence:

7 Laura Domínguez

8 ldo@soton.ac.uk

9 **Keywords: Comparative Fallacy, native speaker, interlanguage, control group, Universal**
10 **Grammar**

11 Abstract

12 In this paper we argue that the Comparative Fallacy (Bley Vroman 1983), which warns against
13 comparisons between native speakers and learners in Second Language Acquisition research, is not
14 justified on either theoretical or methodological grounds and should be abandoned, as it contravenes
15 the explanatory nature of SLA research. We argue that for SLA to be able to provide meaningful
16 explanations, grammatical comparisons with a baseline (usually of native speakers although not always
17 the case) are not only justified but necessary, a position which we call the ‘Comparative Logic’. The
18 methodological choices assumed by this position ensure that interlanguage grammars are analysed on
19 their own right and respecting their own principles. Related issues such as why we focus on the native
20 speaker and why investigating deficits in linguistic-cognitive SLA is essential in our field are discussed
21 as well.

22 1 The nature of comparisons in Second Language Acquisition research

23 The view that comparisons between native speakers (NS) and non-native speakers (NNS), which are
24 pervasive in Second Language Acquisition (SLA) research, should be discouraged is not new (Firth &
25 Wagner 1997; Klein 1998). Recently, however, concerns about the use of these comparisons have been
26 raised among some researchers working within the so-called linguistic-cognitive approaches to SLA¹
27 arguing that comparing learners with natives falls into a ‘Comparative Fallacy’ (CF) as described by
28 Bley Vroman (1983) and help promote a monolingual bias in our field. The CF rests on two key claims:
29 a) the linguistic system of the learner (the interlanguage grammar (ILG)) is a system on its own right
30 and b) comparisons between ILG and other systems (including the target grammar) are not legitimate
31 under any circumstances (see also Lakshmanan & Selinker 2001). We argue, however, that these two
32 claims are independent from each other. Indeed, many SLA researchers have explicitly claimed that
33 the language of the second language acquirers “represents a linguistic system in its own right and
34 should be investigated as such (Huebner 1983:33)”; this view is consistent with Selinker’s (1972)

¹ Linguistic-cognitive approaches include researchers working on cognitive–interactionist, instructed, psycholinguistically oriented, usage-based, and Generative SLA according to Ortega (2014, 2019) who claims that their “*main disciplinary goal is to illuminate the human capacity for language, and most would also share post-positivist logics, quantitative rigor, and generalizability as values in their research* (Ortega 2019:23).”

35 original conception of ‘interlanguage’ and it is widely accepted in our field. The main concern for
36 researchers (from all fields) is the legitimacy of NS-NNS comparisons (Firth & Wagner 1997; Klein
37 1998; Cook 1999)².

38 According to the CF, ILG should be investigated without being compared with NS (the control group)
39 as this may result in a view of learner grammars as ‘degenerate’ systems of less quality (i.e. the
40 comparison necessarily presupposes a prejudice against NNS). In this paper we position ourselves
41 against this view (see also Mack 1997; Montrul 2013) and argue that, despite its increasing popularity
42 among linguistic-cognitive SLA researchers, the CF is not justified on either theoretical or
43 methodological grounds. Furthermore, we propose replacing the CF with the Comparative Logic which
44 justifies comparisons with a baseline as these not only allow researchers to study L2 grammars “in
45 their own right” but are also essential in order to examine and explain the nature of L2 grammars.

46 In this paper we will present and defend the Comparative Logic on the basis of the following
47 arguments:

- 48 • **The Comparative Fallacy is routinely misunderstood.** The methodological decisions to
49 prevent the CF entail much more than not including a control group of native speakers in the
50 design of a study. We will argue that the CF, in fact, constitutes a significant barrier to providing
51 meaningful analyses and explanations and it does not support the fundamental explanatory
52 nature of the field.
- 53 • **Acquiring a language is an incremental process and learners’ grammars develop towards**
54 **a target.** The developmental nature of L2 acquisition means that L2 speakers can be situated
55 along a linguistic continuum (see a similar proposal by Polinsky & Kagan 2007) for Heritage
56 Speaker Grammars) which represents different stages of acquisition and proficiencies. ILG are
57 representations of specific points in the process of acquiring a second language as learners move
58 closer to an endpoint. Since native controls are speakers who have a complete (or endstate)
59 grammar (see Meisel 2011), it is legitimate to regard a NS grammar as a possible endpoint
60 (target) in the L2 acquisition continuum. Thus, comparisons between the current state and the
61 target (the endstate grammar) as well as the current state and (a possible) next state (i.e. NNS-
62 NNS comparisons) are necessary in order to understand the fundamental nature of ILG and L2
63 acquisition. Without such comparison the data can be described but both meaningful analyses
64 and predictions for subsequent development are virtually impossible.
- 65 • **Native controls are necessary in experimental SLA to validate the tasks.** Evidence from the
66 behavior of native controls is key as it ensures that the instruments are appropriate and that the
67 theoretical assumptions are correct. We advocate for the elicitation of data from a variety of
68 tasks so any conclusions on the nature of ILG are based on more than one source of evidence.
69 Native speakers are not chosen as the baseline because they are perfect, privileged or infallible

² Tensions between the so-called cognitive approaches to SLA (those which investigate linguistic systems) and sociocultural perspectives (those which focus on the socialisation aspect of language learning) are not new (see Zuengler & Cole 2005 for a review). A good example of the types of criticism that cognitive SLA has endured over the years can be found in the arguments put forward in Firth & Wagner (1997) and the defence in Gass (1998) and Long (1997). We see recent criticisms of the type expressed in Ortega (2014, 2019) as another turn of the screw in the quest for dismissing any serious inquiry into second language acquisition which has a theoretical interest and focuses on investigating grammatical knowledge.

70 but because they are often the control group that is methodologically appropriate. This is why
 71 native speakers undergoing attrition are not appropriate controls for SLA studies (their
 72 grammars do not represent the endstate of language acquisition anymore), although they may
 73 be appropriate controls in other contexts.

74 • **The control group needs to be appropriate for each specific study.** The control group and
 75 the experimental group need to be matched on a number of variables to ensure that they differ
 76 only with respect to the condition to be investigated. Since some variation in the behavior of
 77 NS is expected, it is essential that both groups speak the same variety (i.e. be exposed to the
 78 same evidence available in the input), have the same level of literacy (same educational
 79 background) as well as same background characteristics (see Dąbrowska (2010); Andringa,
 80 Olsthoorn, van Beuningen, Schoonen & Hulstijn (2012); Hulstijn & Andringa (2014); Hulstijn
 81 (2011)). The challenge is to decide what group of NS to include for the comparison with NNS
 82 to be appropriate.

83 Debates on the usefulness of native controls go beyond the legitimacy of the CF as many believe that
 84 comparisons between native speakers and non-native speakers are *unfair* (on moral grounds) as
 85 learners/bilingual speakers are expected to conform to native norms unfairly. This is particularly
 86 critical in the case of learners of English as this language carries added connotations of colonisation,
 87 power and privilege, notions which are not the main concern of most SLA researchers. It is important
 88 to note that the original formulation of the CF discusses comparisons between grammars (interlanguage
 89 systems) without specifically referring to the speakers of the target language (TL) as native speakers.
 90 However, one main objection clearly concerns the use of native speakers. There are various reasons
 91 for this, one of them being that ‘native speaker’ carries negative connotations outside the strict SLA
 92 remit. In particular, concepts such as NS and NNS are used to represent the people themselves even
 93 though for linguistic-cognitive SLA approaches (as well as for Bley-Vroman), the focus is on the
 94 linguistic system, not the speakers in their social context.

95 One consequence of the misunderstanding of what the object of study is (grammars vs people) is that
 96 any analysis or evaluation in deficit/error terms can erroneously be extended to the speaker as a person.
 97 In turn, this can be used to claim that SLA researchers think of learners as being deficient speakers
 98 (Ortega 2014; The Fir Group 2016). Although issues around native prestige have been debated in
 99 related fields for some time (see e.g. discussions on the superior native speaker (Phillipson 1992) and
 100 native-speakerism in English Language Teaching (Holliday 2006, 2015)), these are now emerging
 101 within our field. The extreme position goes as far as arguing that grammars (or ILG) are not a legitimate
 102 object of study (see Ortega 2019), a claim which is neither in the spirit of the original formulation of
 103 the CF nor does it fit within the main goals of our field.

104 Although we argue against using the CF to make methodological decisions in our research, we also
 105 recognise that those working on formal/cognitive SLA approaches should pay attention to the
 106 terminology employed and the rationale for including comparisons with native speakers when this is
 107 the appropriate choice.³ For instance, referring to ‘NS-NNs comparisons’ may no longer be completely
 108 appropriate in certain contexts as this is likely to be interpreted to mean that it is the speakers
 109 themselves who are being compared. It has to be clear that we are talking about systems, grammars,

³ A type of social injustice exists in Academia as English is the dominant vehicular language for transmission and dissemination of research findings. Today, most of the high-impact journals in our field overtly or covertly support a policy by which authors must comply with English native-like writing form and style (see Costello 2020 for the privilege of English in academic publishing).

110 interlanguages, and abstract grammatical representations. For that reason, we propose that in certain
111 contexts ‘end-state grammars’ instead of ‘native speaker’ can be useful to avoid this type of confusion.

112 Being mindful of how we make our research findings available to non-specialist audiences is also
113 important (see discussion in Domínguez, Hicks & Slabakova 2019), in particular when discussing
114 notions such as ‘deficits’, ‘incomplete acquisition’ and ‘not target-like’ which can be easily
115 misconstrued. Criticisms based on the CF and the monolingual bias have increased at a time when SLA
116 researchers working on theoretical issues are urged to share their research findings with people who
117 are not familiar with our goals and methods, including researchers in other disciplines, the general
118 public, funding agencies and the learners/bilingual speakers themselves. We recognise the difficulty in
119 explaining notions such as deficits and errors, incompleteness, underproduction etc. but rather than
120 rejecting the use of NS we ask to engage in a debate on how the field can overcome this challenge.

121 2 Describing, analysing and explaining L2 grammars

122 When the CF was proposed in 1983 the field of SLA looked quite different to what it does today, both
123 in terms of its goals and methodology. At the time, there was an interest in investigating the
124 systematicity of interlanguage grammars (ILG), the oral language produced by L2 speakers (Nemser
125 1971; Selinker 1972).⁴ One of the main assumptions of interlanguage studies, inspired by generative
126 studies (Sorace 1996), is that ILG are systems governed by rules. This implies that ILG are systematic,
127 although some variation in the behavior of learners is expected as well (see Andersen 1977; Hyltenstam
128 1977; Klein & Dittmar 1979; Meisel, Clahsen, & Pienemann, 1981; Clahsen, Meisel, & Pienemann
129 1983; Ellis 1985; Tarone 1979; Schachter 1986).⁵

130 Brey Vroman’s (BV) rationale for proposing the CF was based on his criticism against how
131 systematicity was being investigated at that time by studies using oral production data as evidence.
132 BV’s focus is on how researchers can best describe the ILG without involving the target language. In
133 interlanguage studies, grammars are systematic if they conform to certain rules and expectations which
134 need to be established by the researcher and are based on analyses of the target language (TL). How
135 can the researcher know what the learner is thinking or what the ‘internal logic’ of the ILG is? For BV
136 this question cannot be answered with the analytical tools employed at that time, mainly searching for
137 contexts in which a specific form *should* be used (the so-called obligatory context). Pica (1983:70)
138 explains that “*Suppliance in Obligatory Contexts (hereafter also SOC) is used to determine accurate*
139 *suppliance of morphemes in linguistic environments in which these morphemes are required in*
140 *standard English*”. The notion of SOC has been instrumental in morpheme studies (Brown 1973; Dulay
141 & Burt 1974) which have focused on tracking the emergence and use of morphological forms in
142 English. SOC was criticised at that time because it cannot reveal whether the learners have acquired
143 all patterns and distributions of use of the target forms⁶. As Pica (1983) argues, Target-like Use Analysis
144 (TUA) can provide this insight as it also includes the number of non-obligatory contexts in which the
145 target form is supplied inappropriately. A review of these two analyses by Pica (1983), however, shows

⁴ For an overview of some main interlanguage studies see Han & Tarone (2014) and Tarone (2018).

⁵ It is important to note that Selinker (1972) originally proposed that ILG can only be studied when the learner engages in oral communication, a view which was not shared by Corder (1981) who argued in favor of investigating the learner’s judgment on grammaticality as well.

⁶ Methodological issues in morpheme studies have also been raised by Andersen (1977), Hakuta (1976), Hatch & Wagner-Gough (1976), Lightbown, Spada & Wallace (1980) and Stauble (1981).

146 that when applied to the same dataset, they render different results so different interpretations can be
 147 made depending on how the target forms are quantified. The point that BV is trying to make, however,
 148 is that SOC is not learner-based as it does not emerge from observations of the learner data alone but
 149 by comparisons with the target grammar. In particular, he criticises the methodological approach
 150 described in Tarone, Frauenfelder & Selinker (1976) as it is based on “*the mistake of studying the*
 151 *systematic character of one language by comparing it to another*” as well as the fact that “*it obscures*
 152 *the internal structure of the learner’s system (p.6).*” This is the Comparative Fallacy (CF).⁷

153 BV points out a number of problems with the methodology that Tarone et al. (1976) used to investigate
 154 systematicity, including that they cannot discern subcases of obligatory context and that the
 155 applicability of their measure is unknowable because one cannot tell whether the learner is faced with
 156 a binary choice as they assume. He argues that the binary nature used in SOC studies lumps together
 157 many possible options which the learners may have entertained but which cannot be revealed by the
 158 limited nature of the options made available by the researcher. He also notes that the linguistic analysis
 159 that the researcher brings to the ILG may not be available to the learners. This implies that the mere
 160 speculation of an obligatory context is a case of a comparative fallacy. Since the description of the ILG
 161 has to be done independent of the TL, the use of obligatory context (or any assumption that a certain
 162 form should be used) is, indeed, discouraged as well. If there is no possibility of any expectation of use
 163 of a form, then other key notions such as accuracy or errors should not be used either. This, in turn,
 164 implies that even describing whether learners use a form using percentages (e.g. reporting that a certain
 165 form is used an x number of times), which is common practice in the field, has to be abandoned too.
 166 The point is that adopting the CF has a knock-on effect on the whole range of methodological choices
 167 and types of analyses available to researchers well beyond NS-NSS comparisons.

168 Other concepts and tools that should not be used for the same logic are proficiency scales (beginners,
 169 intermediate, advanced, near-native etc), omissions, overproductions, simplifications and all of the
 170 other typical characterisations of interlanguage grammars (for an example of an analysis without these
 171 concepts see Klein’s (1998) description of the Basic Variety). In fact, even investigating if a form or
 172 structure has been acquired is a case of the CF as this question already imposes a view of the learner
 173 system based on what is observed on the TL and not their internal logic. It is clear that adopting BV’s
 174 own interpretation that ILGs are systems in their own right is at odds with one of the main assumptions
 175 in our field: that L2 speakers are learners engaged in the process of learning the grammar of a second
 176 language and that in this process they entertain different linguistic systems until they reach the end-
 177 state (the target grammar).

178 For this reason, it is important to understand that adopting the CF has important methodological
 179 consequences involving the tools that researchers can or cannot use in their research. It is often the
 180 case that researchers who choose not to use comparisons with native controls still analyse the learner
 181 data in terms of accuracy and expected use/acceptance of forms, even those this necessarily assumes
 182 the existence of a baseline and, thus, promotes the CF. For instance, Schwarz (1997) agrees that UG-
 183 oriented SLA suffers from the comparative fallacy because the ILG is judged against norms from the
 184 target language. However, she also claims that “*From this perspective, that properties of the TL do not*
 185 *get acquired requires explanation.*” Implicitly, Schwarz still assumes that L2 acquisition involves
 186 acquiring features present in the grammar of another group of speakers who are not the learners (i.e.

⁷ We note that although there is no explicit mention of the native language by Bley-Vroman, in principle, comparisons between ILG and any other language (native or not) would also fall foul of the CF. This would include comparing L2 and L3 speakers, or two groups of bilingual speakers (e.g. second vs third generation heritage speakers) or even comparing the same group of learners at different points in time in a longitudinal study. The specific objection to imposing native norms on learner grammars, although related, is in fact independent of the CF even though they are usually interconnected.

187 absence of a required feature is an error). Schwartz's view, with which we agree, still constitutes a case
 188 of the comparative fallacy according to Bley-Vroman's own definition.

189 Furthermore, by focusing on descriptions of ILG only, BV avoids the fact that his proposed
 190 methodology makes it virtually impossible to provide meaningful explanations about the nature of ILG
 191 and the process of acquiring a second language. Thus, the main problem arising from adopting the CF
 192 is that it does not fulfil the explanatory goal of the field. At the time when BV proposed the CF, the
 193 focus may have been on providing descriptions of ILG but this does not meet the main goals of the
 194 field⁸ today which include to (1) describe, (2) analyse and (3) explain the process of acquiring
 195 grammatical systems (see Norris & Ortega 2000; Gass 1998). Adopting the CF in its strong form is
 196 problematic as researchers could only (1) describe ILG but not (2) analyse or (3) explain (evaluate on
 197 theoretical grounds) the evidence. This position strongly contradicts the main goals of the discipline as
 198 stated above. A soft version of the comparative fallacy is also possible: the only goal of SLA is to (1)
 199 describe and (2) analyse ILG avoiding (3) to explain (i.e. evaluate on theoretical grounds) the evidence.
 200 This version also explicitly excludes comparisons with controls and it is in line with the original spirit
 201 of the CF (i.e. to provide the right kind of descriptions emerging from the learner data only). However,
 202 it is also in contradiction with the main goal of our field as it necessarily precludes an interpretation
 203 and evaluation of any finding. For instance, if a group of learners are found to use the definite article
 204 in some contexts (a description of the data without quantifying the use by means of an obligatory
 205 contexts), we would not be able to interpret this finding to be low or high if we do not know what the
 206 expected use is as set by speakers who already have that form in their grammatical systems. The only
 207 way that research can provide meaningful and appropriate analyses of ILG and test hypotheses which
 208 investigate the acquisition process is by comparing ILG with the target grammar.

209 The Comparative Logic is the only position that can achieve the three goals of SLA research: (1)
 210 describe, (2) analyse and explain ILG. This position justifies the use of controls and comparisons
 211 between grammars from learners and a baseline on purely scientific grounds. The baseline for L2
 212 studies is often formed by native speakers but this is not necessarily always the case (e.g. two groups
 213 of learners to investigate L3 acquisition; comparing second vs third generation bilingual heritage
 214 speakers; comparing the native language of monolingual and bilingual speakers undergoing attrition
 215 etc). As we will argue in the following sections, native speakers become legitimate members of a
 216 control group because of the nature of their grammatical systems, not because they are ideal or infallible
 217 speakers. It is also possible that certain grammatical areas may be subject to a higher level of variation
 218 than others even for native speakers. This is why it is informative to collect these data from a control
 219 group in experimental SLA studies.

220 In summary, in this section we have shown two main problems with adopting the CF; first, researchers
 221 lose the main methodological tools and concepts which are necessary to analyse and explain learner
 222 grammars (error, accuracy, overproduction, etc); second, the possibility of providing meaningful
 223 explanations is virtually impossible if there is no link between the learner and the target grammars. We
 224 have argued that comparisons, including NS-NNS comparisons, are necessary to meet the goals, a
 225 position which we have called the Comparative Logic.

226 3 Development in SLA: accepting that L2 speakers are learners

⁸ See Zuengler & Miller (2006) for a discussion on the main opposing perspectives in SLA research focusing on the long-standing debate between cognitive and sociocultural approaches.

227 The view that ILG are systems in their own right, which can be traced back to at least Selinker's (1972)
 228 original definition of interlanguage, is widely adopted in our field. Bley Vroman agrees with this view
 229 as well but also argues that ILGs need to be analysed independently of any other system as this is the
 230 only way that the own logic of ILG can be revealed; for this reason, he claims that the comparison
 231 with the grammar of speakers of the target language (TL) makes ILGs degenerate versions of the native
 232 grammar. It is important to note that the word chosen by BV is 'degenerate' which means degraded,
 233 abnormal and of lower quality. In our view, degenerate is an unfortunate choice of term as it is a
 234 measure of quality (i.e. non-native speakers produce language of substandard quality) which does not
 235 naturally arise from the objective description of that system.

236 Some researchers have taken the view that if L2 grammar lacks a grammatical feature or contains an
 237 error, that means that the speakers themselves are deficient in some way (see Firth & Wagner (1997))⁹.
 238 Although this misconception has been already addressed by some (see Gass 1998) criticisms of this
 239 kind towards linguistic-cognitive SLA research still remain (Ortega 2014, 2019). Reconciliation on
 240 this matter necessarily entails an understanding of how 'deficit' is understood in linguistic-cognitive
 241 SLA and why it is important that we investigate both what learners can and cannot do in the process
 242 of acquiring a second language. Although a deficit-view of acquisition (both for first and second
 243 language acquisition) exists, this is to mean that learners make errors or show incomplete knowledge
 244 of a certain grammatical aspect of the TL, not that the learners themselves are deficient in any way.
 245 Both fossilization (Selinker 1972; Han 2004) and incompleteness (Schachter 1988, 1990; Sorace 1993)
 246 have been routinely used to describe aspects of learner grammars. These terms only make sense
 247 because ILGs are evaluated against a target (complete) grammar where target means that it represents
 248 the outcome of language acquisition under ideal input conditions (what we will characterise in the next
 249 section as the 'end state'). We have already argued that since the CF prevents researchers from making
 250 any evaluations of ILG that would conclude that the system is degenerate (incomplete or deficient),
 251 concepts which are widely used in our field such as errors, omissions, overgeneralisations,
 252 simplifications etc. would need to be abandoned as well. In our view, this is the wrong approach as we
 253 would stop using the tools that allow researchers to carry out explanatory research in second language
 254 acquisition. For this reason, it is our view that any research committed to offer precise descriptions and
 255 explanatory answers will necessarily be subject to, at least, the soft version of the comparative fallacy
 256 as any explanation arising from the description of ILG would necessarily need to address the
 257 deficit/error issue we just noted.¹⁰

258 In fact, adopting the idea that the CF exists intrinsically threatens the notion of interlanguage itself, as
 259 interlanguage was proposed as a means to account for the process involved (often shown by different
 260 stages) in learning a target language. In traditional interlanguage studies, assuming that learners
 261 develop a second language (i.e. they move closer to the TL) does not necessarily mean that an ILG is
 262 not a system in its own right but, rather, that the learner is in the process of acquiring a full grammatical
 263 system with all the features expected in that system. For instance, Spanish has grammatical gender
 264 which triggers a type of agreement between nouns, adjectives and determiners (e.g. *la gata negra*/the
 265 black female cat). Thus, it is reasonable to expect that learners of Spanish will have to learn this feature
 266 which is likely not to be present at the early stages of acquisition. Until that feature is present in their

⁹ For a similar argument in the context of heritage language acquisition see Domínguez, Hicks & Slabakova (2019).

¹⁰ As an example, the Basic Variety was proposed with this objection in mind and it is the result of a description of a learner system without references to a TL. This approach is rather limited in its explanatory power and the kinds of predictions it can make about the SLA process more generally.

267 grammar, the process of acquiring Spanish (the target grammar) can be said to be incomplete.
 268 Researchers interested in finding out how learners go about the challenge of acquiring a new feature
 269 (gender) which does not exist in their native language need to know whether learners use the right
 270 gender (masculine or feminine) appropriately. It would not be possible to do this without a reference
 271 to how gender is used by speakers of the target language.

272 In this respect, one basic assumption in SLA studies is that we are investigating a process whereby a
 273 speaker *develops* a second/n language through a specific route. An ILG represents specific points in
 274 the process of acquiring a second language (see Meisel, Clahsen & Pienemann 1982 for a discussion
 275 on developmental stages in L2 acquisition). This process necessarily entails a progression which, in
 276 turn, necessarily assumes that certain features of the target grammar can/should/will be absent. In this
 277 respect, one can argue that this represents a ‘deficit’ view of language development, similar to what is
 278 observed in the process of acquiring a native grammar in the case of children, as a learner starts the
 279 process of acquiring a language with very little knowledge of the grammar which is being learnt. Deficit
 280 in this context does not mean that the grammar of a learner is of lower quality, degraded or degenerate,
 281 nor is it an evaluation of the speakers themselves. It means that the system entertained by the learner
 282 does not (yet) show the features and properties of the target grammar. Importantly, this ‘deficient’ or
 283 ‘incomplete’ view is not in opposition to the view that learner grammars are systems in their own
 284 right¹¹. Both interpretations can be true. This point becomes clear when analysing overregularisations
 285 such as when learners use the English past tense marker *-ed* with an irregular verb (e.g. using ‘goed’
 286 instead of ‘went’). The use of ‘goed’ is both an error (i.e. it is not how the verb ‘go’ marks past tense
 287 in English) and the result of the speakers’ grammar respecting a certain grammatical principle of their
 288 own system (e.g. use ‘ed’ with all verbs to mark past tense).

289 We would like to reiterate our point that without comparisons with the TL there can be no analysis,
 290 and that without analysis there cannot be any explanations. Notions such as accuracy and errors are
 291 fundamental to understand the processing of acquiring a language in all contexts. There are numerous
 292 examples of how different SLA frameworks make notions such as accuracy and error central to their
 293 analyses. Without these, there would be no field. A good overview of some ways in which
 294 interactionist, emergentist and generative scholars measure SLA is found in Norris & Ortega (2000).
 295 For instance, these authors show that detecting the use of a form is important for interactionist
 296 approaches to SLA. However, this is not the measure used for acquisition as learners have to show that
 297 they are also able to use that form *appropriately* and fluently. The only way in which it makes sense
 298 to describe the use of a form as appropriate is if some criteria for such use has been established.

299 With regard to emergentist approaches, Norris & Ortega (p. 728) explain that accuracy is one of the
 300 main factors used for establishing the parameters of acquisition in this framework. As in the case of
 301 appropriateness above, accuracy can only be established if a comparison with a ‘correct’ use of the
 302 form is established. In these two frameworks, comparisons between learner grammars and the grammar
 303 of the TL are necessary to fulfil our goal of understanding the process underlying SLA. Interestingly,
 304 the term ‘nativelike’ is only mentioned by Norris & Ortega (p. 727) when they describe generative

¹¹ One obvious consequence arising from the developing nature of ILG is that learners are often classified according to different proficiencies. This classification assumes a comparison as well, not between learners and native speakers, but between learners and other more or less advanced learners. Since a beginner learner is only a beginner in comparison to a more proficient learner, this type of comparison should also present a case of the CF as per BV’s definition. This means that across-group comparisons between learners (not just native speakers) is to be avoided as well.

305 approaches to SLA: “*Generative linguistic studies of SLA are likely to rely almost exclusively on the*
 306 *outcomes of grammaticality judgment tasks of various kinds, where acquired means nativelike levels*
 307 *of rejection of illegal exemplars of the target grammar.*” Although we agree with these authors that the
 308 term nativelike is often used by Generative SLA studies, the same concern with the appropriate and
 309 accurate use of target form is shared by all of the frameworks reviewed by these authors. For all these
 310 researchers, the use of target forms is analysed by comparison with a group of speakers which perform
 311 target-like. That is, one fundamental notion of acquisition is that it assumes conformity with native
 312 use/judgement in all approaches. For instance, in a study promoting task-based learning, Pica, Kang &
 313 Sauro (2006:320) describe ILG as being full of omissions, substitutions, and inconsistencies and a
 314 varying degree of accuracy. They do this without explicitly comparing learner behavior with a group
 315 of native control even though this is the only one in which they can discuss accuracy. An important
 316 body of research has been concerned with the role of corrective feedback in SLA. Studies on corrective
 317 feedback assume that L2 learners make errors. For instance, Ellis, Lowen, & Erlam (2006:340) argue
 318 that “*Corrective feedback takes the form of responses to learner utterances that contain an error. The*
 319 *responses can consist of a) an indication that an error has been committed, b) provision of the correct*
 320 *target language form, or c) metalinguistic information about the nature of the error, or any*
 321 *combination of these.*” Superficially, one could conclude that the focus of the investigation is to show
 322 that learners fail to acquire a second language. Similarly, approaches which investigate NS and NNS
 323 interactions (see e.g. Lyster & Saito 2010) do so on the assumption that the NS plays a crucial role in
 324 second language development: it is through the interaction with a NS that input is rendered
 325 comprehensible to learners. Finally, when Andersen and Shirai (1994: 143) proposed the extremely
 326 influential ‘Aspect Hypothesis’ to explain the L2 acquisition of past tense morphology, they were
 327 trying to explain why learners **fail** to supply past marking in obligatory context much more frequently
 328 with some predicates than with others. The analysis of correct and incorrect compliance of target forms
 329 was the basis of Andersen and Shirai’s analysis later adopted by a large number of studies.

330 These examples show how in all of these approaches, notions such as accuracy, progress and errors are
 331 crucial if it is expected that ILG develops towards a target. As Lardiere (2003) argues, even those
 332 approaches/researchers who are supposed to be respectful of the comparative fallacy (because they
 333 claim that they investigate learners’ interlanguage on its own right) are susceptible of it once they base
 334 their analysis on notions such as obligatory context, accuracy, omissions etc. In our view,
 335 understanding and explaining SLA necessarily requires comparisons with a baseline. We have called
 336 this the Comparative Logic and have argued that it is the most appropriate position in order to both
 337 view ILG as system in their own right and provide meaningful explanations. Analysing and
 338 understanding when success is both possible and when it is not is fundamental in our field.

339 **4 Generative SLA and the role of native controls**

340 In the previous section we showed that comparisons between learner and complete (native) grammars
 341 are commonplace in the field because they are necessary to explain ILG irrespective of the theoretical
 342 framework; however, it is often the case that researches in the generative tradition are the target of
 343 criticism specifically for promoting comparisons between learner and native grammars and for the
 344 (erroneous) belief that the field sees native speaker norms as a goal. This is partly due to the fact that
 345 having evidence from native speakers’ intuitions is clearly part of the methodological design. There
 346 are other reasons which are linked to the main assumptions of the whole generative enterprise which
 347 have been carried over to SLA research. As we have already explained, Generative SLA is concerned
 348 with the abstract linguistic knowledge of speakers, what they unconsciously know about language(s).
 349 The field assumes an innate and biologically determined capacity for language which is unique to
 350 humans. The specialised and abstract module specific to language known as the computational system

351 includes a lexicon and the syntactic operation Merge (Chomsky 1998; Berwick & Chomsky 2008;
 352 Friederici 2017) which builds syntactic structures which are interpreted and pronounced by specific
 353 subsystems. Importantly, there is evidence suggesting that access to this capacity may decline with age
 354 as differences between how speakers acquire a native and a non-native language have been found (see
 355 discussion in White & Genesee 1996).

356 Unlike other cognitive approaches, generative SLA is interested in I-Language, rather than language
 357 as a social or cultural object. I-language is an internalised system, what is also known as a grammar. I-
 358 language is according to Chomsky, Gallego & Ott (2019) “*a system that links meaning and sound/sign*
 359 *in a systematic fashion, equipping the speaker with knowledge of these correlations.*” During the
 360 language acquisition process, assumed to be constrained by Universal Grammar (UG), children
 361 develop a grammar (i.e. they figure out what is correct and what isn’t) and establish form and meaning
 362 pairs as determined by the language faculty (Chomsky 1986). These form-meaning connections thus
 363 exist in the target language which serves as the input for L2 speakers. Typically, the language
 364 acquisition process finishes when children’s grammars reach the so-called ‘steady’ or ‘end state’. The
 365 ‘steady state’ is the full adult grammar resulting from full access to UG and exposure to a full set of
 366 linguistic input; in this respect, one could say that it is what results in ‘ideal conditions’ for language
 367 acquisition in the sense that full convergence with the ‘end state’ is always achieved. For this reason,
 368 we argue that a more appropriate way of calling native speakers in SLA research would be ‘end-state
 369 speakers’ or even more appropriate those who have an ‘end-state grammar’ to avoid any confusion
 370 about what the object of our study is.

371 In the context of L2 speakers, ILG is also an I-Language (see Adjemian 1976; Klein 1998). L2 speakers
 372 have access to UG¹² during the acquisition process but the characteristics of their ‘steady state’, unlike
 373 the case of children, are unclear. It is also not completely obvious whether any intermediate grammars
 374 or ILG have direct access to UG, or whether all L2 speakers reach a similar ‘steady state’ with the
 375 same characteristics.¹³ Comparing the status and characteristics of these intermediate I-languages and
 376 the corresponding ‘end-states’ is useful to evaluate the role and accessibility of UG, the role of the
 377 input during acquisition, L1 influence, etc. Even though updated views of the role of UG have
 378 promoted other types of research questions (the role of linguistic interfaces, representational
 379 impairment vs computational efficiency, feature-reassembly etc), White’s (2003) claim that “*the*
 380 *crucial question is whether or not interlanguage grammars are UG-constrained, rather than whether*
 381 *or not they are native-like*” is still valid today.

382 One specific and very common criticism against generative SLA is based on the (misinformed) claim
 383 that generative approaches to language are based on native speakers are idealised speakers (Leung,

¹² Whether UG is available for L2 acquisition has been a major topic of debate (see Epstein, Flynn & Martohardjono 1996; Borer 1996), the current view being that learners can indeed access UG when they encounter input which cannot be comprehended/parsed by their existing grammar (the Full Access position) (see White 2003).

¹³ Bley Vroman (1990, 2009) argues that the process of acquiring a first and second (foreign) language are fundamentally different. The Fundamental Difference Hypothesis (FDH) proposes that whereas child language acquisition may be constrained by UG (which is domain-specific), foreign language learning is not, and so adults need to resort to general learning mechanisms in an instance of general skill acquisition. He makes this claim on the basis of certain observations which have mostly been contested or not completely accepted including that “complete success is extremely rare, or perhaps even non-existent” (1990:6); “adults not only generally do not succeed, they also fail to different degrees (1990:7); and adults set their own goals and can fail, this leads to variation in the process and outcomes. It is interesting that despite the fact that the FDH and the Comparative Fallacy are based on a clear anti UG position, some UG scholars seem to embrace both.

384 Harris & Rampton 1997); embedded in this criticism is, again, that speaker here refers to the speaker
 385 as a person functioning in the real world, not their abstract linguistic system as we have just explained.
 386 This particular criticism often arises from a misunderstanding of what ‘ideal’ means¹⁴ and the reasons
 387 that led Chomsky to propose this assumption in the first place. The contentious quote from Chomsky
 388 (1965:3) is as follows: “*Linguistic theory is concerned primarily with an ideal speaker-listener, in a*
 389 *completely homogeneous speech-community, who knows its language perfectly and is unaffected by*
 390 *such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and*
 391 *interest, and errors (random or characteristic) in applying his knowledge of the language in actual*
 392 *performance.” This may appear to be a call for a search of the perfect speaker, which is identified with*
 393 *a native speaker (i.e. nativism equals perfection). However, Chomsky is really arguing that in order to*
 394 *understand grammar as a cognitive system (competence) one has to look further than what speakers*
 395 *actually say (performance) as this is modulated by non-linguistic factors. Chomsky is concerned with*
 396 *knowledge of a grammar as an abstraction, an outcome of language acquisition in ideal learning*
 397 *conditions, not as a real object that can be studied. In this respect, Chomsky (1965) also explains that*
 398 *“a generative grammar is not a model for a speaker or a hearer. It attempts to characterize in the most*
 399 *neutral possible terms the knowledge of the language that provides the basis for actual use of language*
 400 *by a speaker hearer.”¹⁵ Critics also ignore the fact that Chomsky later abandoned the competence-*
 401 *performance distinction (and the idealised speaker) in favor of I-language and E-language (Chomsky*
 402 *1986) as this distinction, among other reasons, can account for linguistic variation.*

403 There have been some attempts to deconstruct and even rethink the need to assume an ideal
 404 speaker/hearer both in formal and experimental contexts. For instance, Chesi & Moro (2015) discuss
 405 the competence vs performance distinction proposing that there is both an idealised native speaker and
 406 a real native speaker. The native speakers are not idealised speakers themselves but have access to the
 407 grammar which is the object of study. This distinction is also useful for explaining why the behavior
 408 of native controls does not always agree with the predictions made by linguistic theory (which are
 409 based on the most idealised competence systems). Similarly, in the SLA literature Duffield (2003)
 410 distinguishes between two types of linguistic competence (underlying and surface competence) to
 411 account for knowledge of gradient grammaticality (when a structure is more acceptable than another).
 412 Underlying competence is categorical whereas surface competence is more probabilistic as it includes
 413 several factors such as sensitivity to frequency of constructions. More recently, Slabakova, Rothman
 414 & Kempchinsky (2011) have provided empirical support for Duffield’s dual competence system and
 415 Sorace & Keller (2005) have also made a similar distinction between hard (syntactic) and soft
 416 (interface-based) constraints which yield different levels of acceptability.¹⁶

¹⁴ It is only by accident that the ‘idealised’ speaker coincides with the native speaker. Some native speakers are not appropriate as control groups in L2 studies, namely those undergoing grammatical attrition (see Domínguez 2013).

¹⁵ Traditionally, generative SLA research is often difficult to translate and apply to the real world (as opposed for instance to pedagogy-oriented SLA). As an example, it is not immediately obvious how examining UG accessibility can be of any use to foreign language teachers or even to the learners themselves.

¹⁶ It is generally agreed that grammaticality taps into speaker’s competence (or I-language) and is not open to gradience whereas acceptability involve the speakers’ performance and is gradient in nature. This is because speakers are judging sentences according to their perception of those sentences (Bard et al 1996). For a full discussion on the differences between grammaticality and acceptability plus how the parser can also affect acceptability see Leivada & Westergaard (2020). For a detailed discussion on how the interpretation of gradient judgments affects sentence acceptability see Francis (2022).

417 The use of native speakers as native controls has been often justified on methodological grounds.
 418 Research which focuses on judgement data as the main source of evidence requires a control group in
 419 the experimental design. Since the control group is very often a group of native speakers of the TL,
 420 although not exclusively, the comparison between native and non-native behavior is often made
 421 explicitly. As Sorace (1996:380) notes: “*For the correctness of judgments to be empirically assessable,*
 422 *it should be possible to measure intuitions of degree of grammaticality against some independently*
 423 *established grammaticality scale.*” Sorace’s quote shows that comparisons among groups (including a
 424 baseline of native speakers) are necessary for explaining and assessing the results arising from
 425 linguistic judgements. It is an essential part of the experimental design used in research which
 426 investigates learners’ judgements and intuitions. In this type of experiments a set of variables are
 427 defined and controlled. The control-experimental group comparison is also necessary to determine
 428 whether the results are the effect of the independent variables or not, to establish the baseline of
 429 comparison, to verify the validity of the task and for investigating whether the hypotheses are incorrect
 430 and need to be reformulated. There are numerous examples showing that this has been the case in
 431 Generative SLA. For instance, Grüter (2006) in a key study which found support for the Full
 432 Transfer/Full Access position used a control group to analyse the acquisition of wh-questions in
 433 German. The behavior of the control group is key to show that there is a bias for one of the two possible
 434 readings of a question which was not expected nor found in the L2 data. Without the native control
 435 data some of the learner behavior would have not been explained by the hypotheses.

436 If native speakers are necessary as baselines to control conditions and offer a key measure for
 437 understanding learner behavior, how can researchers meet their methodological needs and avoid the
 438 CF at the same time?¹⁷ This is definitely a challenge for UG-based research which has an explanatory
 439 goal that goes beyond providing descriptions and often elicits intuitions; in fact, such is the difficulty
 440 that we argue that it is virtually impossible. In our view, the key is to separate that comparisons between
 441 native and non-native grammars are necessary from any conclusions that researchers can reach based
 442 on those comparisons (the issue is how ‘deficit’ is/should be approached). In particular, it should be
 443 possible to investigate learner grammars on their own right whilst providing analyses which take into
 444 account the judgements of speakers of the target language. In this respect, we agree with Sorace
 445 (1996:385) that even when the comparison with native speakers are justified “*learners’ judgments*
 446 *themselves should provide the primary criterion for deciding which structures are or are not part of it*
 447 *(the non-native grammar).*” The learner’s data are still the relevant data as argued by Birdson (1989).

448 Those researchers which still choose not to include a group of native speakers as control groups need
 449 to clearly specify how they intend to provide accurate and appropriate descriptions and explanations
 450 of the learner data. For instance, Heil & López (2020) included a group of native speakers as controls
 451 but learners’ and natives’ judgments were not analysed together. The authors showed the results of the
 452 monolingual English group in order to verifying experimental validity as they wanted to avoid the CF.
 453 In the method they use, they provide evaluations of learners’ grammars based on indirect comparisons
 454 with native controls. However, it is difficult not to draw comparisons between these two groups when
 455 both sets of results are presented together in the same tables and there is a clear connection between

¹⁷ One anonymous reviewer suggests that we consider the suitability of using native speakers who are also L2 learners as controls in SLA studies on the grounds that both groups would be bilingual. It is our view that the characteristics of the control group depend on the research question to be investigated and so studies who are interested in investigating the effects of the bilingual experience should take this variable into account when selecting the controls. In some studies, having two control groups (one formed by monolingual speakers only and one formed by monolingual speakers who also know another language) may even be relevant. The reviewer’s suggestion is consistent with our view that the selection of the control group should be carefully considered for each particular study.

456 the behavior of the learners and the native speakers. Furthermore, there are studies in which a direct
 457 statistical comparison between the control and the experimental groups is justified. This comparison,
 458 which is essential in certain studies, should not be ruled out on the basis that it provides a case of the
 459 CF.

460 **5 Variability in the (native) control data**

461 One final argument against NS-NNs comparisons is that the NS themselves do not form a
 462 homogeneous group and variability in the data makes it difficult to set goals for learners based on how
 463 we expect NS to behave. In this section we argue that variation within a community of speakers and
 464 within speakers themselves is nothing unusual and has been successfully accounted for in linguistic
 465 theory. We will also show how some of the concerns raised with respect of variability can be mitigated
 466 by applying more rigorous research methods, in particular better sampling techniques.

467 Formal SLA has borrowed analytical tools from linguistic theory as researchers assume that evidence
 468 of knowledge of grammar is shown by knowing what is both grammatical and ungrammatical. There
 469 is also a long tradition of testing hypotheses in controlled, experimental settings.¹⁸ A priori it may seem
 470 that variation is problematic for a UG approach to language since UG is invariant by nature. However,
 471 variability has been accounted for by several approaches such as Adger's (2006) Combinatorial
 472 Variability model or the Multiple grammars approach (Kroch 1989, 1994; Yang 2002). Another recent
 473 development has brought together generative syntax and variationist sociolinguistics (see review in
 474 Adger, Jamieson & Smith (submitted)) and employs a new methodological approach which moves
 475 beyond the individual and focuses on both linguistic and social aspects of the whole community of
 476 speakers. Under this approach, and following Labov (1982), it is expected that the linguistic rules
 477 shared in the community are of a variable nature. Sentences which would be ungrammatical for some
 478 speakers of English can be part of the grammar of speakers of certain varieties for which the standard
 479 and the regional variety are both possible. For instance, Henry (1996) shows cases of word order
 480 variation with imperatives in Northern Irish English as shown in (1a) and (1b):

- 481 1. a. You go away
 482 b. Go you away

483 Despite the fact that intra-speaker variation is often observed as shown in example (1), there is still an
 484 expectation that speakers would conform to certain rules; that is, certain aspects of the grammar are
 485 not subject to intra-speaker variation regardless of differences in gender, class, style, education, age
 486 etc. For instance, sentence (2b) with a missing subject would not be acceptable by any speaker of
 487 English:

¹⁸ In theoretical syntax the research method is to obtain a judgement of the acceptability of a sentence often by just using the intuition of the author/s of the study. Phillips (2009) and Adger (2014) have both defended this method. Experimental evidence supporting the validity of introspective judgements can be found in Sprouse, Schütze & Almeida (2013) as they show that these data have medium to large effect sizes. Sprouse (2011) argues that the results of an acceptability judgement task conducted via AMT are almost indistinguishable from the results from the same task conducted in a lab with control from the researcher. So same quality and AMT has the key advantage that it can recruit participants for the control group from a wide pool and varied backgrounds. In contrast, Gibson & Fedorenko (2013) point out some of the problems including that this method leads to incorrect generalisations due to cognitive biases from the part of the researcher. They argue for a quantificational approach using Amazon.com's Mechanical Turk.

- 488 2. a Lena says that [she] will come soon
 489 b. *Lena says that [] will come soon

490 There is some tension between conformity and variability when investigating the linguistic behavior
 491 of speakers. We expect speakers of English to conform to core syntactic properties (such as the use of
 492 overt/null subjects) in some cases more clearly than others. It is important to highlight that cases such
 493 as the examples shown in (1) are cases of *true* variability in the speakers' grammars (I-language).
 494 However, the SLA literature also describes a type of variability which is linked to performance and to
 495 other methodologically related issues. For instance, Sorace (1996:377–378) mentions several
 496 extralinguistic factors that are likely to influence how participants go about completing grammaticality
 497 judgements including parsing strategies, context and mode of presentation, pragmatic considerations,
 498 mental states and linguistic training. Schütze (1996/2016) also shows that literacy is a relevant factor.
 499 These and other similar factors, which are external to the mental representation of the grammar, are
 500 important for SLA researchers and can affect the results arising from grammaticality/acceptability tasks
 501 giving rise to extra-linguistic variation. Researchers should try to minimise this by choosing the
 502 appropriate design and research method.

503 In particular it is important that, for some structures, researchers allow for the possibility of using
 504 gradience or a range of responses (usually a Likert scale) rather than restricting the responses to yes/no
 505 answers (see discussion in Schütze 1996/2016). In some cases, it may be necessary to elicit evidence
 506 through various types of tasks and make comparisons based on a range of answers rather than a fixed
 507 point (see e.g. Hyltenstam & Abrahamsson 2000, 2009) and how they judged the performance of L2
 508 speakers against the whole range of responses provided by native speakers). Recruiting participants to
 509 be part of the control group is an important task which needs careful attention from the part of the
 510 researcher (see Lipsey 1990; Quené 2010) so that the sample is both as homogeneous and
 511 representative as possible. Special attention needs to be paid so that the control group and the
 512 experimental group are matched on the key variables to ensure that they differ in respect of the
 513 condition to be investigated only. Other adjustments such as that both groups speak the same variety
 514 and are exposed to the same evidence available in the input should be taken into account as well.

515 It has also been argued that other factors such as processing and experience may be subject to variation.
 516 When investigating individual differences in L2 acquisition, Andringa, Olsthoorn, van Beuningen,
 517 Schoonen & Hulstijn (2012) assume variation of listening proficiency for both non-native and native
 518 speakers. They found that the success comprehension process for native speakers depends on their
 519 ability to deal with the pressure of online speech processing. Those speakers with more accumulated
 520 experience processing complex texts were the best listeners. This suggest that NS should be matched
 521 with NNS of similar literacy levels. A similar argument has been made by Hulstijn & Andriga (2014)
 522 as they argue that it may not be possible to single out a single factor responsible for variation in their
 523 native control data as effects of (working-memory capacity, reasoning ability and reaction-speed in a
 524 nonverbal task together could explain effects of age and length of exposure. In general, these and other
 525 studies investigating individual differences reach the conclusion that NS-NNS are legitimate as long
 526 as the right NS are included in terms of literacy, educational background, experience, background
 527 characteristics etc. Individual variation can also be an effect of the task. In this respect, Hulstijn
 528 (2011:236) shows how individual differences in some tasks employed are mainly restricted to
 529 differences in the speed with which linguistic information can be processed (as a function of age),
 530 whereas in other tasks it is “mainly by differences in intellectual skills and amount of reading and
 531 writing activities, as reflected by education, occupation, and leisure-time activities.”

532 Finally, the type of predictions and expected results can have an effect on the results as well. For
 533 instance, in our investigation of the use of preterite and imperfect forms in Spanish by both native and
 534 non-native speakers we asked all the participants to complete a series of oral and comprehension tasks
 535 (see Domínguez et al 2013). We investigated whether the predictions of Andersen & Shirai's (1994)
 536 Aspect Hypothesis (AH) hold for both groups so it was important to have data showing the use and
 537 acceptability of the target forms for the native speakers as well as the learners. According to the AH
 538 preterite tends to be used with telic events rather than with atelic events; on the other hand, the
 539 imperfect is preferred with atelic events. The results of two oral production tasks, an interview with an
 540 investigator and a picture-based story retell show that, despite some variation in the amount of preterite
 541 and imperfect forms produced by the controls, the averages conform to the expected results. For
 542 instance, in the interview, the least-controlled task, the native controls used the preterite with
 543 achievement (telic) verbs on average 80% of the time whereas they use this form on average 32% of
 544 the time with state (atelic) verbs. Although most native speakers used the preterit between 80%-95%
 545 of the time with achievements, the range of use was wide from 57% to 100%. The range of use of the
 546 preterite with states was equally wide from the lowest use of 7% to the highest use of 55%. Despite
 547 this variation, the means were useful as they corroborated our predictions and showed differences with
 548 the pattern of use shown by the learners. We were able to conclude that the pattern of use of preterit
 549 and imperfect predicted by the AH is already represented in the pattern of use of these forms in the
 550 native input, so learners have access to that kind of evidence though the course of acquisition.

551 In Domínguez & Arche (2014) we reported variability in the data of the (native) control group even
 552 though this was not expected. All the participants completed a content-matching acceptability task to
 553 investigate preference of SV and VS orders with different types of verbs (accusative and unergative)
 554 and different types of pragmatic contexts (narrow focus on the subject or not). The theoretical analysis
 555 adopted predicted that native controls would prefer the VS structure with narrowly focused subject
 556 with unergative verbs (smoke, dance, sneeze and cry). However, the aggregated means of all the native
 557 participants showed that these speakers only chose this structure 45% of the time. A closer look at the
 558 individual results revealed that this was not a case of optionality as native speakers had clear patterns
 559 of behavior as roughly half of them preferred SV and roughly other half preferred VS in this context.
 560 Interestingly, the advanced learner group also showed variability in their responses but in this case, the
 561 same participant would choose both options. Unlike the native controls, learners did show optionality
 562 in their responses. Based on the responses of the native control data we were able to suggest that the
 563 input can be vague with respect to SV and VS structures in Spanish which can lead to difficulties
 564 (optionality) for learners.

565 In this section we have argued that variation in the data is not unexpected and can be accounted for
 566 both theoretically and empirically. A more careful selection process for the control group can mitigate
 567 problems arising from extra-linguistic variation and ensure that the sample is representative and
 568 appropriate.

569 **6 Will these issues ever be resolved? Some reflections for the future**

570 A review by Zuengler & Cole (2005) shows that criticisms against the goal and methodology employed
 571 by cognitive approaches to SLA have been raised for quite some time. In that review it was clear that
 572 the criticism came from scholars from the socio-cultural tradition (e.g. Firth & Wagner 1997). It is now
 573 the case, however, that questions on the role of the native speaker are being asked from within the
 574 cognitive field. We have analysed Bley-Vroman's Comparative Fallacy and examined the validity of
 575 its assumptions in the context of SLA research today. We have concluded that by ignoring the target
 576 grammar, the CF does not enable researchers to achieve the main goals of our field. This is because

577 making methodological choices on the basis of the CF entails much more than not including control
 578 groups of native speakers. Those who choose to avoid the CF would not be able to make any a priori
 579 predictions that would impose their own analysis/expectations on the learner data; they would not be
 580 able to analyse the data in terms of what is not produced, whether forms are absent, or overused or
 581 simplified etc. The analysis they produce will not be able to make references to errors or accuracy
 582 either. Since these are notions which are essential to account for the nature of the acquisition process,
 583 we conclude that adopting the CF will prevent researchers from providing meaningful explanations on
 584 learners' grammar.

585 We have also argued that the only position which ensures that the goals of the field are met (describing,
 586 analysing and explaining the process of learning a second language) is the Comparative Logic, the
 587 view that comparisons with a control group or baseline, are necessary. The field, almost 30 years after
 588 BV proposed the CF, is well-equipped to make comparisons between learner and non-learner grammars
 589 in a way that respects the principle that ILGs are systems in their own right. Nevertheless, careful
 590 attention needs to be paid to the methodology chosen and, in particular, the sampling process for
 591 inclusion of participants in the control group or the baseline for comparison. Researchers should
 592 consider not just what types of tasks to employ but also how variation in the cognitive skills, literacy,
 593 experience etc of the participants in the control group could lead to variability in the results.

594 We have also argued, as others have before us, that for cognitive SLA errors are an important source
 595 of information when investigating the learners' mental grammars. SLA is a process by which learners
 596 entertain different interlanguages or I-languages which may not include all of the features of the target
 597 grammar until they reach the 'end' or 'steady state'. The 'steady state' is the adult grammar which
 598 results from the interaction of UG, exposure to input and certain cognitive principles during child
 599 language acquisition. Interlanguage is a type of I-language, an abstract, subconscious and internalised
 600 grammar with characteristics similar to learner grammars (ILG). For this reason, we completely agree
 601 with Gass (1998:84) when she claims that the scope of inquiry of SLA is to study acquisition and so
 602 L2 speakers in this context are necessarily *learners* and not *users* of the language¹⁹. In the same spirit,
 603 we emphasised in this article that in order to answer relevant questions about the nature of ILG we
 604 need to focus on the grammatical systems and not the speakers. Crucially, our enterprise does not
 605 preclude others from studying social aspects associated with learning a second language.

606 We are mystified that anyone could conclude that our field *promotes* native speaker norms and that
 607 there is a monolingual bias in SLA (see e.g. Kachru 1994). We hope that this article has shown that
 608 there is no privileged status or prestige associated with the notion of a native speaker per se, nor that
 609 native speakers are a model or inspiration for learners (see Davies (2003) for this view). Criticisms of
 610 this sort are particularly common when Generative SLA is targeted, which is often criticised for
 611 focusing too much on correctness and the native norms. We have shown that this is due to a
 612 misunderstanding of our goals and scope of inquiry. Since the emphasis is of Generative SLA is on
 613 understanding grammars (as opposed to communication or language use) and we directly judge
 614 learners' intuitions as grammatical or not, some may think that the field sees correctness as a goal when
 615 this is not clearly the case. Nevertheless, we admit that there needs to be more clarity from our part on
 616 our goals and methods, particularly when sharing our research with non-experts. In this sense, a clearer

¹⁹ The user-learner distinction is key to understand existing opposing approaches to English language learning and teaching: whereas SLA researchers investigate how L2 speakers go about learning a language, English as a Lingua Franca (ELF) researchers view these speakers as users engaged in communicative practices (see Seidlhofer 2001; Jenkins, Cogo, Dewey 2011)). We would like to emphasise that in our view a learner of a language can also be a user of that language in other contexts and that both approaches are not mutually exclusive.

617 rewording of our research questions would be a step forward. For instance, Generative SLA does not
 618 investigate *if an L2 speaker can become a native speaker* but rather *if an end or steady-state grammar*
 619 *can be attained based on partial input after the onset of the critical period*. The problem we see with
 620 this is that the latter is harder to understand and it is not as attractive as the former, particularly as
 621 researchers are under pressure to get funding, make our research impactful to non-specialists and seek
 622 collaborations with other disciplines.²⁰

623 We believe that this is a serious issue for cognitive approaches to SLA and Generative approaches in
 624 particular. Although some good attempts to made formal SLA useful to foreign language teaching exist
 625 (Whong, Gill & Marsden 2013; Leal & Slabakova 2019; Rankin & Whong 2020), a large body of our
 626 research does not have an immediate application outside the academic remit, mostly because our
 627 concerns are theoretical in nature. This may be seen as a limitation compared to other approaches,
 628 when it clearly is not, nor does it justify a radical methodological change. Without research which
 629 engages with theoretical questions, there cannot be any scientifically inspired applications. Gregg
 630 (1996:75) already cautioned that L2 theories may only have intellectual value since the problems
 631 tackled are fundamentally theoretical (as opposed to practical problems). Furthermore, Newmeyer and
 632 Weinberger (1988) also argue that “*progress in L2 acquisition theory, as in any other scientific*
 633 *discipline, comes by focusing on the explanatory problem, and not by looking over one’s shoulder at*
 634 *the possible applications*.” The apparent (lack of) immediate applicability issue has become quite real
 635 recently for researches working on theoretical issues. As pressure mounts to make our results
 636 meaningful and impactful in the real world we make ourselves vulnerable as opportunities for
 637 misunderstanding multiply. Something as simple as proposing as a vision of SLA based on
 638 transdisciplinarity (The Douglas Fir Group 2016)²¹ is likely to instigate even more criticism against
 639 cognitive and formal approaches to SLA as we are singled out for not taking into account the learners’
 640 social context and that they are people who function in the real world. It is in the sense that
 641 transdisciplinary in the SLA context is a trap and not a vision all researchers see as beneficial for the
 642 field (see also Han 2016).

643 7 Conflict of Interest

644 The authors declare that the research was conducted in the absence of any commercial or financial
 645 relationships that could be construed as a potential conflict of interest.

646 8 Author Contributions

647 LD and MJA contributed to conception and design of the study. LD wrote the first draft of the
 648 manuscript. Both authors contributed to manuscript revision, read, and approved the submitted version.

649 9 References

²⁰ We think of this situation as a paradox. The field is urged to open up and make its research impactful to others as a way to prosper and become relevant but by doing so researchers accidentally create opportunities for misunderstanding, criticism and alienation (for instance not understanding that we are interested in analysing grammars and not people in real situations). This, in turn, raises concerns about our theoretical and methodological assumptions which can make our field less prosperous and less relevant in the end.

²¹ The transdisciplinary agenda “*seeks to integrate the many layers of existing knowledge about the processes and outcomes of additional language learning by deriving coherent patterns and configurations of findings across domains* (The Douglas Fir Group 2016:20)”. The starting point of this agenda is the social-local reality of L2 learners.

- 650 Adger, D. (2006). Combinatorial variability. *Journal of linguistics*, 503-530.
- 651 Adger, D. (2014). Syntax. *WIREs Cognitive Science* 6 (2), 131-147.
- 652 Adger, D., Jamieson, E & Smith, J. (submitted) Sociolinguistics and Minimalist Syntax. In K
653 Grohmann & E. Leivada, (Eds.) *The Cambridge Handbook of Minimalism*. Available at: [lingb](#)
- 654 Adjemian, C. (1976). On the nature of interlanguage systems. *Language learning*, 26(2), 297-320.
- 655 Andersen, R. W. (1977). The Impoverished State of Cross-Sectional Morpheme
656 Acquisition/Accuracy Methodology. *Working Papers on Bilingualism*, No. 14.
- 657 Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition
658 principles. *Studies in second language acquisition*, 133-156.
- 659 Andringa, S., Olsthoorn, N., van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants
660 of success in native and non-native listening comprehension: An individual differences
661 approach. *Language Learning*, 62, 49-78.
- 662 Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic
663 acceptability. *Language* 72(1), 32-68.
- 664 Berwick, R. C., & Chomsky, N. (2008). Poverty of the stimulus' revisited: Recent challenges
665 reconsidered. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 30,
666 No. 30).
- 667 Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Springer-Verlag
668 Publishing.
- 669 Bley-Vroman, R. (1983). The Comparative Fallacy in interlanguage studies: the case of
670 systematicity. *Language learning*, 33(1), 1-17.
- 671 Bley-Vroman, R. (1990) The logical problem of foreign language learning. *Linguistic*
672 *Analysis* 20:3-49
- 673 Bley-Vroman, R. (2009). The evolving context of the Fundamental Difference Hypothesis. *Studies*
674 *in Second Language Acquisition*, 31(2), 175-198.
- 675 Borer, H. (1996). Access to Universal Grammar: the real issues. *Behavioral and Brain*
676 *Sciences*, 19(4), 718-720.
- 677 Brown, R. (1973). *A first language: The early stages*. London: George Allen & Unwin.
- 678 Chesi, C., & Moro, A. (2015). The subtle dependency between competence and performance. *Angel*
679 *J. Gallego & Dennis Ott (Eds.), 50 years later: Reflections on Chomsky's Aspects.*
680 *MITWPL* 50, 33-45.
- 681 Chomsky, N. (1965). Persistent topics in linguistic theory. *Diogenes*, 13(51), 13-20.
- 682 Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing
683 Group.
- 684 Chomsky, N. (1998). Some observations on economy in generative grammar. *Is the best good*
685 *enough*, 1(1), 115-127.
- 686 Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language:
687 Insights, questions, and challenges. *Catalan Journal of Linguistics*, 229-261.
- 688 Clahsen, H., Meisel, J. M., & Pienemann, M. (1983). *Deutsch als Zweitsprache: Der Spracherwerb*
689 *ausländischer Arbeiter* (Vol. 3). John Benjamins Publishing Company.
- 690 Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press, USA.

- 691 Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL quarterly*, 33(2),
692 185-209.
- 693 Costello, E. (2020). ‘Requires proofing by a native speaker’–colonization and
694 scholarship. *Insights*, 33, 1-7.
- 695 Dąbrowska, E. (2010). Naive v. expert intuitions: An empirical study of acceptability judgments.
- 696 Domínguez, L. (2013). *Understanding Interfaces: Second language acquisition and first language*
697 *attrition of Spanish subject realization and word order variation* (Vol. 55). Amsterdam: John
698 Benjamins Publishing.
- 699 Domínguez, L., & Arche, M. J. (2014). Subject inversion in non-native Spanish. *Lingua*, 145, 243-
700 265.
- 701 Domínguez, L., Tracy-Ventura, N., Arche, M. J., Mitchell, R., & Myles, F. (2013). The role of
702 dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism*, 16(3),
703 558-577.
- 704 Domínguez, L., Hicks, G., & Slabakova, R. (2019). Terminology choice in generative acquisition
705 research: The case of “incomplete acquisition” in heritage language grammars. *Studies in*
706 *Second Language Acquisition*, 41(2), 241-255.
- 707 Duffield, N. (2003). Measures of competent gradience’. *The Lexicon Syntax Interface in Second*
708 *Language Acquisition*, 30, 97.
- 709 Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition
710 1. *Language learning*, 24(1), 37-53.
- 711 Ellis, R. (1985). Sources of variability in interlanguage. *Applied linguistics*, 6(2), 118-131.
- 712 Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the
713 acquisition of L2 grammar. *Studies in second language acquisition*, 28(2), 339-368.
- 714 Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical
715 and experimental issues in contemporary research. *Behavioral and Brain Sciences*, 19(4), 677-
716 714.
- 717 Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in
718 SLA research. *The modern language journal*, 81(3), 285-300.
- 719 Francis, E. J. (To appear, January 2022). *Gradient Acceptability and Linguistic Theory*. Oxford:
720 Oxford University Press.
- 721 Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. MIT
722 Press.
- 723 Gass, S. (1998). Apples and oranges: or, why apples are not orange and don't need to be a response
724 to Firth and Wagner. *The Modern Language Journal*, 82(1), 83-90.
- 725 Gibson, E., & Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics
726 research. *Language and Cognitive Processes*, 28(1-2), 88-124.
- 727 Gregg, K. R. (1996). The logical and developmental problems of second language
728 acquisition. *Handbook of second language acquisition*, 49, 81.
- 729 Gruter, T. (2006). Another take on the L2 initial state: Evidence from comprehension in L2
730 German. *Language Acquisition*, 13(4), 287-317.
- 731 Heil, J., & López, L. (2020). Acquisition without evidence: English infinitives and poverty of
732 stimulus in adult second language acquisition. *Second Language Research*, 36(4), 415-443.

- 733 Hakuta, K. (1976). A case study of a Japanese child learning English as a second language 1,
734 2. *Language learning*, 26(2), 321-351.
- 735 Han, Z. (2004). *Fossilization in adult second language acquisition* (Vol. 5). Multilingual Matters.
- 736 Han, Z. (2016). A "Reimagined SLA" or an Expanded SLA? A Rejoinder to The Douglas Fir Group
737 (2016). *The Modern Language Journal*, 100(4), 736-740.
- 738 Han, Z., & Tarone, E. (Eds.). (2014). *Interlanguage: Forty years later* (Vol. 39). John Benjamins
739 Publishing Company.
- 740 Hatch, E., & Wagner-Gough, J. (1976). Explaining sequence and variation in second language
741 acquisition. *Language Learning*, 4, 39-47.
- 742 Henry, A. (1996). Imperative inversion in Belfast English. *Amsterdam Studies in the Theory and
743 History of Linguistic science* Series 4, 79-94.
- 744 Holliday, A. (2006). Native-speakerism. *ELT journal*, 60(4), 385-387.
- 745 Holliday, A. (2015). Native-speakerism: Taking the concept forward and achieving cultural belief.
746 In *(En) countering native-speakerism* (pp. 11-25). Palgrave Macmillan, London.
- 747 Huebner, T. (1983). Linguistic systems and linguistic change in an interlanguage. *Studies in Second
748 Language Acquisition*, 33-53.
- 749 Hulstijn, J. H., & Andringa, S. (2014). The effects of age and level of education on the ability of
750 adult native speakers of Dutch to segment speech into words. In *Above and Beyond the
751 Segments* (pp. 152-164). John Benjamins.
- 752 Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for
753 research and suggestions for second-language assessment. *Language Assessment
754 Quarterly*, 8(3), 229-249.
- 755 Hyltenstam, K. (1977). Implicational patterns in interlanguage syntax variation. *Language
756 learning*, 27(2), 383-410.
- 757 Hyltenstam, K., & Abrahamsson, N. (2000). Who can become native-like in a second language?
758 All, some, or none? On the maturational constraints controversy in second language
759 acquisition. *Studia linguistica*, 54(2), 150-166.
- 760 Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language:
761 Listener perception versus linguistic scrutiny. *Language learning*, 59(2), 249-306.
- 762 Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a
763 lingua franca. *Language teaching*, 44(3), 281-315.
- 764 Klein, W. (1998). The contribution of second language acquisition research. *Language
765 learning*, 48(4), 527-549.
- 766 Dittmar, N., & Klein, W. (1979). *Developing Grammars: The Acquisition of German Syntax by
767 Foreign Workers: With 9 Figures*. Springer.
- 768 Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language variation and
769 change*, 1(3), 199-244.
- 770 Kroch, A. (1994). Morphosyntactic variation. In *Papers from the 30th regional meeting of the
771 Chicago Linguistics Society: Parasession on variation and linguistic theory* (Vol. 2, pp. 180-
772 201). Chicago Linguistics Society Chicago.
- 773 Labov, W. (1982). Objectivity and commitment in linguistic science: The case of the Black English
774 trial in Ann Arbor. *Language in society*, 165-201.

- 775 Lakshmanan, U., & Selinker, L. (2001). Analysing interlanguage: how do we know what learners
776 know?. *Second Language Research*, 17(4), 393-420.
- 777 Lardiere, D. (2003). Revisiting the comparative fallacy: A reply to Lakshmanan and Selinker,
778 2001. *Second Language Research*, 19(2), 129-143.
- 779 Leal, T., & Slabakova, R. (2019). The relationship between L2 instruction, exposure, and the L2
780 acquisition of a syntax–discourse property in L2 Spanish. *Language Teaching Research*, 23(2),
781 237-258.
- 782 Leivada, E., & Westergaard, M. (2020). Acceptable ungrammatical sentences, unacceptable
783 grammatical sentences, and the role of the cognitive parser. *Frontiers in psychology*, 11, 364.
- 784 Leung, C., Harris, R., & Rampton, B. (1997). The idealised native speaker, reified ethnicities, and
785 classroom realities. *Tesol Quarterly*, 31(3), 543-560.
- 786 Lightbown, P., Spada, N., & Wallace, R. (1980). Some effects of instruction on child and adolescent
787 ESL learners. *Research in second language acquisition*, 162-172.
- 788 Lipsey, M. W. (1990). *Design sensitivity* (Vol. 370). Newbury Park, CA: Sage.
- 789 Long, M. H. (1997). Construct validity in SLA research: A response to Firth and Wagner. *The Modern
790 Language Journal*, 81(3), 318-323.
- 791 Lyster, R., & Saito, K. (2010). Interactional feedback as instructional input: A synthesis of
792 classroom SLA research. *Language, Interaction and Acquisition*, 1(2), 276-297.
- 793 Mack, M. (1997). The monolingual native speaker: not a norm, but still a necessity.
- 794 Meisel, J. M. (2011). *First and second language acquisition: Parallels and differences*. Cambridge
795 University Press.
- 796 Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in
797 natural second language acquisition. *Studies in second language acquisition*, 3(2), 109-135.
- 798 Montrul, S. (2013). *The Ups and Downs of the Monolingual Bias and the Bilingual Turn in Second
799 Language Acquisition: Getting Down to Some Specifics*. Talk presented at the Seminar on
800 Second Language Acquisition Stockholm University, Sweden, 5 November 2013.
- 801 Nemser, W. (1971). Approximative systems of foreign language learners. *IRAL* 9, 115-123.
- 802 Newmeyer, F. J., & Weinberger, S. H. (1988). The ontogenesis of the field of second language
803 learning research. In *Linguistic theory in second language acquisition* (pp. 34-45). Springer,
804 Dordrecht.
- 805 Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C J. Doughty & M. C. Long (Eds).
806 *The handbook of second language acquisition*, 716-761.
- 807 Ortega, L. (2014). Ways forward for a bi/multilingual turn in SLA. In S. May (Ed.) *The
808 multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 32–53). New
809 York: Routledge/Taylor & Francis.
- 810 Ortega, L. (2019). SLA and the study of equitable multilingualism. *The Modern Language
811 Journal*, 103, 23-38.
- 812 Phillips, C. (2009). Should we impeach armchair linguists. *Japanese/Korean Linguistics*, 17, 49-64.
- 813 Phillipson, R. (1992). ELT: the native speaker's burden?. *ELT journal*, 46(1), 12-18.
- 814 Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second
815 language data. *Studies in second language acquisition*, 69-78.

- 816 Pica, T., Kang, H. S., & Sauro, S. (2006). Information gap tasks: Their multiple roles and
 817 contributions to interaction research methodology. *Studies in Second Language Acquisition*,
 818 301-338.
- 819 Polinsky, M., & Kagan, O. (2007). Heritage languages: In the ‘wild’ and in the
 820 classroom. *Language and linguistics compass*, 1(5), 368-395.
- 821 Quené, H. (2010). How to design and analyze language acquisition studies. *Experimental methods*
 822 *in language acquisition research*, 27, 269.
- 823 Rankin, T., & Whong, M. (2020). *Grammar: A Linguists' Guide for Language Teachers*.
 824 Cambridge University Press.
- 825 Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua
 826 franca. *International journal of applied linguistics*, 11(2), 133-158.
- 827 Schachter, J. (1986). In search of systematicity in interlanguage production. *Studies in second*
 828 *language acquisition*, 119-133.
- 829 Schachter, J. (1988). Second language acquisition and its relationship to Universal
 830 Grammar. *Applied linguistics*, 9(3), 219-235.
- 831 Schachter, J. (1990). On the issue of completeness in second language acquisition. *Interlanguage*
 832 *studies bulletin (Utrecht)*, 6(2), 93-124.
- 833 Schwartz, B. D. (1997). On the basis of the Basic Variety. *Second Language Research*, 13(4), 386-
 834 402.
- 835 Schütze, C. T. (1996/2016). *The Empirical Base of Linguistics: Grammaticality Judgments and*
 836 *Linguistic Methodology*. Berlin: Language Science Press.
- 837 Selinker, L. (1972). Interlanguage. *IRAL* 10 (3), 209-231.
- 838 Slabakova, R., Rothman, J., & Kempchinsky, P. (2011). Gradient competence at the syntax-
 839 discourse interface. *EuroSLA Yearbook*, 11(1), 218-243.
- 840 Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non native
 841 grammars of Italian. *Second Language Research*, 9(1), 22-47.
- 842 Sorace, A. (1996): The use of acceptability judgments in L2 acquisition research. In W. Ritchie and
 843 T. Bhatia (Eds), *The Handbook of Second Language Acquisition*. New York: Academic Press,
 844 375–409.
- 845 Sorace, A., & Keller, F. (2005). Gradience in linguistic data. *Lingua*, 115(11), 1497-1524.
- 846 Sprouse, J., Schütze, C. T., & Almeida, D. (2013). A comparison of informal and formal
 847 acceptability judgments using a random sample from Linguistic Inquiry 2001–
 848 2010. *Lingua*, 134, 219-248.
- 849 Stauble, A. M. E. (1981). *A comparative study of a Spanish-English and Japanese-English second*
 850 *language continuum: Verb phrase morphology*. University of California, Los Angeles.
- 851 Tarone, E. (1979). Interlanguage as Chameleon 1. *Language learning*, 29(1), 181-191.
- 852 Tarone, E. (2018). Interlanguage. In. CA Chapelle. *The Encyclopedia of Applied Linguistics*, 1-7.
- 853 Tarone, E., Frauenfelder, U. H., & Selinker, L. (1976). Systematicity/variability and
 854 stability/instability in interlanguage systems. *Language Learning*, 93-134.
- 855 Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *The*
 856 *Modern Language Journal*, 100(S1), 19-47.

- 857 White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University
858 Press.
- 859 White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in
860 adult second language acquisition. *Second language research*, 12(3), 233-265.
- 861 Whong, M., Gil, K.-H., & Marsden, H. (2013). *Universal Grammar and the second language*
862 *classroom*. Dordrecht, The Netherlands: Springer
- 863 Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford University Press on
864 Demand.
- 865 Zuengler, J., & Cole, K. (2005). Language socialization and second language learning. *Handbook of*
866 *research in second language teaching and learning*, 1, 301-316.
- 867 Zuengler, J., & Miller, E. R. (2006). Cognitive and sociocultural perspectives: Two parallel SLA
868 worlds? *TESOL quarterly*, 40(1), 35-58.