Est. 1841 | YORK ST JOHN UNIVERSITY

Hall, Christopher J, Joyce, Jack and Robson, Chris (2016) Investigating the lexico-grammatical resources of a non-native user of English: The case of can and could in email requests. Applied Linguistics Review, 8 (1). pp. 35-59.

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/1646/

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version: https://www.degruyter.com/view/journals/alr/8/1/article-p35.xml?language=en

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repositories Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Christopher J. Hall*, Jack Joyce and Chris Robson

Investigating the lexico-grammatical resources of a non-native user of English: The case of *can* and *could* in email requests

DOI 10.1515/applirev-2016-1001

Abstract: Individual users of English as a first or second language are assumed to possess or aspire to a monolithic grammar, an internally consistent set of rules which represents the idealized norms or conventions of native speakers. This position reflects a deficit view of L2 learning and usage, and is at odds with usage-based approaches to language development and research findings on idiolectal variation. This study problematizes the assumption of monolithic ontologies of grammar for TESOL by exploring a fragment of genre-specific lexico-grammatical knowledge (the *can you/could you* V construction alternation in requests) in a single non-native user of English, post-instruction. A corpus sample of the individual's output was compared with the input he was exposed to and broader norms for the genre. The analysis confirms findings in usage-based linguistics which demonstrate that an individual's lexico-grammatical knowledge constitutes an inventory of constructions shaped in large part by distributional patterns in the input. But it also provides evidence for idiosyncratic preferences resulting from exemplar-based inertia in production, suggesting that input is not the sole factor. Results are discussed in the context of a "plurilithic" ontology of grammar and the challenges this represents for pedagogy and teacher development.

Keywords: usage-based linguistics, plurilithic Englishes, second language acquisition (SLA), English as a Lingua Franca (ELF), idiolect

1 Introduction

The project reported here is motivated by the observation that the conceptualization of grammar underlying much pedagogical thinking in TESOL is

***Corresponding author: Christopher J. Hall,** Department of Languages and Linguistics, York St. John University, York, United Kingdom, E-mail: c.hall@yorksj.ac.uk
http://orcid.org/0000-0001-9038-1238
Jack Joyce, Department of Sociology, University of York, York, United Kingdom
Chris Robson, Department of Language and Linguistics, Malmo Hogskola, Malmo, Sweden

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

ontologically inconsistent with the nature of the actual grammars that learners can, and do, develop through instruction and in post-instruction usage. In TESOL and theories of Second Language Acquisition (SLA), there is a widespread assumption that, with respect to grammar, the appropriate target for learners is performance which conforms to native speaker (NS) norms, i. e. underpinned by knowledge of what is normally referred to as "*the grammar* of English". In other words, the only relevant conceptualization of grammar for learning and teaching is at the community level, as an inevitably idealized group property (the "standard variety"). Yet, as Larsen-Freeman (2014: 209) remarks, "the fact is that the learner's system and any idealized system will never converge". Furthermore, as argued by several scholars (notably Cook 1991 and Firth and Wagner 1997), this assumption inevitably implies a deficit model of learning: learners are defined "in terms of what they are not" (Kramsch 1998: 28). Our strategy is to problematize such monolithic ontologies of grammar for TESOL by exploring the *plurilithic* outcomes (Hall 2013; Pennycook 2009) and ongoing development of lexico-grammatical resources in individual non-native speakers (NNSs) of English, post-instruction. As a first step in this process, we present a corpus-based characterization of a tiny fragment of the lexico-grammatical resources of a single individual, in order to illustrate an alternative usage-based ontology which could underpin more realistic approaches to grammar in the classroom.

The paper is organized as follows. We first discuss dominant approaches to grammar in SLA, and show how these are based on monolithic ontological assumptions about language in mainstream linguistics. We then briefly summarize previous corpus work on individual users of L1 and L2 English. In the body of the paper we describe our own informant and corpus, and identify the grammar fragment to be focused on, the *can you/could you* V construction alternation in requests. The main section presents an account of the way the constructions are configured in the individual's grammar, based on usage-based approaches to language but also recognising the formative roles of cross-linguistic influence from the L1, possible instructional biases, and pragmatic factors. We conclude by sketching the grammar fragment in terms of a constructional hierarchy and discussing the implications of our ontological approach for teacher development.

1.1 Ontologies of grammar in SLA

In SLA, NNS English grammar has been typically described with reference to "deviations" from NS norms or as features of Interlanguage states on a learning

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

trajectory towards an assumed "final state" of NS competence (cf. Han and Tarone 2014 for a recent re-appraisal). The issue of variable outcomes in grammar learning has been addressed in work on "ultimate attainment", focusing on the assessment of proficiency following instruction (e. g. Birdsong 2004; Lardiere 2006). But whether the nature of the target final state is simply asserted (as in generativist approaches) or is ascertained using corpora (as in functionalist approaches, e. g. Ellis et al. 2014), the explicit or implicit assumption has generally been that it is monolingual NSs of a "standard variety" who provide the yardstick. A logical consequence of this enduring monolithic ontology of target grammar is that *accuracy* is taken as an essential component in the conceptualization and assessment of language proficiency (Hall 2014). Accuracy, i. e. the extent to which an individual's production conforms to externally specified norms (normally NS "standard varieties"), is the main criterion for assessing grammatical attainment and, in consequence, for evaluating the relative effectiveness of different approaches to grammar instruction. Although some in SLA argue that the "A" in the "CAF" dimensions of L2 performance and proficiency (complexity, accuracy, and fluency) should also be taken to include appropriateness and acceptability (Housen et al. 2012: 4), success in learning English grammar is still generally understood in terms of the extent to which NS grammatical norms are accurately reproduced.

Cook's (1991, 2002) notion of multi-competence represents a powerful counter-argument to the dominant position, stressing that learners/users inevitably develop and use qualitatively different kinds of competence from that of monolingual users. But as Hall et al. (2006) point out, the ontology of language assumed in multi-competence work has not been fully specified, and its characterizations of monolingual competence are essentially monolithic, leading to a view of multilinguals as "possessing two or more separate, internally consistent language systems" (p. 225). Despite pockets of non-conformity (notably Harris 1981), the dominant ontology of language in linguistics, both theoretical and applied, has until recently been one in which English (and other named languages) exist as monolithic, internally consistent community systems.

Usage-based linguistics (e. g. Tomasello 2003) is an approach that provides a radically alternative conceptualization of grammar, as overlapping sets of structural regularities determined by individual mental analyses of the frequency and distribution of form-meaning pairings in the input received during usage events. Developing a grammar involves the assembly of "more or less schematic constructions on the basis of recurring lexically specific exemplars of the construction in question" (Eskildsen 2015: 56). This developmental process has been demonstrated for several constructions in L2 English (Bardovi-Harlig 2002; Mellow 2006; Eskildsen 2008; several papers in Collins and Ellis 2009;

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Gries and Wulff 2009; Wulff and Gries 2011; Deshors 2014; Deshors and Gries 2014; Eskildsen 2015). Eskildsen's (2008) study, for example, concentrates on the piecemeal, item-based development of constructions using modal *can* in an L1 Spanish speaker, which he traces back to specific usage events featuring the fixed expression *I can write*. He argues that the constructions develop as a function of the learner's situated experience, with "no evidence of a journey along the interlanguage continuum toward structural completion and increasingly native-like conformity" (p. 352).

1.2 Community grammars and NS norms

The grammar of "standard" varieties of English is an idealized social construct, abstracted away from the behaviour of many individuals and interpenetrated with non-linguistics factors. Many scholars argue that it is in fact an ideological construct, rather than a linguistic one (e. g. Firth and Wagner 1997; Gumperz and Cook-Gumperz 2005; Davis 2013). This is a particular problem for generative linguistics, which adopts Chomsky's (1986) concept of *I-language*, an individual's mentally-represented linguistic (grammatical) resources, but in practice idealizes away from I-language to describe grammatical rules and principles on the basis of (educated) NS intuitions about "well-formedness" (cf. Hall 2013). In Cognitive Linguistics, a grammar is an inventory of "conventional linguistic units" (Langacker 1987: 62), i. e. symbolic resources which are shared (to a greater or lesser degree) by a speech community. Well-formedness is understood as the degree to which an expression conforms to linguistic convention. These cognitively-oriented approaches in mainstream linguistics assume a homogeneity of outcomes in L1 acquisition which is being increasingly questioned. Experimental evidence from a variety of constructions, over several languages, demonstrates that grammatical knowledge varies between NSs, most probably because of differences in usage experience, principally in education and literacy (Dąbrowska 2012). This implies different degrees of *entrenchment* of constructions across individuals, i. e. variation in the extent to which the associated neural event type is stored and accessed as a unit in memory as "an established routine that can be carried out more or less automatically once it is initiated" (Langacker 1987: 100).

To what extent can learners of L2 English be expected to come to know the "conventional linguistic units" (CLUs) of the language, and what is the nature of the knowledge developed? Deliberate learning of rules can, of course, result in declarative knowledge of aspects of the community grammar, but it is broadly accepted that such knowledge will rarely, on its own, constitute an adequate

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

basis for effective use (cf. Bialystok 1981; VanPatten 2010). Neurobiological work on the distinct roles of procedural and declarative memory in L1 and L2 development (cf. Ullman 2014) suggests that instructed learning is less likely than "immersion-like" learning to lead to "native-like" processing. This again relates to the issue of CAF. Explicit learning in declarative memory will bolster accuracy in the CLUs used by NSs, but this is likely to be at the expense of *fluency*, a dimension of language use more associated with automatization (entrenchment) in procedural memory, and *complexity*, understood as increasing sophistication in grammar through restructuring of the implicit system. Research on the usage of English as a Lingua Franca (ELF) demonstrates that the implicit system, developing on the basis of uninstructed experience, is likely to diverge substantially from the NS norms modelled in instruction (Seidlhofer 2011, Mauranen 2012, Hall forthcoming).

ELF scholars have argued that a unique focus on NS norms as a model for learning is likely to be unhelpful for many circumstances of use (Seidlhofer 2011). Corpus work on ELF has shown that users often develop their own "natural" norms in order to communicate effectively in lingua franca contexts (Mauranen 2012:6). Such contexts, in which NS interactants are generally absent or at least in a minority, now probably account for a large proportion of global English usage events. ELF interactional data provide evidence of group norms emerging on the fly (e. g. Hynninen 2013), but the extent to which these norms may be retained by individuals across usage events, regulating the composition of their own lexico-grammatical resources, is not yet clear.

1.3 Corpus work on individual grammars

Our understanding of the nature of plurilithic Englishes is very poor, and the best prospect for detailed knowledge, namely through examination of corpus data, is as yet only very scantly fulfilled. Beyond authorship identification studies in literary criticism (e. g. Love 2002) and forensic linguistics (e. g. Coulthard 2004), there are very few corpus studies of individual NS usage. Mollin (2009) investigated the collocational preferences of Tony Blair on the basis of a corpus of transcripts of his public speaking and Barlow (2013) studied the idiolectal features of several White House press secretaries. A finding of the latter study that we will be returning to below is that while individuals tend to align their production with that of their interlocutors, they still maintain robust, idiosyncratic structural preferences which can resist patterns in the input.

There are almost no studies which focus on the lexico-grammatical resources of individual NNS English users post-instruction. The only work that

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

we are aware of (Vetchinnikova 2015 and Vetchinnikova forthcoming) uses ELF data from NNS academic writers and a prolific blog poster to show that: (a) individual users tend to re-use multi-word units (MWUs) that have appeared in the input to which they have been exposed; and (b) individual users' lexical patterning is more "fixed" (less variable) than at the community level. The first finding contradicts claims that in ELF contexts users show a preference for the "open choice principle" over the "idiom principle" (Sinclair 1991). The second finding, on the other hand, confirms Barlow's (2013) observation that input is not the sole determinant of an individual's lexico-grammatical resources. On the basis of interactional analysis in ELF contexts, Hynninen (2013) also found that NNSs maintained their "idiolectal preferences" across usage events.

1.4 An ontological framework

In the following analysis, we adopt the ontological framework for language introduced in Hall (2013) and applied in Hall et al. (2015). The relevant ontological categories are as follows.

- *I-language*: the mentally represented linguistic resources of an individual (*I* suggesting "internal", "individual");
- *I-register*: the I-language resources associated with use in a specific social/functional domain or genre;
- *E-language*: expressed, situated utterances or texts (*E* suggesting "external");
- *N-language*: named linguistic resources like "(Standard) English" (*N* suggesting also "normed," and "national");
- *P-language*: idealized linguistic systems independent of cognition and social use (the *P* stands for *Platonic*, as in the concept of Platonic ideal).

The categories I-, E-, and P-language were all originally introduced by Chomsky (1986). As alluded to previously and argued in greater detail in Hall (2013), the object of study in generativist linguistics is not actually I-language, but rather an idealized form of it filtered by "grammaticality judgements" which are inevitably biased by N-language conceptualizations. Hall (2013) interprets I-languages as repertoires of linguistic resources, often drawn from what are socially constructed as different N-languages, and deployed in different domains/genres of use. In bilinguals, such I-registers will be "polylingually constituted" (Jørgensen 2008) and involved in translanguaging (Otheguy et al. 2015).

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

1.5 Research questions and research strategy

Our object of interest is NNS I-language, studied not according to the implicitly monolithic N- or P-language principles of traditional linguistic analysis, but instead using a more psychologically plausible epistemology based on inferences from E-language, in the form of corpus data. Specifically, we wanted to address the following questions:

- To what extent are individual NNS lexico-grammatical resources (I-language) determined by input (E-language)?
- To what extent do they align with (E-language) genre norms?
- What other factors of linguistic awareness and experience might explain the individual's I-language configuration?
- How can I-language grammars best be characterized?

Our strategy was, accordingly, to collect longitudinal usage data from an individual NNS user of English and to identify a representative lexico-grammatical feature to examine within a particular I-register. Computer-mediated communication (CMC) has obvious advantages as a source of data for such a project, principally: (a) it already exists in device histories such as email inboxes; (b) it includes both input and output; (c) it is already transcribed; and (d) much of it is genre-specific. We therefore sought to identify informants with the appropriate linguistic profile who would be prepared to share with us a large sample of their recent CMC interactions with other NNSs and NSs.

2 Method

We contacted potential NNS informants who had recently graduated from a taught postgraduate degree for experienced international managers, delivered mostly online. Only a few candidates had CMC data for longer than six months, and only one was prepared to share a sizeable amount of data with us and commit the significant amount of time required to provide data on correspondents. Given our focus on the nature of I-language, however, only one participant was sufficient for our purposes.

2.1 The informant and his correspondents

Our informant Antonio (a pseudonym) gave us access to the emails in his inbox for the previous 18 months and his outbox for the previous 36 months. At the

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

time of data collection, he was working at managerial level in an industrial corporation in South Africa, where he had lived for over 25 years. His first language is Italian. He studied English at school in Italy and spent two years in the UK, returning to Italy with his English NS wife for eight years. In South Africa he learned sufficient Afrikaans for many purposes, but he reported English as his dominant language. He estimated that he used English around 60 % of the time, followed by Afrikaans at 35 %, and Italian at only 5 %. On 10-point proficiency self-assessment scales from the *LEAP-Q* questionnaire (Marian et al. 2007), Antonio rated himself with a mean score of: 8.75 for English; 3.5 for Afrikaans; and 8.25 for Italian. He reported having a "moderate foreign accent" in English and that he was taken as a NNS 50 % of the time. There are certainly NNS features in the corpus data, such as:

- Subject-Auxiliary Inversion in embedded WH-clauses (e. g. *Do you know where can we find the drawings?*)
- NNS prepositional complements (e. g. *interested on straightening*; *result on the suspension*)

But the samples of Antonio's writing that we have seen use sophisticated vocabulary and syntax and appear to be communicatively effective. The fact that he worked for an Anglophone company at a very high level of responsibility confirms that Antonio is an expert user.

Antonio also provided information on a subset of his correspondents, totalling 292 individuals. In terms of L1, Afrikaans (43 %) and English (41 %) dominate. Next most frequent are NSs of Sesotho and Zulu (13 % combined), followed by NSs of several European languages (3 %). Many of his email exchanges were thus performed in an ELF context. Antonio coded the majority of his correspondents (61 %) as work colleagues, followed by service providers (22 %), friends and relatives (11 %), and work clients/suppliers/partners (7 %). These data indicate that most of the messages (89 %) were sent to and/or received from correspondents in the business domain or in personal business/service transactions. Lastly, Antonio reported knowing most of his correspondents (59 %) for less than two years, but the largest single category (37 %) was between two and five years.

2.2 The corpora

The Outbox (OB) corpus of sent mail comprises 57,291 word tokens, representing 5,661 types, in 2,179 messages sent over a period of 18 months. The Inbox (IB) corpus of mail received is larger, comprising 74,145 word tokens, representing

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

6,525 types, but drawn from a smaller number of messages (1,600) received over a period of 36 months (coterminous with the OB data). The messages were stripped of all headers, footers, copies of other messages, junk mail, and extensive quoted text. The OB constitutes the main corpus, with the IB as reference corpus. A second reference corpus was the EnronSent (EN) corpus (Styler 2011). This is a very large subset of a substantial collection of emails from a multinational corporation, in the public domain since 2003. It comprises almost 14 million words from over 96,000 messages. The corpus is untagged and there is no metadata on authors or timeframe. We took the OB data to be a more-or-less representative E-language reflex of Antonio's I-language in this domain (an I-register). Similarly, we took the IB data to represent the kind of E-language to which Antonio was exposed in this domain prior to and during the period sampled, and used the EN data to identify broader international norms for the genre. Table 1 provides a summary of the corpora.

Table 1: Composition and coverage of the main and reference corpora.

	Messages	Word tokens	Word types	Time period
Outbox (OB)	2,179	57,291	5,661	18 months
Inbox (IB)	1,600	74,145	6,525	36 months
EnronSent (EN)	96,107	13,810,266	Unknown	Unknown

3 Data analysis

The corpus data were analysed using the AntConc corpus analysis software (Anthony 2011). We adopted a data-driven approach, allowing the software to yield focus constructions for analysis. The first step was to conduct keyword analyses of the OB data in order to get an impression of the most frequent content addressed in the emails and the language used to do so (cf. Scott and Tribble 2006: ch. 4). We used the IB as a reference corpus to see how Antonio's language differed from the kind of language he was exposed to in this genre. We then repeated the analysis using the EN as a reference corpus, in order to see how Antonio's language differed from broader genre norms. Keyness was calculated using the log-likelihood (LL) statistic as well as the mutual information (MI) measure. The patterns emerging were similar on both measures, so we report only LL statistics here. The IB comparison revealed that Antonio's language (the OB) differed from that of his correspondents (the IB) mostly with

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

regard to stylistic features, especially politeness markers (e. g. *thank*; component parts of the phrase *I trust this finds you well*), formal language (e. g. *herewith*) and the making of requests (e. g. *wondering*). The EN comparison repeated this pattern, but also reflected the lack of shared context through keywords which reference Antonio's specific domain (e. g. product and company names).

The modal verb *can* features as a highly significant keyword in both comparisons, with a log likelihood of 89.79 with reference to IB and 128.05 with reference to EN (for both, $p < 0.0001$). Several corpus studies have observed that learner usage of *can* is elevated compared with NSs, and may be selected as a default over other modals associated with similar functions in complex linguistic contexts (Deshors and Gries 2014). In the OB, *can* accounts for 1% of the entire corpus, compared with 0.5% in IB and 0.3% in EN. A visual inspection of concordance lines for *can* in the OB suggested that it was being used mostly for requests. Since *can* alternates with *could* in indirect requests in English, and these are a prominent feature of this domain and genre, we decided to focus on this construction alternation. We therefore subjected Antonio's use of the two modals to further scrutiny, examining their overall distribution and comparative frequency. *Could* is used relatively more frequently in the OB than in the reference corpora (0.18% vs. 0.13% for both the IB and EN), consistent with the finding of Neff et al. (2003) that Italian NS learners also used this modal more frequently than did NSs of English. Yet it did not feature prominently in our keyword analysis, with a keyness score of 1.11 in the IB comparison (not statistically significant) and 4.5 in the EN comparison (only moderately significant).

Further inspection of *can* concordance lines in the OB and IB reveals that Antonio uses the modal disproportionately in interrogative structures. In the IB, *could* is used roughly equally in declarative and interrogative structures, whereas in the OB it is twice as likely to be declarative. This suggests that for Antonio, interrogative *can* might be doing a lot of the work done by interrogative *could* in the IB.[1] Within interrogative clauses, the bulk of *can* and *could* tokens occur in yes/no questions (YNQs), in both the OB and IB, and virtually all of these are requests. Of course, *can/could* constructions have a variety of (dynamic, deontic, and epistemic) functions in YNQs, and analysis of Antonio's usage suggests that he uses them for these functions too. E.g.:

1. Can I go ahead with the repair? [deontic: permission]
2. Can you beat this quotation? [dynamic: ability]

1 We do not have corresponding data from EN, given the obstacles to carrying out a full analysis of modal verbs in such a large database without PoS tags.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

But the request function predominates in both the OB (79%) and IB (82%), and of these, the majority are in second person.

Antonio's more frequent use of *Can you* V (henceforth CnYV) than *Could you* V (henceforth CdYV) in requests is reflected neither in the input to which he is exposed (IB) nor the norms for the genre (EN), as illustrated in Figure 1. A Fisher's exact test for independence[2] reveals that the proportions in the OB differ significantly from those for the IB ($p < 0.0001$). The following section considers the factors that might account for this strong idiolectal preference.

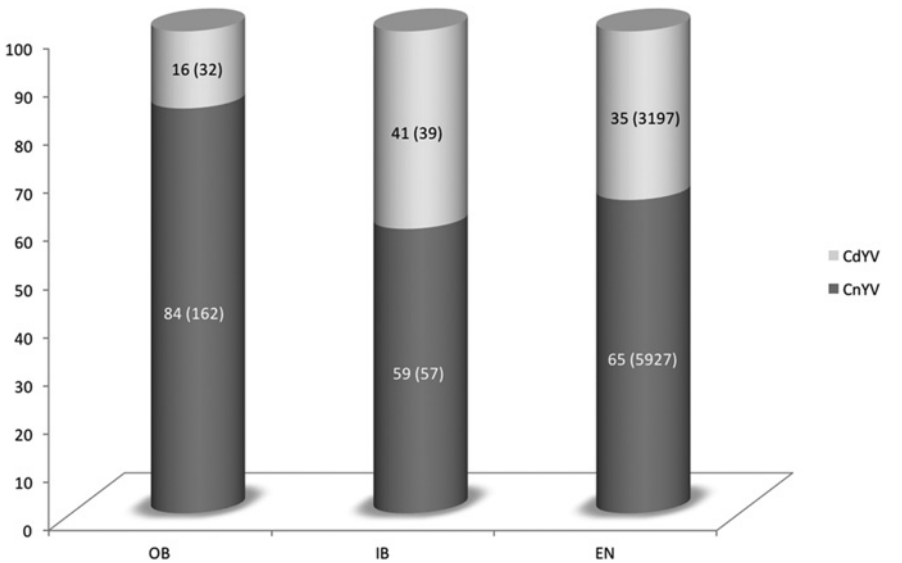

Figure 1: Construction token percentages (with raw frequencies in parentheses) across the three corpora.

4 Accounting for the CnYV preference

A number of factors no doubt contribute to the greater entrenchment of the CnYV construction over the CdYV alternative in Antonio's I-register for English email requests. We consider the most likely here.

2 Used in preference to other correlation tests because of the small numbers involved (cf. Gries and Stefanowitsch 2004).

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

4.1 Initial learning context

Antonio's development of the constructions will have been influenced by the instructional/usage contexts in which he encountered them during his early experiences with English. Regrettably, we do not have access to relevant data, so cannot conduct the kind of analysis that Eskildsen (2008) performed to track his informant's development of modal *can* during instruction. Nevertheless, there is some research on other Italian NSs' learning of the English modals *can* and *could* which is germane to our study. Research using cross-sectional learner corpora has reported that *can* is used significantly more frequently by Italian NS learners than by NSs of English, in both academic writing (Neff et al. 2003) and CMC settings (Dalziel and Helm 2008). Neff et al. suggest that their elevated numbers for *can* are due in part to its use in epistemic senses (e. g. "I think that this [...] can be the only way to improve our society") where NSs would use *could* or *might* (p. 216). They attribute this to cross-linguistic influence (CLI) from the Italian modal *potere*, which has broader functional scope than English *can*. We might expect Antonio to have initially forged a strong interlingual association (cf. Hall et al. 2009) between his mental representation of the entrenched present indicative forms of *potere* and their perceived L2 translation equivalent *can*, leading to its use in constructions where NSs might use *could*, *might*, *may* or other modals. As Neff et al. (2003: 216) suggest, this might also be expected given that *can* is generally the first modal taught to learners. It may also be over-represented in teaching materials, compared with NS usage (Römer 2004).

Most accounts of lexical CLI demonstrate a weakening of associations between L1 and L2 items as exposure to the L2 increases (cf. Hall and Reyes 2009). L2 items can become functionally dissociated from their original L1 hosts as a result of learners' participation in meaningful usage events in which the L2 forms are encountered in context. Some frequently experienced items will become stored as part of semi-fixed expressions (SFEs) as well as on their own. So, although an initial association between *can* and indicative *potere*, together with instructional experience, might make this modal more available for production than *could*, both modals will be experienced in the input and will gradually become stored also with associated lexical material (e. g. specific verbs) for functions which are divorced from the original meanings associated with the modals on their own. In this way, CnYV and CdYV will get stored as SFEs and directly activated for requests. But why does Antonio come to use the former construction so much more than the latter in post-instruction usage?

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

4.2 The pragmatics of requests

According to conventional pragmatic analysis (e. g. Leech 2014) and limited NS and NNS judgement data (Carrell and Konneker 1981), *could* is considered more polite than *can*. An informal inspection of EFL textbooks and Internet resources confirms that this is also the message that learners are likely to receive. There is a parallel alternation to CdYV/CnYV for requests in Italian, with *potere* in the indicative taken as less formal than the conditional, matching the way the isolated modals *can* and *could* are presented in instruction. This suggests that influence from the L1 should be positive (a case of "transfer to somewhere" in Andersen's [1983] words), leading to straightforward initial mapping to CdYV and CnYV.

But closer analyses of usage data have revealed that *can/could you* V are now pragmatically unmarked alternates for requests, at least for NSs (Watts 2003; Curl and Drew 2008). Nevertheless, responding to follow-up questions after data collection, Antonio expressed the belief that CdYV was more formal/polite, and that he used it more with correspondents he did not know so well. We therefore inspected the (albeit patchy and coarse-grained) data on his relationship with correspondents to check for any correlation.[3] Over 75 % of *both* constructions were used in requests to work colleagues and service providers, who together account for over 80 % of all correspondents. The other two groups yielded numbers which were too small for meaningful interpretation; however, it is perhaps significant that CdYV was used five times more often than CnYV with friends, suggesting that the reported association of this construction with greater formality/politeness is not a factor in Antonio's usage. This is consistent with the findings of Watts (2003) and Curl and Drew (2008) and inconsistent with the speaker's own intuitions.[4]

Additionally, a correlation with length of relationship revealed that, contrary to his own intuitions, Antonio uses CnYV more with newer acquaintances than older ones, and vice-versa for CdYV. The result was significant at $p < 0.001$ according to a Fisher exact test. Also, presence of *please* was more frequent than absence in both constructions, indicating politeness across requests. Peterson *et al*. (2011) found that in the EN corpus, more formal language was used overall in requests vs. non-requests, so perhaps genre-specific factors eliminate any formality distinction associated with the construction alternation outside the

3 Unfortunately we do not have data on his relative seniority with respect to work colleagues.
4 We have no ready account for the striking discrepancy between Antonio's intuitions and his actual usage. Future studies could track user intuition data more systematically in order to shed light on the issue of post-instruction language awareness in successful L2 users.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

genre. In sum, we find no evidence that Antonio's request pragmatics can account for his preference for CnYV over CdYV.

4.3 Verb use

We turn now to an analysis of the lexical verbs used in the constructions, to explore the extent to which the OB patterns may be a function of the collocational strength of particularly frequent verbs and the overall verb frequency patterns in the input to which Antonio has been exposed and in his own output. An examination of collocational association, using the t-score measure of confidence, shows that the six most frequent verbs in the CnYV construction are particularly significantly associated with the *can you* bigram (t-score > 2).[5] For verbs in CdYV, on the other hand, only *forward* is significantly associated with *could you*. Table 2 lists the association strengths, with their t-score, by order of frequency (absolute and normalized per 100 instances to allow comparison across constructions and corpora). Of the six most frequent verb types in the OB, only two are strongly associated with the modal in the IB: *assist* in CnYV (t = 2.7) and *advise* in CdYV (t = 2.4). See Table 3.

Table 2: Frequency (normalized per 100 in brackets) and strength of collocational association (t-value) of the six most frequent verb types in each construction in the OB.

CnYV type	Frequency	t	CdYV type	Frequency	t
assist	23 [40.35]	4.9	*forward*	05 [29.41]	2.4
provide	16 [28.07]	3.6	*advise*	04 [23.53]	–
(re)send	14 [24.56]	3.6	*tell*	04 [23.53]	–
forward	09 [15.79]	2.7	*assist*	02 [11.76]	–
advise	08 [14.04]	3.1	*provide*	02 [11.76]	–
(double)check	06 [10.53]	2.6	*(re)send*	02 [11.76]	–

The verb *assist* is the main exemplar for CnYV for both the OB and the IB. There are fewer types and tokens overall for CdYV, but there is still partial alignment at the high frequency end. This time *advise* is the most frequent verb in the IB, and it is also the second most frequent verb for Antonio in this frame. Overall, these data reveal partial coincidence of frequent verb types across the IB and the OB, but with some significant differences.

5 Given the low token verb frequencies in CdYV, we calculated t-values within constructions rather than using distinctive collexeme analysis to compare preferences for verbs across constructions (Gries and Stefanowitsch 2004).

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Table 3: Most frequent verb types across the constructions in the IB corpus, compared with OB (normalized per 100 in brackets).

CnYV type	Frequency IB	Frequency OB	CdYV type	Frequency IB	Frequency OB
assist	07 [20.00]	23 [40.35]	*advise*	05 [18.52]	04 [23.53]
forward	04 [11.43]	09 [15.79]	*confirm*	03 [11.11]	–
(double) check	03 [08.57]	06 [10.53]	*(re)send*	03 [11.11]	02 [11.76]
give	03 [08.57]	03 [05.26]	*assist*	02 [07.41]	02 [11.76]
make	03 [08.57]	01 [01.75]	*indicate*	02 [07.41]	01 [03.70]
(re)send	03 [08.57]	14 [24.56]	*let know*	02 [07.41]	–
review	03 [08.57]	01 [01.75]	*provide*	02 [07.41]	02 [11.76]
do	02 [05.71]	01 [01.75]			

Some of the same verbs also crop up with proportionally high frequency in the EN corpus, suggesting that they are typical of the genre overall, as Table 4 shows.[6] But the most frequent couple of verbs across the constructions in both the OB and IB are relatively infrequent in EN: *assist* occurs in CnYV 12 times (normalized 0.22) and in CdYV only twice (normalized 0.07); and *advise* occurs

Table 4: Most frequent verb types across the constructions in the EN corpus, compared with combined OB/IB frequency (normalized per 100 in brackets).

CnYV type	Freq EN	Freq OB/IB	CdYV type	Freq EN	Freq OB/IB
(re)send	337 [6.28]	17 [18.48]	*(re)send*	167 [5.76]	05 [11.36]
help	311 [5.80]	04 [04.35]	*let*	134 [4.62]	02 [04.55]
give	257 [4.79]	06 [06.52]	*give*	081 [2.79]	01 [02.27]
get	247 [4.61]	04 [04.35]	*forward*	079 [2.72]	05 [11.36]
tell	213 [3.97]	03 [03.26]	*email*	062 [2.14]	–
(double)check	191 [3.56]	09 [09.78]	*(double)check*	057 [1.96]	01 [02.27]
let	161 [3.00]	01 [01.09]	*have*	050 [1.72]	–
provide	139 [2.59]	17 [18.48]	*get*	044 [1.52]	01 [02.27]
make	125 [2.33]	03 [03.26]	*tell*	040 [1.38]	04 [09.09]
confirm	112 [2.09]	06 [06.52]	*provide*	039 [1.34]	04 [09.09]
do	089 [1.66]	03 [03.26]	*help*	033 [1.14]	–
take	089 [1.66]	03 [03.26]	*take*	030 [1.03]	01 [02.27]
call	085 [1.59]	–	*confirm*	027 [0.93]	03 [06.82]
have	080 [1.49]	04 [04.35]	*do*	025 [0.86]	01 [02.27]
email	76 [1.42]	–	*call*	023 [0.79]	–
forward	74 [1.38]	13 [14.13]	*look into*	021 [0.72]	–

6 We used cut-offs of ≥ 74 for CnYV and ≥ 33 for CdYV in EN as proportional to the minimum OB frequencies used for inclusion in Table 2.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

31 times (normalized 0.58) in CnYV and 11 times (normalized 0.38) in CdYV. This suggests that in Antonio's context, there is a pattern of usage of verbs in this construction alternation which is not present in the broader genre: a set of "natural" local norms (Mauranen 2012: 6).

Overall, these analyses show: (a) considerable coincidence of frequent verbs in the constructions across corpora, but with some notable differences between OB and IB, and OB/IB and EN; and (b) an especially tight association between a handful of frequent verbs and *Can you* in Antonio's production.

We now take into account the distribution of less frequent verb types used in the construction. Figure 2 plots the token frequencies of OB verbs in CnYV from highest to lowest. It will be observed that the type/token distribution at least partially approximates a Zipfian curve (Ellis 2013: 369–370), in which the highest frequency items account for a disproportionately high number of tokens: in CnYV, Antonio uses a small number of verbs a lot, and a large number of verbs only once. It has been widely documented for infants that the acquisition of constructions is typically driven by the initial dominance in production of a single, prototypical item or small set of items, typically verbs (cf. Tomasello 2003: 117–122). Evidence is emerging for similar learning trajectories in L2 (e. g. Ellis and Ferreira-Junior 2009) and is essentially what Eskildsen (2008) observed in his study of an L2 English learner's development of *can*. From a usage-based perspective, learning and usage are two sides of the same coin, so patterns of comprehension and production will continually influence the configuration of a speaker's I-language, in both L1 and L2. We may therefore expect to see item-based learning patterns

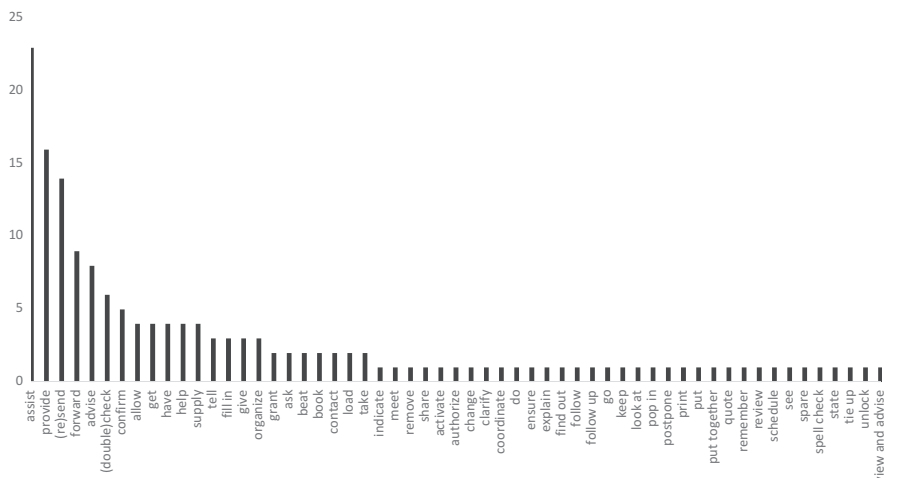

Figure 2: Verb token frequencies for CnYV in the OB.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

recapitulated in usage patterns, such as the ones under analysis here. In line with this expectation, Vetchinnikova (forthcoming) found similar Zipfian effects with adjectives used by a NNS blogger in the *it is* ADJ *that* construction.

Psycholinguistic work on accommodation and structural priming (cf. Pickering and Ferreira 2008) demonstrates that speakers tend to reproduce grammatical constructions that have been receptively processed in preceding discourse, especially when they are associated with specific lexical material, typically verbs (Gries 2005). This work addresses the psychological mechanism through which input from usage events can continue to shape the dynamic language system. But usage, of course, includes a speaker's output as well as input, and this would appear to be a major factor in Antonio's elevated use of CnYV compared with the lower levels of usage in the IB and EN. Barlow (2013) documents similar individual divergence from community norms in his study of idiolectal variation in NS White House press secretaries. He speculates that despite the tendency to align one's production to patterns in the input, this can be offset by "an inbuilt inertia and preference for familiar routines leading to a consistency in frequency of usage of language expressions by individual speakers over a period of time" (pp. 471–472) with the result that "the patterns of production data do not follow the patterns of comprehension data as closely as might have been thought" (p. 472). This is also consistent with Vetchinnikova's (forthcoming) finding that individual lexical patterning in ELF appears to be more "fixed" than community norms.

Accordingly, we suggest that the main determinant of Antonio's elevated usage of CnYV is his idiolectal preference for the exemplar *assist*, by far the most frequently occurring verb in CnYV for both OB and IB (see Tables 2 and 3). This preference might also be due, in part, to CLI, since the verb is cognate with Italian *assistere*, whereas its near-synonym *help* is generally translated by the non-cognate form *aiutare*. *Assist* is also a prototypical one for the construction, distilling the request function from the object of the request and shifting the latter into another construction, typically a gerund complement of a prepositional phrase headed by *with*. A selection of concordance lines from the OB in Figure 3 illustrate the pattern.

y, can you check it out with [name]? [Name] can you	assist	with filling in the dates for the coastal trucks
tes maximum on a personal matter. Can you please	assist	with scheduling a suitable time." I need help on
th weekly report for your perusal. Can you please	assist	with affixing another white board in my office? (
s), regarding the change of name? Can you please	assist	with completing the details as much as we can? Th
I FYI Sorry mate I left you out . Can you please	assist	with finalizing this spread sheet, with providing

Figure 3: Concordance lines illustrating the *Can you assist with* V-*ing* SFE in the OB.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Antonio's preference for the more indirect *Can you assist with* V-*ing* over direct *Can you V* may also stem in part from the enhanced overall politeness/formality he adopts in this genre, as reflected in the markers of formality that emerged in the keyword analysis and his frequent use of *please* in both constructions. This would also account for his preference for *assist*, designated "a formal word" by the corpus-based *Longman Dictionary of Contemporary English*, over *help*, "the most general verb" of the synonym group. But the co-occurrence of the SFE with *Sorry mate* in the last concordance line in Figure 3 suggests that the construction has subsequently become dissociated from active pragmatic choices. In sum: Antonio's original preference for more formal language in email requests (attested across the genre, as evidenced by Peterson et al.'s [2011] work on the EN data), together with the frequency of *Can you assist* in the email requests he received, and possible CLI from *assistere*, has led to a proliferation of *assist* in his own use of CnYV, which in turn sets up a bias for new verbs to be used in the same construction.

5 Discussion and conclusions

Examination of one NNS's language experience reveals that he uses CnYV for email requests much more frequently than would be expected on the basis of patterns in the broader genre and the input he receives. We explored the possibility that the preference was due to Antonio's initial learning experience and/or differences in his construal of the constructions' pragmatics, but concluded that neither alone was able to account for the high CnYV numbers. Detailed examination of the distribution of token frequencies of verb types occurring in the constructions' frames reveal that a single verb, *assist*, although relatively uncommon in the overall genre, is particularly common in the input Antonio receives, and is disproportionately used in his output, probably due to CLI and the inertia in production suggested by Barlow (2013). This verb is a prototypical one for the request function, resulting in more indirect, and therefore more polite requests, and its use in preference to less formal *help* is consistent with the general levels of formality observed in his usage. We concluded that Antonio's preferential use of CnYV is, in large part, a function of his input- and CLI-induced preference for the verb *assist*. This has led to the entrenchment of the MWU *Can you assist* in his email I-register, which in its turn has the effect of biasing construction selection for new verbs.

We are now in a position to sketch out at least one tiny fragment of the lexico-grammatical resources (the I-language) of this particular non-native user

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

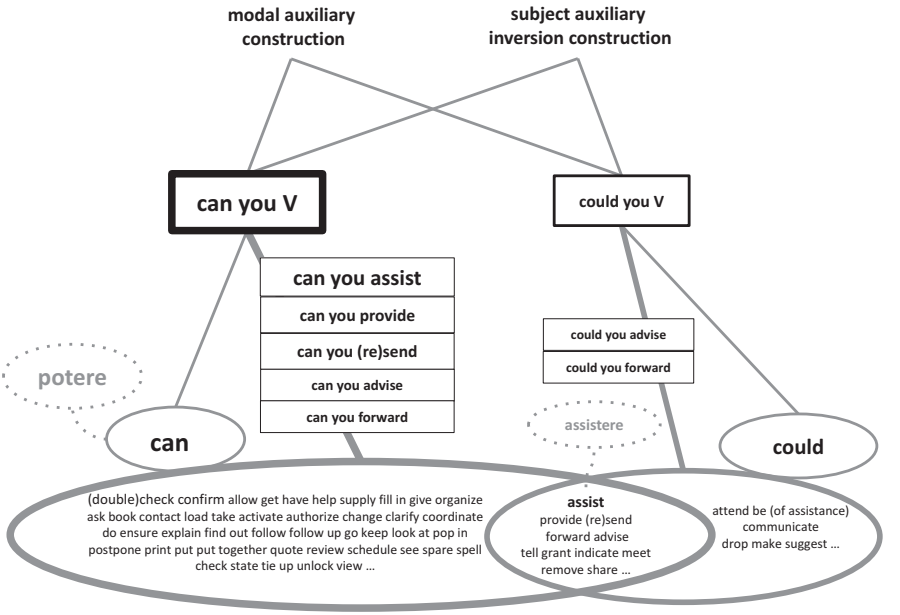

Figure 4: A fragment of Antonio's lexico-grammar, showing the constructions involved in email requests using the *can/could* V alternation.
Key: Ellipses contain individual word forms; rectangles contain MWUs and SFEs. Lines indicate inheritance and activation. Size of font and thickness of borders reflect relative frequency levels. L2 word forms are indicated in dotted ellipses.

of English. Figure 4 provides such a sketch, in the form of an inheritance network or structured taxonomy inspired by those used in Construction Grammar (Goldberg 2003). At the base of the hierarchy are lexical items, which are linked as subparts to the more complex constructions they participate in on the next level up (MWUs and SFEs). In turn, the MWUs and SFEs are linked as instantiations to more abstract constructions at the top of the hierarchy. We have used size of font and thickness of borders to indicate (approximately) the relative frequency levels of different constructions. The diagram is inevitably a vast oversimplification. For example, it is evidently part of a polylingually-constituted set of resources, but we have only included the Italian verbs *potere* and *assistere* from Antonio's other language resources, given their potential significance in the development of the network. Furthermore, we have omitted semantic/pragmatic information, whereas a full representation would need to specify form-meaning pairings for each construction.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

The representation of lexico-grammatical resources sketched in Figure 4 assumes a very different ontology of grammar from those underpinning most approaches to SLA and TESOL. In the P-language conceptualizations of many textbooks and teachers, only the constructions at the top of the hierarchy are represented, in the guise of "grammar rules". Corpus-based approaches recognize lower levels, referring, for example, to the use of CdYV for formal requests (e. g. in the relevant entry in the *Longman Dictionary of Contemporary English*). But corpus-based grammars are idealized representations of N-language based on NS E-language, not successful NNS I-language. The results of our analysis suggests that the I-language lexico-grammatical resources that successful NNS users come to know, use, and adapt post-instruction can best be characterized as construction inventories shaped by individuals' initial learning context, identity, the usage events in which they participate, and their own inertia in production. We believe that if teachers can develop ontologies of grammar which embrace the actual outcomes of the ongoing learning process, they will be better able to critically evaluate the untenable assumption that a monolithic NS-based target is a realistic pedagogical precept and that learning trajectories can be expected to converge on so-called NS norms.

Proponents of usage-based linguistics have stressed the importance of engaging learners in input-driven, participatory activities (e. g. Robinson and Ellis 2008: 495), yet the application of the approach to TESOL is problematic. Some have advocated the direct use of construction grammar in the classroom (Holme 2010). This requires further research, however: for example, the skewing of verb frequencies in learner input may not facilitate learning of target constructions when interactional opportunities are limited (Year and Gordon 2009). Also, despite the increasing prominence of ELF research, many usage-based linguists still assume learning targets based exclusively on NS community norms, referring uncritically to "correct usage" (Holme 2010: 373) or "a type of English which is really used and understood by native speakers" (Römer 2004: 186). Similarly, for Ellis (2002) the role of teachers is to ensure "that the learner representations that emerge [from experience] approximate those of native speakers" (p. 179). Wulff and Gries' (2011: 70) definition of accuracy, as the selection of constructions which match the contextual preferences of "a particular target variety and genre", and their view of proficiency as "a gradual phenomenon that transcends a NS-NNS divide", reflect a less monolithic stance; but it is implicit in their case studies and discussion that NS patterns are the expected and desired learning target.

We believe that pedagogical approaches which anticipate plurilithic learning outcomes such as the one we have explored here will be superior to those which are based on the expectation of convergence with NS norms. But we are a

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

long way from understanding how to incorporate usage-based insights into pedagogical practice. A logical prior step, the one which motivated the present study, is to problematize traditional monolithic ontologies of grammar for (applied) linguists and teachers alike (cf. Dewey 2012). We believe that this can be achieved, in part, by helping teachers apprehend alternative conceptualizations of grammar, which take seriously the lexico-grammatical resources developed, deployed and constantly updated by successful NNS *users*. One example of a resource which attempts to do this is the free online course *Changing Englishes* (Hall and Wicaksono 2013). For this kind of sea change in thinking to occur, approaches to grammar teaching and accuracy in mainstream teacher development programmes will have to be fundamentally reconceived.

This exploratory study constitutes a first step in theorizing and evidencing the kind of usage-based I-language view of grammar which we believe can underpin such a reconceptualization. To advance the broader project and elaborate on the findings reported here, more and more detailed characterizations of idiolectal L2 language resources are needed. Such studies will provide data on: different constructions and construction alternations; the same constructions in different I-registers; and constructions in I-languages developed under different learning conditions and deployed in different usage contexts. Overall, future work will need to employ more sophisticated methodologies capable of accurately distinguishing and assessing the contributions of different factors, in addition to input, which shape individual users' lexico-grammatical resources. A particular focus on the independent role of output (Barlow's [2013] notion of inertia) in constructional entrenchment is required. In sum: there is a great deal of work still to be done, both empirical and applied, before the suggested ontological shift has any chance of coming to pass.

Acknowledgements: We are greatly indebted to Antonio for making available to us his email data and for the time and care he took to code correspondent information and answer our questions about his usage and learning experiences. We are grateful also to Svetlana Vetchinnikova, Vander Viana, and two anonymous reviewers for their very helpful observations and suggestions. Their feedback has substantially enriched our account, but of course any shortcomings in the analysis and interpretation of the data are entirely our own.

References

Andersen, Roger. 1983. Transfer to somewhere. In Susan Gass & Larry Selinker (eds.), *Language transfer in language learning*, 177–201. Rowley, MA: Newbury House.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Anthony, Laurence. 2011. *AntConc* (Version 3.2.4) [Computer Software]. Tokyo, Japan: Waseda University. http://www.antlab.sci.waseda.ac.jp/.

Bardovi-Harlig, Kathleen. 2002. A new starting point? Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition* 24. 189–198.

Barlow, Michael. 2013. Individual differences and usage-based grammar. *International Journal of Corpus Linguistics* 18(4). 443–478.

Bialystok, Ellen. 1981. The role of linguistic knowledge in second language use. *Studies in Second Language Acquisition* 4(1). 31–45.

Birdsong, David. 2004. Second language acquisition and ultimate attainment. In Alan Davies & Catherine Elder (eds.), *Handbook of applied linguistics*, 82–105. Oxford: Blackwell.

Carrell, Patricia L. & Beverly Konneker. 1981. Politeness: Comparing native and nonnative judgments. *Language Learning* 31(1). 17–30.

Chomsky, Noam. 1986. *Knowledge of language. Its nature, origin, and use*. New York: Praeger.

Collins, Laura & Nick Ellis (eds.). 2009. Input and second language construction learning: Frequency, form, and function. [Special issue]. *The Modern Language Journal* 93(3). 329–470.

Cook, Vivian J. 1991. The poverty-of-the-stimulus argument and multi-competence. *Second Language Research* 7(2). 103–117.

Cook, Vivian J. (ed.). 2002. *Portraits of the L2 user*. Clevedon, UK: Multilingual Matters.

Coulthard, Malcolm. 2004. Author identification, idiolect, and linguistic uniqueness. *Applied Linguistics* 45(4). 431–447.

Curl, Traci S. & Paul Drew. 2008. Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction* 41(2). 129–153.

Dąbrowska, Ewa. 2012. Different speakers, different grammars. Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2(3). 219–253.

Dalziel, Fiona & Francesca Helm. 2008. Exploring modality in a learner corpus of online writing. In Carol Taylor Torsello, Katherine Ackerley & Erik Castello (eds.), *Corpora for university language teachers*, 287–302. Bern: Peter Lang.

Davis, Hayley G. 2013. The language myth and standard English. In Roy Harris (ed.), *The language myth in western culture*, 41–54. London: Routledge.

Dewey, Martin. 2012. Towards a *post-normative* approach: Learning the pedagogy of ELF. *Journal of English as a Lingua Franca*, 1(1). 141–170.

Deshors, Sandra C. 2014. A case for a unified treatment of EFL and ESL. A multifactorial approach. *English World-Wide*, 35(3). 277–305.

Deshors, Sandra C. & Stefan T. Gries. 2014. A case for the multifactorial assessment of learner language. The uses of *may* and *can* in French-English interlanguage. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy,* 179–204. Amsterdam: John Benjamins.

Ellis, Nick. 2002. Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24. 143–188.

Ellis, Nick. 2013. Construction grammar and second language acquisition. In Thomas Hoffmann & Graham Trousdale (eds.), *The Oxford handbook of construction grammar*, 365–378. Oxford: Oxford University Press.

Ellis, Nick C. & Fernando Ferreira-Junior. 2009. Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal* 93(3). 370–385.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Ellis, Nick C., Matthew B. O'Donnell & Ute Römer. 2014. Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistics Approaches to Bilingualism* 4(4). 405–431.

Eskildsen, Søren. 2008. Constructing another language—Usage based linguistics in second language acquisition. *Applied Linguistics* 30(3). 335–357.

Eskildsen, Søren. 2015. What counts as a developmental sequence? Exemplar-based L2 learning of English questions. *Language Learning* 65(1). 33–62.

Firth, Alan & Johannes Wagner. 1997. On discourse, communication, and (some) fundamental concepts in second language acquisition research. *Modern Language Journal* 81(3). 285–300.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.

Gumperz, John J. & Jenny Cook-Gumperz. 2005. Language standardization and the complexities of communicative practice. In Susan McKinnon & Sydel Silverman (eds.), *Beyond nature and nurture*, 268–286. Chicago: Chicago University Press.

Gries, Stefan T. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34(4). 365–399.

Gries, Stefan T. & Anatol Stefanowitsch. 2004. Extending collostructional analysis. A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1). 97–129.

Gries, Stefan T. & Stefanie Wulff. 2009. Psycholinguistic and corpus-linguistic evidence for L2 constructions. *Annual Review of Cognitive Linguistics* 7. 164–187.

Hall, Joan K., An Cheng & Matthew T. Carlson. 2006. Reconceptualizing multicompetence as a theory of language knowledge. *Applied Linguistics* 27(2). 220–40.

Hall, Christopher J. 2013. Cognitive contributions to plurilithic views of English and other languages. *Applied Linguistics* 34. 211–231.

Hall, Christopher J. 2014. Moving beyond accuracy: from tests of English to tests of 'Englishing'. *ELT Journal* 68(4). 376–385.

Hall, Christopher J. Forthcoming. Cognitive perspectives on English as a Lingua Franca. In Jennifer Jenkins, Will Baker & Martin Dewey (eds.), *Routledge Handbook of English as a Lingua Franca*. London: Routledge.

Hall, Christopher J. & Areli Reyes. 2009. Cross-linguistic influence in L2 verb frames: the effects of word familiarity and language proficiency. In Alessandro Benati & Karen Roehr (eds.). *Issues in second language proficiency*, 24–44. London: Continuum.

Hall, Christopher J., Denise Newbrand, Peter Ecke, Ulrike Sperr, Vanessa Marchand & Lisa Hayes. 2009. Learners' implicit assumptions about syntactic frames in new L3 words: The role of cognates, typological proximity and L2 status. *Language Learning* 59(1). 153–202.

Hall, Christopher J. & Rachel Wicaksono. 2013. *Changing Englishes: An interactive course for teachers*. Online at: www.yorksj.ac.uk/changing-englishes.

Hall, Christopher J., Rachel Wicaksono, Shu Liu, Yuan Qian & Xiaoqing Xu. 2015. Exploring teachers' ontologies of English. Monolithic conceptions of grammar in a group of Chinese teachers. *International Journal of Applied Linguistics*. doi: 10.1111/ijal.12107.

Han, ZhaoHong & Elaine Tarone (eds.). 2014. *Interlanguage: Forty years later*. Amsterdam: John Benjamins.

Harris, Roy. 1981. *The language myth*. London: Duckworth.

Holme, Randal. 2010. A construction grammar for the classroom. *International Review of Applied Linguistics* 48. 355–377.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Housen, Alex, Folkert Kuiken & Ineke Vedder. 2012. Complexity, accuracy and fluency. Definitions, measurement and research. In Alex Housen, Folkert Kuiken & Ineke Vedder (eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 1–20. Amsterdam: John Benjamins.

Hynninen, Niina. 2013. *Language regulation in English as a Lingua Franca. Exploring language-regulatory practices*. PhD thesis, University of Helsinki. https://helda.helsinki.fi/bitstream/handle/10138/38290/language.pdf?sequence=1.

Jørgensen, Normann. 2008. Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism* 5(3). 161–176.

Kramsch, Claire. 1998. The privilege of the intercultural speaker. In Michael Byram & Michael E. Fleming (eds.), *Language learning in intercultural perspective: Approaches through drama and ethnography*, 16–31. Cambridge: Cambridge University Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Volume 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Lardiere, Donna. 2006. *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum.

Larsen-Freeman, Dianne. 2014. Another step to be taken—Rethinking the end point of the Interlanguage Continuum. In ZhaoHong Han & Elaine Tarone (eds.), *Interlanguage: Forty years later*, 203–220. Amsterdam: John Benjamins.

Leech, Geoffrey. 2014. *The pragmatics of politeness*. Oxford: Oxford University Press.

Love, Harold. 2002. *Attributing authorship: An introduction*. Cambridge: Cambridge University Press.

Marian, Viorica, Henrike K. Blumenfeld & Margarita Kaushanskaya. 2007. The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research* 50(4). 940–967.

Mauranen, Anna. 2012. *Exploring ELF. Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.

Mellow, J. Dean. 2006. The emergence of second language syntax: A case study of the acquisition of relative clauses. *Applied Linguistics* 27(4). 645–670.

Mollin, S. 2009. "I entirely understand" is a Blairism. The methodology of identifying idiolectal collocations. *International Journal of Corpus Linguistics* 14(3). 367–392.

Neff, JoAnne, Mercedes Diez, Rosa Prieto & Carmen Sancho. 2003. Contrasting learner corpora; the use of modal and reporting verbs in the expression of writer stance. In Sylviane Granger & Stephanie Petch-Tyson (eds.), *Extending the scope of corpus-based research. New applications, new challenges*, 211–230. Amsterdam: Rodopi.

Otheguy, Ricardo, Ofelia García & Wallis Reid. 2015. Clarifying translanguaging and deconstructing names languages: A perspective from linguistics. *Applied Linguistics Review* 6(3). 281–307.

Pennycook, Alistair. 2009. Plurilithic Englishes: Towards a 3D model. In Kumiko Murata & Jennifer Jenkins (eds.), *Global Englishes in Asian contexts. Current and future debates*, 194–207. Basingstoke, UK: Palgrave Macmillan.

Peterson, Kelly, Matt Hohensee & Fei Xia. 2011. Email formality in the workplace: A case study on the Enron corpus. *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, 86–95. Portland, Oregon, 23 June 2011.

Pickering, Martin J. & Victor S. Ferreira. 2008. Structural priming: A critical review. *Psychological Bulletin* 134(3). 427–459.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM

Robinson, Peter & Nick Ellis. 2008. Conclusion: Cognitive linguistics, second language acquisition and L2 instruction—issues for research. In Peter Robinson & Nick Ellis (eds.), *Handbook of cognitive linguistics and second language acquisition*, 489–545. London: Routledge.

Römer, Ute. 2004. A corpus-driven approach to modal auxiliaries and their didactics. In John M. Sinclair (ed.), *How to use corpora in language teaching*, 185–199. Amsterdam: John Benjamins.

Scott, Mike & Christopher Tribble. 2006. *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

Seidlhofer, Barbara. 2011. *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Styler, Will. 2011. *The EnronSent Corpus*. Technical Report 01-2011. Boulder, CO: University of Colorado at Boulder Institute of Cognitive Science.

Tomasello, Michael. 2003. *Constructing a language. A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Ullman, Michael T. 2014. The declarative/procedural model. A neurobiologically motivated theory of first and second language. In Bill VanPatten & Jessica Williams (eds.), *Theories in second language acquisition: An introduction*, 2nd edn, 135–158. London: Routledge.

VanPatten, Bill. 2010. The two faces of SLA: Mental representation and skill. *International Journal of English Studies* 10(1). 1–18.

Vetchinnikova, Svetlana. 2015. Usage-based recycling or creative exploitation of the shared code? The case of phraseological patterning. *Journal of English as a Lingua Franca* 4(2). 223–252.

Vetchinnikova, Svetlana. Forthcoming. On the relationship between the cognitive and the communal: A complex systems perspective. In Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds.), *Changing English: Global and local perspectives*. Berlin: Mouton de Gruyter.

Watts, Richard J. 2003. *Politeness*. Cambridge: Cambridge University Press.

Wulff, Stefanie & Stefan T. Gries. 2011. Corpus-driven methods for assessing accuracy in learner production. In Peter Robinson (ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance*, 61–87. Amsterdam: John Benjamins.

Year, Jungeun & Peter Gordon. 2009. Korean speakers' acquisition of the English ditransitive construction: The role of verb prototype, input distribution, and frequency. *The Modern Language Journal*, 93(3). 399–417.

Authenticated | c.hall@yorksj.ac.uk author's copy
Download Date | 6/30/16 10:35 AM