From corpora to experiments: investigating word order at the interfaces in SLA

This presentation shows how experimental and corpus data can be (and should be) triangulated to fully understand interlanguage grammars at the interfaces (lexicon-syntax, syntax-discourse, syntax-phonology). We do so by investigating inversion: Subject-Verb (SV) and Verb-Subject (VS) alternations in L1 Spanish–L2 English.

The corpus study: Departing from learner corpus research methods (Myles 2015), we used the ICLE corpus (Granger et al. 2009) and its equivalent English native corpus (LOCNESS). We briefly report on an L1 Spanish–L2 English study on word order alternations: preverbal subjects (S-V) vs. postverbal subjects ((XP-)V-S). Results revealed that subject position in L2 English is constrained by the same interface principles as in native English (see Lozano and Mendikoetxea 2010). The corpus data thus confirmed a well-known finding in previous experimental studies at the *lexicon-syntax* interface: postverbal subjects are restricted to a subset of intransitive verbs (unaccusatives of existence/appearance/change of location). But the corpus data revealed two facts that had gone unnoticed in experimental studies, namely, that SV/VS is also constrained at the *syntax-discourse* interface (the information status of the subject: topic/focus) and at the *syntax-phonology* interface (its weight: heavy/light). Postverbal subjects are focus and heavy, while preverbal subjects are normally topic and light, which entails that learners are sensitive to the external interface constrains, as in native English (contra the Interface Hypothesis, cf. Sorace 2011). The deficits learners show are to do rather with the syntactic realization of the preverbal constituent (XP). They optionally insert a null expletive (Ø-V-S) [1] or overuse an expletive *it* as the generic expletive (*it*-V-S) [2], while the production of grammatical expletive *there*, which is relatively low (*there*-V-S) [3], and the production of locative inversion (PP-V-S) [4], which is high, are target-like.

The experiment: as a follow-up to the corpus study, an online experiment was designed to (dis)confirm the corpus data and to specifically test L1 Spanish–L2 English learners' knowledge of the preverbal XP. Learners had to judge, on a five-point Likert scale, the acceptability of 32 contextualised sentences which, crucially, were modelled and designed after those produced by similar learners in the corpus studies: preverbal XP (*Ø/it/there/PP*) + intransitive V (unaccusative/unergative) + postverbal heavy & focus subject. 4 of those verbs were our top inversion verbs in the corpus (unaccusatives: *exist, appear, begin and come),* while the other four where verbs for which no inversion structures were found *(*unergatives: *talk, work, play and speak).* Results from learners at all levels of proficiency plus an English control group (total n=417) show a very robust pattern, which mostly matches the results obtained in the corpus study but also show patterns that the corpus could not reveal: While learners significantly prefer VS with unaccusatives to unergatives at all proficiency levels (Fig. 1), acceptance of unaccusatives unexpectedly decreases with proficiency. Importantly, learners show optional behaviour in the selection of the relevant preverbal expletive, which leads to variable selection of ungrammatical expletives (*it*, *Ø) as well as their grammatical counterpart (*there*). But they discriminate grammatical (*there*, PP) vs ungrammatical (*Ø, it*) preverbal XP with unaccusative VS (Fig. 2A), thus showing early sensitivity to grammatical XP, but ungrammatical XP decreases with proficiency. By contrast, they show sensitivity to the ungrammaticality of unergative XPs only with increasing proficiency (Fig. 2B).

We will discuss how these and other findings on optional expletive selection have significant implications for a theory of SLA at the interfaces as proposed by, e.g., Sorace & Serratrice 2009, Sorace 2011, 2012 (but see Montrul 2011, Slabakova & Ivanov 2011, White 2009, 2011). Finally, we argue that corpus-based learner results can, and should, be validated against corpus-external findings (see Gilquin & Gries 2009).

[1] … * because *exist* the science technology and industrialisation. [*Ø*-V-S]
[2] … * **it** *has appeared* some cases of women who have killed their husbands. [*it*-V-S]
[3] … **there** *exist* hidden tactics to seize the money of others. [*there*-V-S]
[4] **In the main plot** *appear* the main characters: Volpone and Mosca. [*PP*-V-S]

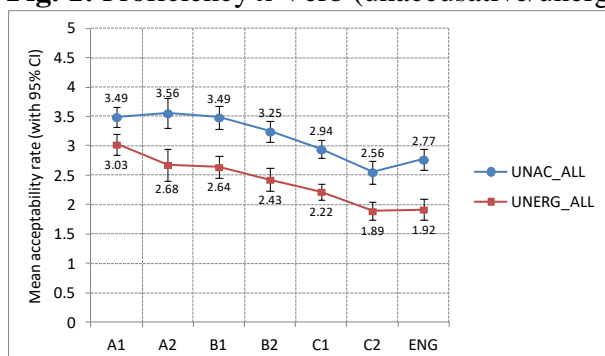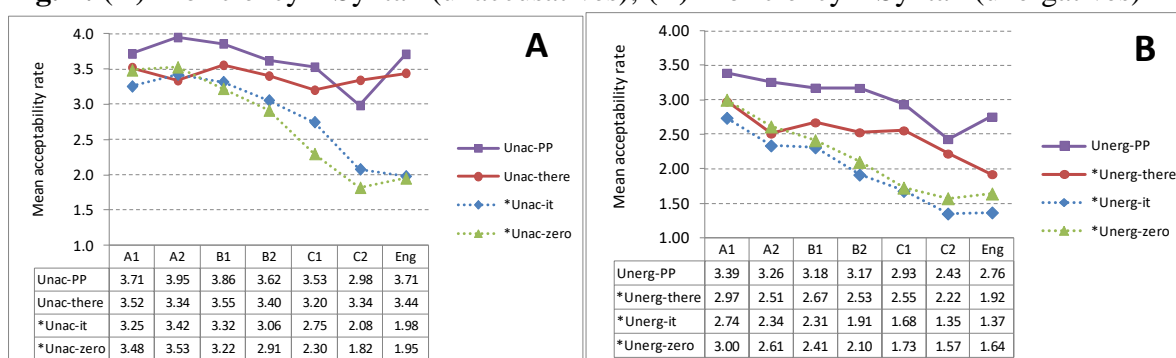**Fig. 1:** Proficiency x Verb (unaccusative/unergative)

Fig. 2: (**A**) Proficiency x Syntax (unaccusatives); (**B**) Proficiency x Syntax (unergatives)

	A1	A2	B1	B2	C1	C2	Eng
Unac-PP	3.71	3.95	3.86	3.62	3.53	2.98	3.71
Unac-there	3.52	3.34	3.55	3.40	3.20	3.34	3.44
*Unac-it	3.25	3.42	3.32	3.06	2.75	2.08	1.98
*Unac-zero	3.48	3.53	3.22	2.91	2.30	1.82	1.95

	A1	A2	B1	B2	C1	C2	Eng
Unerg-PP	3.39	3.26	3.18	3.17	2.93	2.43	2.76
*Unerg-there	2.97	2.51	2.67	2.53	2.55	2.22	1.92
*Unerg-it	2.74	2.34	2.31	1.91	1.68	1.35	1.37
*Unerg-zero	3.00	2.61	2.41	2.10	1.73	1.57	1.64

Gilquin, Gaëtanelle & Gries, Stephan (2009). Corpora and experimental methods: a state-of-the-art review. *Corpus Linguistics and Linguistic Theory* 5(1). 1-26.

Granger, S., Dagneaux, E., Meunier, F., & Paquot, M. (2009). International Corpus of Learner English version 2. Louvain: Presses universitaires de Louvain.

Lozano, C. & Mendikoetxea, A. (2010). Interface conditions on postverbal subjects: a corpus study of L2 English. *Bilingualism: Language and Cognition*, 13(4):475-497.

Montrul, S. (2011). Multiple interfaces and incomplete acquisition. *Lingua*, 121(4), 591–604.

Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 309–332). Cambridge: Cambridge University Press.

Slabakova, R., & Ivanov, I. (2011). A more careful look at the syntax-discourse interface. *Lingua*, 121(4), 637–651.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: beyond structural overlap. *International Journal of Bilingualism*, 13(2), 195-210.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33.

Sorace, A. (2012). Pinning down the concept of "interface" in bilingualism: a reply to peer commentaries. *Linguistic Approaches to Bilingualism*, 2(2), 209–216.

White, L. (2009). Grammatical theory: Interfaces and L2 knowledge. In W. C. Ritchie & T. K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition* (pp. 49-68). Bingley: Emerald.

White, L. (2011). Second language acquisition at the interfaces. *Lingua*, 121(4), 577–590.