

Rethinking interviews as representations of spoken language in learner corpora

Pascual Pérez-Paredes^a – Geraldine Mark^b
University of Murcia^a / Spain
Cardiff University^b / United Kingdom

Abstract – Following the call to examine the role of learner corpora in SLA research (Bell and Payant 2021), this paper discusses spoken learner corpora —specifically those collected through interviews— and considers the aspects of spoken learner language that they represent. The interview is both an elicitation technique and a complex genre. The overlapping of the two conceptualisations under the same term may give rise to problems of definition about the nature of the language collected and, as a consequence, to difficulties in interpretation when assessing the characteristics of spoken learner data. In this paper, we use original research to exemplify some of the areas that need some rethinking in terms of future reconceptualisation about how spoken data are collected and analysed. This research shows the potential impact of the degree of interviewer/interviewee engagement with the task, suggesting that not enough attention has been paid to the genre of interview in learner corpus research.

Keywords – learner corpus research; spoken language; task; interview; representativeness

1. INTRODUCTION¹

This paper touches broadly on the ubiquity of the ‘interview’ as a means of gathering learner language and the challenges of using interviews to represent everyday spoken language. It seeks to show the limitations of some standard practices of gathering spoken learner data, focusing particularly on the practice of interviewing and the language produced in interview

¹ The authors would like to thank the anonymous reviewers and the editors for their invaluable insights and constructive feedback. Their thoughtful comments have significantly contributed to the enhancement and clarity of this manuscript.



tasks. In doing so, it considers the challenges of defining the interview as a genre. Our research asks whether the data resulting from interview tasks offer a valid representation of spoken learner language and draws attention to its use as a spoken learner language benchmark. It argues that, if corpus linguists are to claim that data collected through interviews are indeed representative of a given mode (e.g., written vs. spoken), we need to pay specific attention to ensuring that the interactive features of everyday spoken language are represented. The research further contends that, if the data are deemed not to be representative of spoken language, we need to understand the language in use that these data do represent (Crawford 2022) and must be wary of using them to investigate the spoken interactional proficiency of learners. We offer some evidence from spoken learner corpora about how the role of interviewers may impact the final product used by researchers when discussing L2 spoken data.

The paper is a relevant contribution to the field of learner language research, as it affects the ways in which the scope of the findings and claims that derive from the analysis of learner corpora are conceptualised and framed. We seek to contribute to methodological innovation in the field by recommending further reconceptualisation of the use of interview data and interviewers' role in collecting such data, as well as pointing to other innovative means for gathering spoken learner language.

In what follows, we first look at definitions and representations of the interview as a genre, and how it is used as a tool in data collection and analysis in L1 and L2 spoken corpora (Section 2). We discuss what we understand by 'spoken language', particularly in the field of language learning and teaching, and the interview as a representation of this. We then turn our attention to the use of the interview in collections of learner language, looking specifically at the role of the interviewer and the effect on interaction. In Section 3, we demonstrate this through original research looking at examples from existing L2 and L1 corpora —such as the *Louvain International Database of Spoken English Interlanguage* (LINDSEI; Gilquin et al. 2010)² and the *Louvain Corpus of Native English Conversation*

² <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html>

(LOCNEC; De Cock 2004)³— and conclude that the interview offers a narrow representation of spoken language that does not necessarily allow for evidence of the interactional features of everyday spoken language (Section 4).

2. THE INTERVIEW IN CORPUS LINGUISTICS: DEFINITIONS AND REPRESENTATIONS

In this section, we show how interview data are represented in L1 corpora (Section 2.1) and contrast this with the ways in which interviews are used in learner spoken data collection (Section 2.2). We consider both the impact of the task and the interviewer on the final L2 product (Section 2.3).

2.1. Interviews and their use in L1 spoken corpora

The super-genre ‘interview’ subsumes diverging assumptions of how speakers construct interaction within the boundaries of the communicative situation where the interview takes place. Despite the apparent simplicity with which one might think that the interview genre can be defined, the concretisation of the genre in corpus linguistics (henceforth CL) presents some challenges. McCarthy and Carter (1994: 191) have defined the interview as a genre that is

sufficiently broad to take in a variety of sub-types from minimally interactional, maximally transactional events (e.g., formal political interviews) to maximally interactional personal encounters (e.g., chat shows, therapeutic interviews).

At one end of the scale, the ‘maximally interactional’ is typically focused on relationship creation, where shared involvement and reciprocity and, for example, expression of stance (Biber *et al.* 1999), is evident in the discourse. At the other end, the ‘maximally transactional’ focuses principally on information provision and exchange. This distinction allows corpus linguists to conceptualise interaction as a key variable to be factored into the data collection

³ <https://corpora.uclouvain.be/catalog/corpus/locnec>

process, the ways in which interviews are designed during the corpus building process, and the analysis of the genre.

Terminological problems arise when it comes to categorising interviews in CL. Let us take, for example, the *British Component of the International Corpus of English (ICE-GB)*.⁴ In ICE-GB, around three per cent of the spoken data (20,000 words) are contributed by broadcast interviews (Lee 2002), which are texts that display minimal interaction. In turn, face-to-face conversations represent around 15 per cent of the spoken data in the corpus. Broadcast interviews are classified as ‘public dialogue’, while face-to-face conversations are conceptualised as ‘private dialogue’ (Lee 2002). According to Lee (2002), both broadcast interviews and face-to-face conversations belong to the same medium (spoken) and interaction (dialogue), but they represent different super-genres and subgenres, as shown in Figure 1.

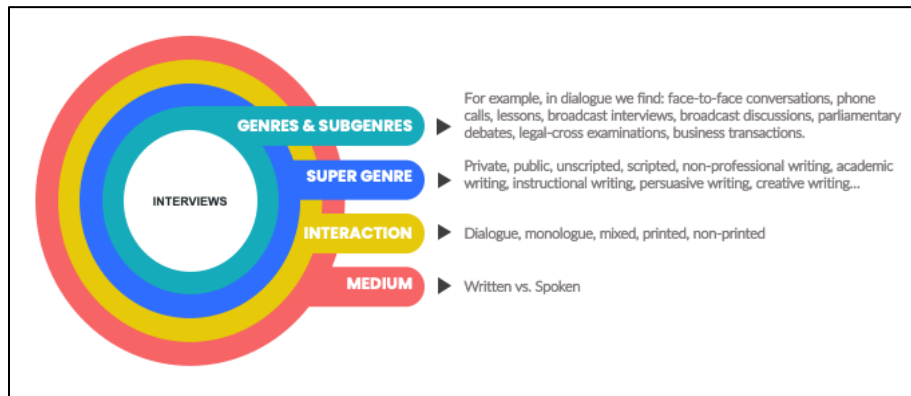


Figure 1: Text classification in ICE-GB. Adapted from Lee (2002)

Broadcast interviews and face-to-face conversations represent, therefore, instances of different dialogic subgenres. In the *Spoken Component of the First British National Corpus (Spoken BNC1994)*,⁵ the super genre ‘interview’ mainly includes job interviews, history interviews, and narrative broadcast interviews. Interviews are either institutionally situated or are reflective of professional practices. In the *Spoken Component of the Second British National Corpus (Spoken BNC2014; Love 2020)*,⁶ which comprises ten million words of

⁴ <http://ice-corpora.net/ice/index.html>

⁵ <http://www.natcorp.ox.ac.uk>

⁶ <http://corpora.lancs.ac.uk/bnc2014>

spoken English gathered from the UK public between 2012 and 2016, the data are classified into 992 activity types, representative of everyday spoken interactions (e.g., chatting with friends or colleagues chatting in coffee breaks). Interviews feature in only two of these activity types and constitute 0.1 per cent of the Spoken BNC2014. As we write this paper, the *Lancaster-Northern Arizona Corpus of American Spoken English* (LANA-CASE), which is the American English counterpart to the BNC2014, is being compiled. This corpus aims to collect 1,000 hours of self-recorded conversation between two or three adults (see Hanks *et al.* 2024). Interviews, however, are not represented in LANA-CASE. Judging from the relative scarcity of interviews in the spoken corpora outlined here, we may conclude that the interview genre is perceived as marginal in representing a benchmark for everyday spoken language in English language corpora, whereas the conversation genre is pervasive and perceived as highly representative of spoken language.

The interview genre, when represented in English corpora, seems not to rely exclusively on interlocutors' shared interpersonal context, but rather favours the construction of texts that depend on public formal interactions. For example, Biber (1995) reports that, from a multi-dimensional analysis perspective, Dimension 1 shows how involved and interactional interviews score in comparison with other registers. A register with a high score on this Dimension exhibits frequent occurrences of private verbs such as *think*, omissions of *that*, present tense verb phrases, contractions, and second person pronouns (Biber 1995: 117). As Biber (1995: 151) notes, high scores show evidence of "highly interactive, affective discourse produced under real-time constraints, whether spoken or written." Telephone conversations, together with face-to-face conversations, show the highest score on this Dimension, while interviews fall below, close to personal letters and spontaneous speeches. In other words, Biber shows that the interview data are not necessarily representative of an involved, interactional register, whereas conversations are.

For Biber *et al.* (1999) and Biber *et al.* (2021), conversations are grounded in a shared context where specification of meaning is avoided and the contextual background provides the backdrop for many of the exchanges that take place between members of the family, friends, or fellow workers. These exchanges are characterised, among others, by the pervasive use of non-clausal elements, personal pronouns, so-called inserts, and low lexical

density. Interlocutors in a conversation dynamically co-construct the discourse, taking turns and adapting, as Biber *et al.* (1999: 1039) state,

their expression to the ongoing exchange [...] the to-and-fro movement of conversation between speaker [...] the occurrence of utterances which [...] either form a response or elicit a response [...] known as adjacency pairs.

In McCarthy's (2010: 7) terms, there is a "shared responsibility" among interlocutors in face-to-face interactions to maintain a "flow across turn boundaries [...] captured by the metaphor of confluence, reflecting the jointly produced artefact which constitutes an efficient and successful interaction." This adaptation to the flow of discourse is achieved additionally through clause co-construction, back-channelling, hedging, and discourse marking, often exemplified through the use of what Hasselgreen (2004) labels 'smallwords' (e.g., *you know, sort of, really, just, well*), which are —as McCarthy (2010: 11) points out— "interactive and flow-sustaining" in everyday conversation.

So far, we have seen that the interview super genre is likely to subsume different assumptions and approaches to how speakers construct interaction within the boundaries of the communicative situation where the interview takes place. A crucial consideration here is that if the role of interaction is not thought out in advance, corpus compilers may collect data which, under the same umbrella term, may represent very different types of subgenres. This can be partly explained by the fact that the field of CL has a long tradition of unobtrusively collecting data (Stubbs 2007) that has already been produced by members of a community, and where the use of interviews as an elicitation method is not particularly frequent. As Koester (2022) argues, interviews are not usually associated with corpus studies other than in the learner corpus domain. In fact, Koester sees interviews as a complement to corpus data rather than as one of the main genres in L1 corpora.

The debate about interviews in CL contrasts with the attention paid to the interviewer's role in social sciences and in sciences related to education or applied linguistics (Mann 2011). In social and education sciences, the literature abounds on the kind of interaction and dynamics that interviewers need to foster during interviews to collect data that are dense enough to support a qualitative analysis (Cohen *et al.* 2017). In fact, there exists a variety of conceptualisations about the nature of interviews in social science research. Mann (2011),

for instance, has discussed two metaphors which are revealing from an epistemological perspective. In one of them, the interviewer is a traveller, which evokes a post-modern constructivist position that contrasts with the interviewer as “the positivist miner extracting nuggets of raw truth” (Mann 2011: 7). For Mann, all interviews are unavoidably meaning-making ventures where the interviewer’s contribution to the co-construction of the interview content must be explicitly acknowledged and may thus become a topic for analysis itself. With this in mind, here we argue that interviews as elicitation instruments for data collection do seem to align with interviews as a genre. The reason for the lack of debate about the role of the interviewer in eliciting language may well lie in 1) the positivist nature of the data collection procedure followed by most researchers in CL (mining nuggets), and 2) the different conceptualisations about the nature of what interview data stand for. The invisibility of the interviewers in co-constructing interview data (Jones 2022) is, perhaps, partly explained by a view in CL stressing that texts can be collected but cannot be manipulated directly by the corpus designers, as this would alter the nature of the observed phenomena.

2.2. Interviews in learner corpus research

In learner corpus design, in stark contrast to L1 corpora, interviews are one of the main genres (Tracy-Ventura *et al.* 2021) that, together with argumentative essays, academic writing, and narratives, are most widely represented (Gilquin 2021). Of the 201 learner corpora identified in the *Learner Corpora around the World* database (LCW),⁷ around a quarter (49) are categorised as spoken. Of these, 37 per cent are described as containing interviews. For Tracy-Ventura *et al.* (2021: 414),

oral corpora often consist of interviews (primarily between a researcher and the participant), narrative retells based on pictures or silent films, or monologues based on a prompt.

These tend to be tasks that EFL learners are familiar with. Both LINDSEI (Gilquin *et al.* 2010) and the *Trinity Lancaster Corpus* (TLC; Gablasova *et al.* 2019)⁸ —to cite two of the

⁷ <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>

⁸ <https://cass.lancs.ac.uk/trinity-lancaster-corpus/>

most relevant spoken learner corpora— represent some of these spoken tasks that learners typically engage in. These two corpora differ, at least, in one key area of corpus design. The LINDSEI tasks were designed to gather, observe, and analyse learner language from a contrastive interlanguage analysis perspective (Gilquin *et al.* 2010), allowing for comparisons between different L1 groups (as well as L1 speakers through the LOCNEC corpus). In turn, TLC is the result of the language produced by exam candidates, comprising tasks from the Graded Examinations in Spoken English (GESE), developed and administered by Trinity College London.

A third important learner corpus, widely used in learner corpus research, is the spoken dialogue component of the *International Corpus Network of Asian Learners of English* (ICNALE SD),⁹ which includes “approximately 270-hour videos of oral interviews conducted with 405 college students in ten regions in Asia” (Ishikawa 2019: 154). The TLC L2 data, as reported in Gablasova *et al.* (2019), include over 2,000 L2 speakers from different cultural and L1 backgrounds and contain 4.2 million words (tokens) of transcribed spoken interaction between exam candidates (L2 speakers of English) and examiners (L1 speakers of English), which makes it the largest corpus of spoken L2 English at the time of writing. The TLC tasks show a combination of controlled tasks and an explicit task description available to the exam candidates. According to Gablasova *et al.* (2019: 133), the interaction “develops dynamically between the L1 and L2 speaker.” As regards topics in TLC, Gablasova *et al.* (2019: 154) state that

in two tasks (presentation and discussion) the topic is selected freely by the candidate, while in the other two tasks (interactive task and conversation), the topics are selected by the examiner.

For Gablasova *et al.* (2019: 147), comparability in terms of linguistic setting and speaking tasks is key when collecting the learner data during the interview, as “all interviews are conducted by trained examiners from Trinity College London following the same principles as in the L2 interviews.”

⁹ <https://language.sakura.ne.jp/icnale/>

LINDSEI contains oral data produced by upper-intermediate and advanced learners of English from different L1 backgrounds. The LINDSEI CD-ROM (Gilquin *et al.* 2010) comprises 11 L1 components: Bulgarian, Chinese, Dutch, French, German, Greek, Italian, Japanese, Polish, Spanish, and Swedish. All components follow the same structure. The interview is made up of the same three tasks: a set topic, a free discussion, and a picture description. The corpus includes 554 interviews, totalling 1,080,232 words. The interviews were transcribed and marked-up according to the same conventions by the different national teams. In the first LINDSEI task, the interviewees were given three topics and were asked to choose one of them, think about it for a few minutes without taking notes, and then talk about it.¹⁰

This first task of the interview is therefore predominantly monologic. After this, the interviewer goes on to ask questions and interacts with the learner. The questions address the topic chosen in the first task and then other subjects, including life at university, hobbies, or travels abroad (Gilquin *et al.* 2010). In the third task of the interview, the interviewees were asked to look at four pictures making up a story and to describe what they saw. For Friginal *et al.* (2017: 43), the LINDSEI interviews illustrate “how learners shift their use of various linguistic features, covering a range of discourse domains” and provide “a wealth of information on how learners actually use language in interviews.”

Alongside LINDSEI, LOCNEC—a comparable corpus of interviews with L1 speakers of English, designed to represent L1 conversation—was compiled to provide a baseline for L1-L2 comparisons. LOCNEC mirrors the tasks and interview approach in LINDSEI and is made up of 50 interviews. Aguado *et al.* (2012) completed an additional 28 interviews at Manchester Metropolitan University in 2006 following the same design criteria.

Friginal and Polat (2015) conducted a multi-dimensional analysis of LINDSEI to identify the dimensions of English learner talk and interpret the resulting dimensions,

¹⁰ In Gilquin *et al.* (2010), the first topic is ‘an experience you have had which has taught you an important lesson. You should describe the experience and say what you have learned from it’. The second topic is ‘a country you have visited which has impressed you. Describe your visit and say why you found the country particularly impressive’. Finally, the third topic is ‘a film/play you have seen which you thought was particularly good/bad. Describe the film/play and say why you thought it was good/bad’.

comparing how they are distributed across the different L1 backgrounds. One of the new identified dimensions shows that the picture description task is functionally distinct from the other two tasks. Pérez-Paredes and Sánchez-Tornel (2019) support this finding in their multi-dimensional analysis investigation of the extended LOCNEC, observing a statistically significant difference between the interactive part and the picture description. However, in the LINDSEI interview, the different tasks do not align with one particular type of interaction, or with any of the subgenres seen in Figure 1, as they arguably represent EFL classroom tasks that spread over different types of interactions and super genres (i.e., narratives and descriptions).

Not all spoken learner corpora, however, prioritise interviews as a means of eliciting data. In the *Michigan Corpus of Academic Spoken English* (MICASE),¹¹ which comprises 1.8 million words from lectures and classroom discussions, 12 per cent of the speakers have an L1 other than English. Likewise, the *TOEFL 2000 Spoken and Written Academic Language* (T2K-SWAL)¹² was designed to provide a basis for test construction and validation of spoken and written registers in US universities (Biber *et al.* 2004) and captures the language as used by students and lectures across study groups, service encounters, or class sessions. Similarly, the *British Academic Spoken English Corpus* (BASE)¹³ represents language used in academic contexts such as seminars or lectures and includes a small amount of L2 learner output (Friginal *et al.* 2017). Also, the *Vienna-Oxford International Corpus of English* (VOICE)¹⁴ captures interactions of spoken English as a lingua franca and also includes interviews, although other elicitation techniques such as seminar discussions, panels, or meetings are more frequent. The VOICE compilation criteria emphasise the lingua franca status of the interactions represented in the data but the L2 learning dimension is not an explicit focus in its design.

The interview genre and the roles of interviewers have not received much attention in specialised CL literature. When considered in the context of their scarcity in L1 spoken

¹¹ <https://quod.lib.umich.edu/m/micase/>

¹² http://universal.elra.info/product_info.php?cPath=42_43&products_id=1497

¹³ https://www.reading.ac.uk/acadepts/ll/base_corpus/

¹⁴ <https://voice.acdh.oeaw.ac.at/>

corpora, as described above, this is not entirely surprising. In the second edition of *The Routledge Handbook of Corpus Linguistics* (O’Keeffe and McCarthy 2022), the term ‘interviewer’ is used in only two occasions, and neither the role of the interviewer nor their ability to influence language during interviews is discussed. In *A Practical Handbook of Corpus Linguistics* (Paquot and Gries 2021), the term ‘interviewer’ is simply not found, while interviews are regarded as genres and interview topics are occasionally referred to, in the context of L1 corpora, as a source of bias that may lead to the overrepresentation of linguistic features (Gut 2012) such as, for instance, the past tense in the *Freiburg English Dialect* corpus (FRED; Anderwald and Wagner 2007). In *The Routledge Handbook of Second Language Acquisition and Corpora* (Tracy-Ventura and Paquot 2021), the term ‘interviewer’ is found only once, when discussing the design of LINDSEI (Gilquin *et al.* 2010).

Despite the abovementioned absence of reference in the literature and the sparse representation of the interview genre in L1 spoken corpora, the interview itself is ubiquitous in L2 spoken data. For example, Bell *et al.* (2021: 218) elected to use an interview task for the spoken component of their corpus study on L2 grammatical development because of 1) its popularity with previous studies and researchers and 2) “potential influences that task condition (monologic vs dialogic) can have on language production.” The reference to “potential influences” is important because, as we have seen, both LINDSEI and TLC include dialogic tasks where interaction might be expected to emerge. However, as Gráf (2017: 29) points out, in LINDSEI, the execution of these tasks, is “left very much to the coordinator’s own experience or initiative.” While Friginal *et al.* (2017) have highlighted the usefulness of interview data to investigate interaction, the nature of the interaction, as pointed out by Gráf (2017), is not totally clear in terms of the linguistic and functional characteristics of the corpus data collected. Gráf (2017) has in fact highlighted some open questions and debates in the research design of standard spoken learner corpora, including the lack of concrete research questions and the lack of specific guidelines in terms of how the interviews are carried out regarding their communicative content. Gráf wonders whether interlocutors are expected to produce certain grammatical or lexical patterns and, if so, how this is supposed to be achieved. Similarly, he has also expressed doubts about the weight of monologic tasks in the design of oral corpora and the interviewers’ active/passive role in the conversational construction of the interview. This leads us to consider the role of the interviewer in the

collection of L2 data and the impact this has on the shape and profile of the data we collect. What role does the interviewer play? Is the interviewer an interlocutor, a conversant, a facilitator, a co-ordinator, a passive listener? As Bell *et al.* (2021: 218) argue, we return to the “potential influences that task condition (monologic vs. dialogic) can have on language production,” and the effect of (degrees of) engagement from the interviewer in the discourse.

2.3. The interviewer in learner corpus research

McCarthy and Carter (1994) explore issues related to integrating discourse and conversational practices into the language learning classroom. They highlight the question of whether learner performance or output engage with the discourse process in a learning context. They explore this by investigating the differences in the same interview tasks firstly undertaken between two L1 speakers and secondly between an L1 and an L2 speaker in a learning context. Both interviewers are given the same brief for the tasks. McCarthy and Carter (1994: 189) note that in the L1:L1 interview the speakers orient themselves towards a more relational-style interview sub-genre, in which their joint goals are interactional and not essential to the “transactional structure of the encounter,” whereas the L1:L2 interview follows a transactional question-answer structure and gets the job of the interview done efficiently, but with less involvement between participants. McCarthy and Carter (1994: 191) point out that the non-intimate interview as a genre “is not well attuned to interactional features: reciprocity and affective convergence are not at all among its goals.” They argue that the interviewer may wish to behave in a more ‘human’ way, but the restrictions from the setting up of the task as a transactional encounter may be a barrier to this. This may result in less interactional output, where features of everyday spoken discourse are not necessarily represented. Unlike in everyday conversation, in learner data collection, the interviewer manages time and topic shifts.

We now return to the pioneering LINDSEI corpus to explore the issues pointed out in McCarthy and Carter (1994), and the effect of these on the data collected. The LINDSEI compilation guidelines (Gilquin *et al.* 2010) specify that the objective of the LINDSEI project was to collect spoken interlanguage during informal interviews. As outlined in Section 2.1, the interviews had to follow a pattern in which the interviewers had some freedom as regards

the actual questions they asked the learners. The guidelines recommend minimal interruption from the interviewer. This is an elicitation task during which the interviewer facilitates the goal of interlanguage collection through a series of questions and answers, not necessarily interactions. The LINDSEI guidelines make this clear and the interviewer and learner turns and tokens are itemised separately, since the learners' turns will be of the utmost importance. The guidelines also refer to the relationship between the interviewer and the interviewee and acknowledge that the status of the interviewer in relation to the interviewee may impact the progress of the interview and its formality. Here we suggest that hedging the impact underestimates the effect of the interviewer on the proceedings and presents one understanding of spoken language in which the monologic or dialogic nature of the data is not of primary importance. Depending on the consistency of the involvement of the interviewer and the degree to which reciprocity is encouraged across the data, the result may, on the one hand, tend towards a representation of spoken language which potentially orients towards written norms (characterised by monologic, transactional responses to questions) or, on the other, towards a representation of spoken language which reflects co-construction, and is dialogic and interactional. For the purposes of learner corpus research, it is problematic if 1) the data purport to contain one representation of language in use in a given context but does not, and 2) the data are used to make judgements about learner language proficiency. This introduces variables which are potentially ignored in traditional learner corpus research.

In practice, the variation in the freedom to interact by interviewers and in the degree to which interviewee and interviewer understand the needs of the genre is critical to the data produced and to its interpretation and comparability (see Gráf 2017). This becomes increasingly important in two research contexts: 1) when interview-elicited data are used to investigate features of L2 spoken interactional language, and 2) when interview-elicited data are used as benchmarks for spoken learner data. To date, the LINDSEI data are frequently used for both these representative and comparative purposes. For example, Larsson *et al.* (2023) use LINDSEI to represent learner speech when exploring development of grammatical complexity in writing, and investigating whether learners move away from speech-like production towards more advanced written production. They state that they “use LINDSEI to represent a benchmark for speech and ICLE to represent a benchmark for

writing” (Larson *et al.* 2023: 8). Similarly, Friginal *et al.* (2017: 45) select LINDSEI because it

is especially well suited to investigations of learner talk because of its large size, representativeness (as noted earlier, 11 L1 backgrounds with approximately 50 interviews each), and the consistency of its implementation.

Castello (2023) also uses LINDSEI and LOCNEC to represent spoken interaction when investigating stance adverbials in discourse and conversation management from spoken English interaction. Likewise, in a study exploring the use of *well* as a discourse marker, Aijmer (2018) uses the Swedish component of LINDSEI and L1 LOCNEC to investigate uses of *well* and finds that the L1 speakers use *well* more frequently than the L2 speakers to signal turn-taking. As Aijmer herself states, she uses the two corpora to examine similarities and differences between the L1 and L2 spoken English and encourages use of the differences as a target for remedial classroom work. Aijmer acknowledges the possible effect of the interview format and notes that other types of interaction (e.g., conversation) may have given different results. This acknowledgment, we believe, gives credence to our argument that equating L2 speakers’ performance in a subset of spoken English (i.e., in this case the LINDSEI interview) with overall L2 speaking performance can be questionable and worthy of further investigation. Aligning with McCarthy and Carter’s (1994) distinction between the kind of interactions that take place in a learner corpus style interview and the same task between two L1 speakers, Crawford (2022: 93) points out that, as examples of dialogic discourse between a learner and an interviewer, LINDSEI is “of limited use for those interested in investigating how learners manage face to face conversations.” By any means, this is not to undermine LINDSEI or other similar data. Well-designed corpora, such as LINDSEI, are highly representative of the language used in a given context (Crawford 2022). However, what we are exploring here is the need to be aware of the limitations and variables at play when using the output from interview tasks as a broad representation of spoken learner language.

In the next section, discussing examples from LINDSEI, we exemplify some of the challenges and limitations of using the interview format for data collection and, generally, for the use of interviews as representations of spoken language. We point to an interviewer

effect on the learner data and demonstrate how considerations of the interaction between the interviewer and the interviewee may inform future protocols for corpus design, collection, and analysis of learner data. We noted above the role of smallwords in everyday conversation and their discourse function in co-construction spoken language (cf. Section 2.1). Since among these, adverbs play an important part and perform multiple roles in the discourse, we have chosen to close in on adverb use and its role in interactivity between participants and across different tasks to exemplify our argument.

3. LIMITATIONS IN THE COLLECTION OF LEARNERS' OUTPUT: SOME EXAMPLES

Adverb functions and their roles in spoken communication have been well documented in corpus studies (Carter and McCarthy 2006; Waters 2013; Beeching 2016; Aijmer 2018; Curry *et al.* 2022, among others). In this section, we demonstrate adverb use across the data. For the purposes of this paper, we have extracted three examples using *really*, *well*, and *maybe* to demonstrate the potential impact of the degree of interviewer/interviewee engagement with the task and the effect on the data produced (see rationale below). We show 1) how varying degrees of interactivity differ in their opportunity for turn-taking, co-construction, and discourse management, 2) how power relations between participants might affect the data collected, and 3) how participants may be struggling to understand the interview genre within this pedagogical interaction. In terms of methodology, we compared three of the LINDSEI learner subcorpora —Spanish, German, and Chinese L1s— alongside a parallel extended version of the L1 English LOCNEC corpus, enlarged with 28 additional interviews (Aguado *et al.* 2012). We arrived at the adverb selection identified above by first using the *Sketch Engine* corpus search tool to extract adverb frequencies using the POS tag for adverbs (RB).¹⁵ The RB tag produced a wide spectrum of forms, not all of which are generally categorised as adverbs (e.g., *yeah*, *not*, *n't*, *but*) in widely considered essential reference grammars (cf. Biber *et al.* 1999 or Carter and McCarthy 2006). We filtered the results using these criteria, resulting in the following highest frequency ranking items: *so*,

¹⁵ <https://www.sketchengine.eu/>

very, well, just, really, quite, and maybe. We then examined the different tasks as separate entities comparing them across datasets and documented the differences in use between each L1 and each task. Alongside statistical tests examining the frequency of the individual adverbs across tasks and L1, we analysed the collocational and colligational patterns of usage for the adverbs across task and L1 and provided a qualitative in-depth view of their functional and positional use.

We first looked at the quantitative differences in adverb contributions between interviewer and interviewee and the degree to which this varied among datasets. We analysed the breakdown of all token counts between the interviewer and interviewee content. In some datasets, the interviewer plays a more (inter)active role than in others. For instance, we found varying degrees of participation from the interviewers, ranging from 19.7 per cent of all tokens in the L1 Chinese data to a 30.5 per cent share in the L1 English data. We then considered the effect of the status of the interviewer on the kind of interactions that took place. In 96.35 per cent of the interviews, the interviewer is an L1 English speaker (sometimes the participants' teacher or their language support assistant). We were also interested in the effect of the interviewer directly on the potential mirroring of linguistic choices from the interviewee. Such considerations may have implications for 1) a valid comparability of the four datasets, and 2) representations of spoken learner language.

The three examples below showcase instances where co-construction of meaning and interaction is not always in place. This approach may seem to favour a conceptualisation of the interview as a data elicitation technique or a method with minimum involvement or implication on the side of the interviewer.

3.1. Example 1: Really as interactional device

We analysed the distribution of *really* across the four groups and found that, in our study, *really* was used more frequently by German L1 and English L1 speakers than by Chinese L1 and Spanish L1 speakers. Table 1 shows the raw frequency mean of *really* per speaker per task for the four groups of language.

	Number (individual tasks)	Mean scores	Standard deviation	Standard error
Chinese L1 speakers	159	0.72	1.688	.134
English L1 speakers	228	4.19	5.145	.341
German L1 speakers	149	4.30	4.242	.348
Spanish L1 speakers	144	1.78	3.278	.274
Total/Average	680	2.89	4.252	.163

Table 1: Frequency of *really* per speaker and task in the LINDSEI data

German L1 speakers displayed the highest average frequency per speaker and task while English L1 speakers showed the highest standard deviation. Chinese L1 speakers displayed the lowest average frequency across all speakers and tasks. An ANOVA test confirmed that the overall frequency differences were significant (Welch's $F(3, 336.458) = 53.76, p = .001$). In the German L1 group, we found a significant difference (post hoc Bonferroni pairwise comparisons) between the picture description task and the set topic ($p = .001$) and between the picture description and the free discussion ($p = .001$). In the L1 Chinese group, post hoc Bonferroni pairwise comparisons revealed that there was no significant difference between the picture description task and the set topic ($p = .194$), the set topic and the free discussion task ($p = .061$), and the free discussion and the picture description ($p = 1.00$). In the L1 English group, post hoc Bonferroni pairwise comparisons showed a significant difference between the picture description task and the set topic ($p = .001$), and the picture description and the free discussion ($p = .001$). No differences were attested between the set topic and the free discussion tasks. In the L1 Spanish group, with a Greenhouse-Geisser correction for sphericity, no significant main effect for task type [$F(1.666, 81.650) = .295, p = .705, \text{partial } \eta^2 = .006$] was found.

As described in Section 2.1, the three tasks in LINDSEI vary in terms of their potential for interactivity, with the first task being predominantly monologic, the second predominantly dialogic, and the third a description of a series of pictures. Figure 2 shows the normalised distribution per one million words for *really* across the four datasets and the three tasks. We note low usage among the Chinese L1 speakers, particularly in task 2 (e.g., CH2), a dip in German L1 use in task 2, high usage in the English L1 group for task 1, and a reduction in tasks 2 and 3. We also note the lowest usage among Spanish speakers, though there is a rise in task 3, which is an opposite trend to what happens in the English L1 group.

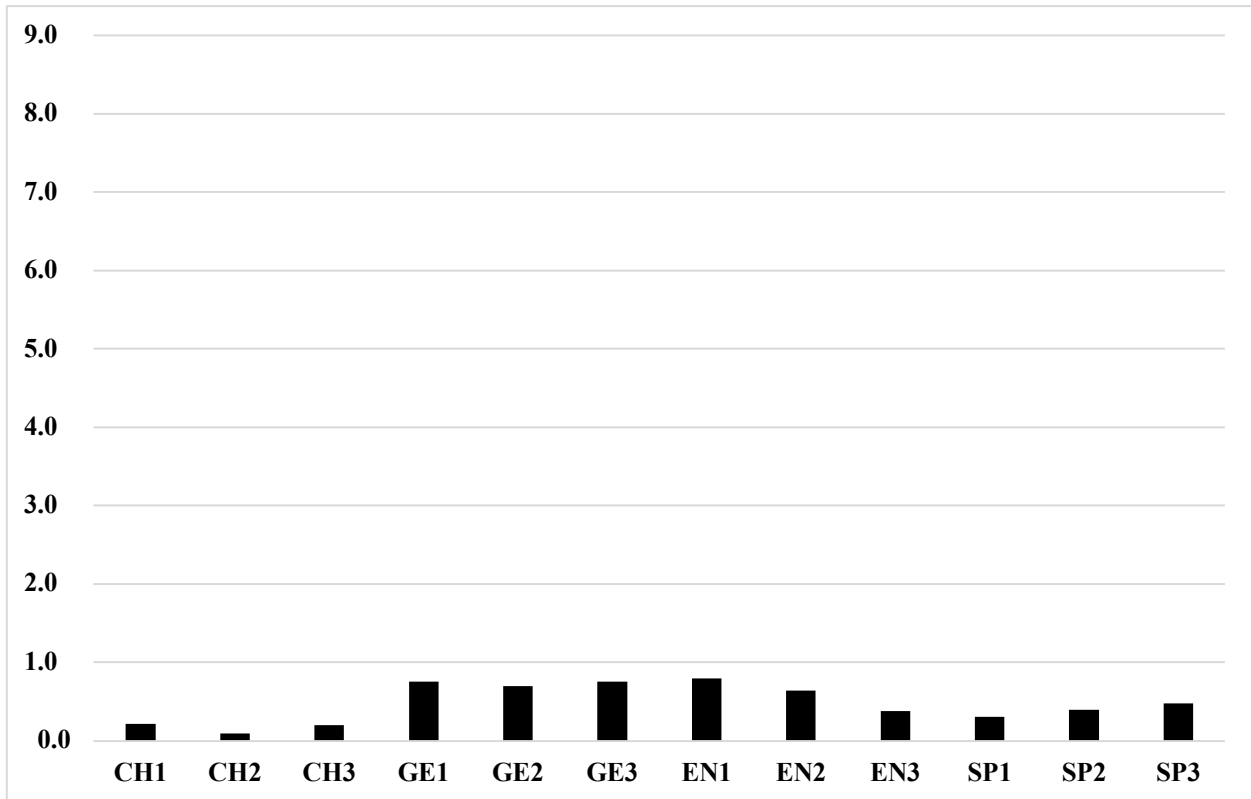


Figure 2: Normalised frequencies (per 1,000 words) of *really* across L1 groups and tasks in the LINDSEI interview

However, an in-depth understanding of the occurrences cannot simply rely on frequency counts of word forms. A qualitative, manual analysis of the functional and positional use of *really* was carried out across the four datasets, using the functional taxonomy illustrated in Table 2.

Function	Example
1 Booster, emphatic, degree	<i>It was really good; I really wanted to see you.</i>
2 Sceptical response	<i>Really? I think that's unlikely.</i>
3 Response token	<i>Oh really.</i>
4 Factually true, actually	<i>It doesn't look like she is really.</i>
5 Hedging	<i>She doesn't really like it.</i>
6 Concessive / summative	<i>It's a bit disappointing really.</i>

Table 2: Functional categorisation of *really* (after Myers 2010)

In all groups, *really* was used most frequently as a booster in a range of positions (e.g., *It was really good; I really wanted to see you*). Our analysis shows that there are no instances of *really* functioning as sceptical response (function 2) and very few instances of it functioning

as a response token (function 3) in the sampled data. Both functions are highly interactive and a feature of back-channelling behaviour in everyday spoken interaction (O’Keeffe and Adolphs 2008) and yet they are not attested in the LINDSEI interview data. We might attribute their absence to 1) the fact that interviewers and interviewees are not encouraged to interact in the tasks in all the sections of the interview, and 2) the possible effect of the relationship between interviewee and interviewer. This may have arguably impacted turn-taking management. Firstly, the opportunity to interact does not always present itself, the transactional nature of the classroom interview genre being a barrier where the interviewer asks a question, and the interviewee provides relevant information by way of response. Secondly, both functions 2 and 3 might require a degree of contradiction of the interviewer by the interviewee which might not be considered appropriate.

By way of contrast, *really* is found 48,492 times in the Spoken BNC2014. Here it occurs ten per cent of the time as a sceptical response or a response token (functions 2 and 3), in the types of contexts illustrated in Figure 3:

want any more scrambled eggs you can have them	really	? I want you to have them okay because then cos
k it's like a hu three hundred and twenty days of rain	really	? a year oh dear we're talking a wet place no you'
the extender that when I arrived a a it didn't recognise	really	? or maybe it was since the extender but we do ha
er had any problem with it so I'd be even tempted to	really	? okay is that the same make? it's the same one r
sarcastic you have no chance in mine at the moment	really	? I've er I've had a temporary kitchen for the last t
had a temporary kitchen for the last twelve years oh	really	? and now why's it temporary? cos you were plan
k my worktops are pardon? my worktops are full real	really	? ye you know it's just I mean it yeah you know I f
Christmas cake I've still got Christmas cake have you	really	? yes wow it's going on forever it's never will neve
s yeah been trying to get them to come out to see us	really	? yes I mean just really easy just get on the ferry y
s actually submitted to the British Museum mus yeah	really	? mm yeah and and and do you just go out in you
n and he was in that and she said it was terrible mm	really	? oh well you c can't you know you can't win them
mm oh she's downstairs well I'll wait for her then oh	really	? you always and I just wai I just er yeah and I go
eah they they're always saying can we go and see ?	really	? yeah you know and every really? month or som
n we go and see ? really? yeah you know and every	really	? month or something yeah yeah you know she's
ald's he would absolutely yeah love it wolf it down oh	really	? yeah he absolutely loves you know Felix and W

Figure 3: Examples of *really* as a sceptical response or a response token in the Spoken BNC2014

Studies on everyday conversation indicate that response tokens perform functions of listenership, ranging from a continuer function—which maintains the flow of discourse using minimal types of responses (i.e., *mmm*, *yeah*)—to a more convergent response, using markers of agreement (*oh right*, *did you?*), or to a more engaged response, such as *really*, *absolutely*, *exactly* (O’Keeffe and Adolphs 2008).

In LINDSEI the lack of engagement between the interviewer and the interviewee is often noticeable through the consistent use of a minimal response such as *mhm*, whose function is to keep the interviewee talking. This is illustrated in extract (1), where Speaker 1 (S1) is the interviewer and Speaker 2 (S2) the interviewee.

Extract (1): CH0104P1

S1: How are you . mm

S2: Fine I'm a little nervous

S1: That's okay . it's not a test . erm so you've chosen a topic

S2: Mm I want choose the topic number one but I wonder

S1: Mm

S2: If . I can change a little bit

S1: Okay

S2: Okay

S1: Mhm

S2: Then I want to talk about my experience in the summer holiday . but . I don't think . it will . teach me a lesson . it just very impressive it make me think a lot

S1: Mhm

S2: Yeah last holiday and I (mm) I stay uni in the university . for . maybe . a month . and because I have a student . and his mother asked me to come here to give her (eh) tutor . tutor jo (eh) ask him ask me . sorry . ask me to teach him . and to improve his English . actually my student is good at the other objects subjects

S1: Mhm

S2: Such as his physics and chemistry's very good but he's (eh) he

In extract (1), there are many opportunities for interaction which are not acted upon by the interviewer. In everyday conversation, S2 might well interpret S1's response as disinterest, as a dispreferred response, but, in this task, the human social element has been removed. In other words, the participants are performing a transaction (McCarthy and Carter 1994) for a pedagogical goal. The only interactional aspect in extract (1) is in the task set-up when S1 asks S2 how they are, to which S2 replies that they are nervous. S1 telling them it's not a test ends the relational interactional element and reverts to the task in hand, marking the transition with *so*. The LINDSEI design does create opportunities for the use of *really* but, in the L2 data, these occur predominantly with a booster or factual meaning function, at a clausal or

phrasal level, rather than at a discourse level with an interactional function. For example, in the Spanish L1 data, 20 per cent of the occurrences use *really* with a factual function (see Table 2), and all of these occur in the third task (the picture description), while 72 per cent of the sampled occurrences were used as a booster, with an equal distribution between phrasal and mid-clause position, as shown in extracts (2) and (3).

Extract (2): SP0107P1

But then if you *really* look into her you can't find anything at all

Extract (3): SP124P2

But I remember it was *really* nice

In summary, if the interviewer does not typically offer anything other than minimal response within the discourse, there are no opportunities for the interviewee to use interactive devices, such as response tokens, as there is nothing to respond to.

3.2. Example 2: Power roles

As has already been discussed, the degree to which the interviewer co-constructs meaning and interacts is variable. In this second example, we do see interactional features emerging in some of the interviews. However, we do not see equal opportunity for their use. Our study shows that use of *maybe* might suggest an interviewer/interviewee relationship effect. For example, the L1 Chinese speakers showed a strong preference for the use of *maybe* to express uncertainty and imprecision. They also used *maybe* to offer possible options or explanations in response to questioning from the interviewer, a use which was predominantly favoured by Spanish and German speakers. This was also favoured in the L1 English sample as a means to give a non-committal response to the interviewer, hence avoiding a contradiction. This is illustrated in extract (4), retrieved from the Chinese data, where S1 is the interviewer and S2 the interviewee. S2 responds to S1's assertion by avoiding a direct disagreement (with the use of *maybe*) while continuing to say the opposite of what S1 has asserted.

Extract (4): CH0115P2

S1: So it's okay to have a little fun as long as you . don't neglect your responsibility

S2: Maybe so . (mm) like a student I should study very hard to . (er) learn more and to . make . my (er) myself more pro like a professor . (uhu)

Extract (5) taken from the German data in part 3 of the interview also demonstrates this non-committal use to avoid disagreement.

Extract (5): GE0144P3

S2: He actually drew her the way she was .. (er) with... all her failures and . (erm) .. blessing no not blessings

S1: . (Erm)

S2: With all failures and

S1: Positive characteristics

S2: Maybe positive characters ... or negate the negative aspects as well

However, extract (5) is an example in which the interviewer (S1) does try to co-construct S2's turn with *positive characteristics*. Rather than disagreeing and rejecting the co-construction, S2 prefaces the partial repetition of S1's turn with a *maybe*, followed by a hesitation and then the opposite reframing of S1's contribution *or negate the negative aspects as well*. As pointed out above, examples such as these, which contain interactional discourse such as co-construction, are not common in the LINDSEI L2 data. Where they do occur, they may be indicative of an unequal power role between interviewer and interviewee and the degree to which the interviewer engages in interaction. This is a variable that may not be consistently applied throughout the data collection but one which—as we have discussed above—will have an effect on the language used and the opportunity to interact. Extract (6), below, is an L1:L1 example retrieved from the LOCNEC data from part 1, which is the task designed to be more monologic. The exchange shows a greater degree of co-construction of meaning between the interviewer and the interviewee, resulting in the occurrence of interactive features such as evaluation, back-channelling, and responses tokens.

Extract (6): LOCNEC53P1

S2: I was gonna actually do that for my project but I looked at it and thought no too much no way it scared me

S1: So it it would have been interesting

S2: Yeah it it's been done since apparently but er

S1: Yeah

S2: Yeah I I opted to go for Billy Joel instead

S1: Mhm

S2: He's a lot more down to earth

S1: Mhm

S2: Really film

S1: But so is the whole book written in that language or

S2: Yeah

S1: Even the descriptions and

S2: Yes it's all from his point of view so he's saying oh yeah he was a bolshy with with and I was saying but occasionally he'll give a translation in brackets just one word

S1: Oh that's nice

In summary, if power roles vary from interview to interview and/or from L1 subcorpus to L1 subcorpus, opportunities of use and issues of comparability between data sets arise.

3.3. Example 3: Making sense of the genre

In this example, we show evidence that speakers do not necessarily understand the demands or purpose of the task and how to engage with it, and that this has an effect on the output. We have already seen that, in the L1 data, there is a greater proportion of participation from the interviewer (30%). For example, in this data, the L1 English speakers are more likely to engage with the exchange, for example, by using adverbs for attitudinal effects to express opinion or soften disagreement, trying to make sense of the task as a conversation. We first illustrate this with the use of *well*. While many of the examples of *well* were found to be used by L1 English speakers as a speech management tool—namely, for pausing, reformulating, and introducing a new turn—there were more examples in the sample of *well* used for attitudinal effect than in the other L1 groups. In extract (7), where the speaking roles are occupied by L1 speakers, the interviewee (S2) uses *well* to contradict the interviewer (S1) and soften the following *no*, followed by an explanation softened with *just*. Similarly, in

extract (8), there is a combination of hesitation and *well* to soften the *no*, followed by an explanation, which is hedged by *I mean*:

Extract (7): LOCNEC35P2

S1: What do you do when it rains when it pours

S2: I get wet

S1: Oh so you you you don't take the bus you

S2: No not usually no

S1: Oh it's very brave

S2: Well no I just don't like the waste of time hanging around for the bus and eh hanging around for the bus

Extract (8): LOCNEC55P2

S1: So it's erm since you you want to do forensics I dunno what that's why you you decided to: er do biology or

S2: Em well no er I mean I'm I'm doing biology because that's it's the one subject I've I've always found easy and I enjoy it

In extract (9), the speakers attempt to construct the interview as a conversation. With the use of *really* at the end of the extract, S2 takes up the interviewer's initial question and threads their answer through the discourse with a final summarising answer ending in *so ... really* in order to answer the question.

Extract (9): LOCNEC1P2

S2: Erm I I'm doing a linguistics minor erm as part of er

S1: And what are you doing

S2: Oh actually it's I don't know if it counts as a minor itself it's part of English literature erm

S1: Ah so you're doing literature and you're doing some courses in linguistics

S2: Yes yeah

S1: Mhm

S2: Er just the one in fact er

S1: Just okay

S2: Yeah

S1: Uhu and er why did you choose literature

S2: Erm well

S1: Good question

S2: I I've always been erm very keen on reading

S1: Mhm

S2: And and in my first year I did English literature and language and French **so** there was reading involved in most of my courses **really**

In contrast, extracts (10) and (11) illustrate the interviewees (S2) explicitly referring to the demands of the task.

Extract (10): CH0105P3

S1: Now could you start now

S2: Okay okay see if I can talk for three or five minutes .. okay may I start now

S1: Yeah

S2: Okay . (mm) I'd like to talk about a film I I I have seen

Extract (11): CH0126P3

S2: Shall I make a make up a story or just tell what happen in this picture

S1: Make up a story

S2: okay ...

In summary, there is variability in the way both interviewers and interviewees orient towards the genre. Some speakers struggle to make sense of what is being asked of them, and whether to engage in co-construction of an interactional nature or to pursue a more transactional question and answer approach. The L1:L1 interviews show evidence of orientation towards the interactional opportunities in the task, whereas the L1:L2 interviews appear to orient towards the transactional.

4. DISCUSSION

Under the term 'interview' we find at least two different conceptualisations: 1) an elicitation technique, and 2) a distinct, albeit complex genre. The overlapping of both conceptualisations under the same term may give rise to problems of definition about the nature of the language collected and, therefore, problems of interpretation when assessing the characteristics of spoken learner data. While learner corpus research may have favoured a miner approach (Mann 2011) to spoken data gathering, it may have inadvertently contributed to the

underrepresentation of some substantial sub-genres in spoken communication, such as face-to-face conversation, where discourse co-construction is key. We argue that, to recreate the communicative situation that takes place during a conversation, it is necessary to rethink the way in which spoken data are collected. We suggest that the way the context and task are set up establishes a particular type of ‘pedagogical interaction’. Discourse Act theory (Allwood 2000; Bunt 2022) argues that spoken communication is multidimensional and complex, relying on a range of activities that, among others, involve task movement, allo-feedback, turn-management, contact management, discourse structuring, partner communication management, or social obligations management. It is unclear how these dimensions shape communication in L2 interchanges (Bunt 2022) where, as in LINDSEI or TLC, L1 speakers are in charge of some of the time and discourse management dimensions. However, the examples in Section 3.3 have shown that the tasks included in LINDSEI did afford, for example, L1 English speakers’ uses of *well* for attitudinal effects. Whether this is the result of power imbalance/balance between speakers of different or the same L1s or a more individual ‘chatty’ approach to the interactions, some of the tasks may facilitate different approaches to participation as speakers in conversations perform a series of functions such as turn grab, turn keep, or turn release (Bunt 2022). The examples in Sections 3.1, 3.2, and 3.3 suggest that for different speakers the interview may draw on different assumptions about the nature of the task and their self-perceived role in the task (i.e., L1 interviewer vs. learner who, following a request from their lecturer, is taking an examination or has volunteered to take part in an interview in her university). Coming back to the use of *really*, in conversations, feedback may refer to different levels of communication such as attention, perception, understanding, evaluation, or execution. We wonder to what extent learners are comfortable trying to engage with the interviewer to give more than a simple answer to a question. Similarly, we may wonder to what extent interviewers feel comfortable facilitating learners’ repetition, co-constructing the discourse, clarifying, offering puzzled faces, nods, or verbal back-channelling, which are common features of real-life spoken interactions. These may create opportunities of use for words such as *really*, *well*, or *actually*, to name but a few.

The examples in Section 3 demonstrate further methodological challenges in collecting and analysing spoken learner data in general. We suggest that some of these challenges are

related to the specific features found in the LINDSEI data that we have explored in this paper but which, we believe, are generalisable to other spoken learner data:

1. Some of the data are more dialogic than others, which has an effect on the degree to which the language produced is more or less transactional or interactional. This has implications for representativeness and corpus design, for instance, in scoping to what degree the resulting data need to display interactional and/or transactional features.
2. The opportunity for interaction is not equal across different interviews and therefore there may be issues of comparability between datasets.
3. The power roles or relationships between the interviewer and the interviewee are likely to influence the language used.
4. The participants' perception of the task is open to interpretation. Language learners in instructed contexts may be more or less familiar with the 'classroom interview' genre. L1 interviewers may be grappling with the restrictions of the task.
5. The data elicited by the interview as a genre presents a limited representation of spoken language. For research that is interested in spoken learner language (broadly defined), and not solely in the performance of learners in spoken language assessments, other types of learner data need to be added in to the mix to complement interview data and present a more comprehensive picture of spoken learner language.

Such considerations must have direct implication not only in terms of comparability but also in the adequate representation of spoken learner language. Learner corpus research is a relatively new field and, while huge strides have been made in understanding learner language so far, there is always further work to be done.

We turn our attention to future research and ask how we can gather a more inclusive broader representation of language learning and learner language (Pérez-Paredes and Mark 2022), particularly in relation to spoken language. We note that the object of focus of previous learner corpus studies can be categorised into discrete features (e.g., lexis, parts of speech, and tenses), composite features (e.g., measures of lexical sophistication and clausal complexity), and constructs (e.g., metadiscourse features and involvement). We point out

that not all areas receive equal attention, and that according to Paquot and Plonsky (2017), only 30 per cent of studies are concerned with discourse and ten per cent with pragmatics. Currently available learner corpora offer a window on a narrow conceptualisation of language learning and language use, which may explain some of the challenges to making corpora more representative of spoken learner language.

LINDSEI and TLC represent different learner language products. While the former shows spontaneous spoken communication, the latter represents a highly practised language test that is familiar to the students before the interview. In this sense, TLC offers a highly contextualised experience which is mediated by the testing nature of the interview, representing the type of language which is fostered by the testing culture of the certificate awarding institution which, in turn, has a trickle-down effect on the types of tasks carried out in classrooms (McCarthy 2020). The former type of corpus, the one represented by LINDSEI, perhaps offers a wider choice of opportunities in terms of use and interaction not necessarily linked to testing practices, and which may offer a more diverse representation of different types of face-to-face interaction. The range of activities used for the collection of the Spoken BNC2014 (Love 2020) such as chatting about general stuff, talking over lunch, academic colleagues chit-chat over coffee, watching TV, discussing fashion, evening catch-up with housemate, talking at book club, dinner conversation about fixing computer, acquaintances having a chat, family advising, etc. may inspire the design of new tasks that may complement our current findings about spoken L2 use. However, as we have seen, the interviewer—who could facilitate unscripted, spontaneous interaction—is encouraged to be absent or is removed from the analysis, for fear of them getting in the way or influence the learner product. As pointed out by Tracy-Ventura *et al.* (2021: 414), “interaction corpora that consist of conversations between learners, or informal conversations between learners and L1/expert target language speakers, are sorely needed.”

There are therefore broad areas where corpus researchers can improve. The first concerns the type of data we are collecting and analysing. As suggested by Friginal *et al.* (2017: 274), future spoken learner corpora may need to address design considerations such as “register coverage [that integrate] more situational contexts, interview questions, and peer response topics or paired activities,” and the role and effect of interaction in the collection

process. Methods of data collection play a pivotal role in shaping the authenticity of language samples within learner corpora. To enhance the representation of authentic interaction, we can make the most of new and developing technologies. This involves designing tasks that mirror real-life interactions and reflect power dynamics present in natural discourse across a variety of contexts and learning scenarios. As suggested in Pérez-Paredes and Mark (2022: 323),

well-designed corpora allow researchers to understand monologic and dialogic communication as they reveal aspects of frequency, collocation, colligation, function and speaker variation that would otherwise remain hidden.

Harnessing technological advancements, such as mobile devices and data collected ‘in the wild’, can provide unfiltered, unscripted language samples, enabling a more genuine and alternative portrayal of spoken language that offers alternatives to the role of interviewers as data collection managers. One such use of technology has been developed by Knight *et al.* (2021) in the construction of the *National Corpus of Contemporary Welsh* (CorCenCC).¹⁶ As Knight *et al.* (2021: 798) point out, the data are gathered through a mobile crowd-sourcing app which is designed to align methods of collection with the Web 2.0 age, and “enables ‘live’ user-generated spoken data collection via crowdsourcing.” A crowdsourcing approach was also taken in the development of the Spoken BNC2014 (McEnery *et al.* 2017; Love 2020). These advancements offer opportunities to bridge the gap between controlled data collection and the intricacies of unscripted, spontaneous linguistic exchanges. Collection methods may allow us to represent a broader conceptualisation of language use both inside and outside of the classroom. This in turn lends us a more inclusive perspective both on language learning product and learning process. Interviewer-led data could be combined with methods that, while remaining ethical and transparent to language learners, can favour the collection of longitudinal data and different types of interaction that are representative of the wealth of turn management options available in conversations which are not staged as interviews or led by L1 interviewers. However, this leads us into a final area of improvement, delving into the applied pedagogical dimension. How do we truly integrate meaningful social

¹⁶ <https://corcenc.org/>

interactions into the classroom? Here we call for a return to the seminal question posed by McCarthy and Carter (1994): How effectively do transactionally oriented tasks represent interactive language use and self-presentation in the classroom? Alongside this, as Curry and Mark (2023) discuss, there is a need to consider how spoken language is represented in educational materials and classroom settings and the subsequent circularity of effect this has on the language used in instructional contexts. By way of example, Fung and Carter (2007: 433) have suggested that the frequency of discourse markers in learner English

reflect[s] the unnatural linguistic input ESL learners are exposed to and the traditional grammar-centred pedagogic focus [on] the literal or propositional (semantic) meanings of words rather than their pragmatic use in spoken language.

Exposure to naturalistic sampling is limited in EFL classroom contexts and more awareness of the spoken register is needed (Mukherjee 2009; Aguado *et al.* 2012). Pérez-Paredes (2019) has suggested the exploitation of annotated spoken pedagogic corpora for secondary school learners to teach pragmatic and lexico-grammatical features of spoken language. EFL textbooks in primary and secondary education levels tend to under-represent spoken interaction (Curry and Mark 2023). In many mainstream coursebooks, task needs are blurred under the generic vague heading ‘speaking’, which often involves monologues and can range from role play to discussion, opinion giving, or response to input. Dialogues on the page may be used to present grammatical or lexical content, rather than attending to dialogic features. Carter and McCarthy (2017) point out that even though spoken grammar has come of age, it is still under the influence of a pedagogy derived from written language. They offer a host of suggestions for future exploration of spoken language including “increased exploitation of spoken learner corpora” as well as “a challenge to the ways in which grammatical and discourse patterns and socio-cultural context are captured” (Carter and McCarthy 2017: 11).

5. CONCLUSION

In this paper, we have highlighted some of the challenges in using the interview as a research tool for the collection of learner spoken data, in an attempt to learn from data collection and analysis thus far and to provide a platform for future work. Our paper contributes to increasing methodological reflection in applied linguistics and CL in what McKinley and

Rose (2017) have identified as a need for researchers to create the spaces in which to discuss not only results but also the methodological considerations that affect their praxis. In spoken corpus research, the design and the construction of corpora, as well as the vagaries of recording, transcription, coding, and marking of spoken data, have received considerable focus (Knight and Adolphs 2022). However, a discussion of the unrealised potential of what the analysis of existing spoken corpora can offer researchers—in the way of insights into the collection and investigation of future learner corpora—is of continued interest and relevance.

Some of the main takeaway messages from our research are the following:

1. The terms ‘interview’ and ‘spoken learner language’ are perhaps too broad and may give rise to a variety of conceptualisations about the nature of the learner data collected. Interpretations of existing corpus data need to consider the complexity involved in defining the nature of interviews (McCarthy and Carter 1994), the nature of the tasks, and the degree of interaction and involvement present.
2. We may be in danger of drawing erroneous conclusions about learner proficiency and the ability to use or not use the interactional features that are characteristic of spoken language when, in reality, the opportunity to do so does not present itself. The interview as a task may not provide the learner the opportunity to use their spoken repertoire.
3. The inconsistency of the occurrence of interaction between the participants in the interview is key to the learner data quality, the resulting genre, and comparability issues with overlapping subgenres across learner corpus research. We have examined the use of a selection of adverbs of relevance in spoken communication in some subsets of LINDSEI and the extended LOCNEC, showing the potential impact of interviewer/interviewee engagement with the task. The role of interviewers is key in terms of the quality and the nature of the collected data. More attention to and awareness of their influence on the resulting data is needed in learner corpus research.
4. We argue that data collected through interview tasks is not fully representative of spoken learner language as discussed. Future spoken corpora will need to explore avenues to represent oral interaction that complement the existing interviewer-led

collection methods, particularly in non-testing situations. This will undoubtedly benefit our insights into the nature of language use, language acquisition (Tracy-Ventura and Myles 2015; Tracy-Ventura *et al.* 2021), and conversational pragmatics (Bunt 2022), and may inform L2 pedagogy with interactional data that may contribute to represent a wider range of sites of L2 learner engagement in language learning and teaching (Carter and McCarthy 2017; Tyler and Ortega 2018).

REFERENCES

- Aguado-Jiménez, Pilar, Pascual Pérez-Paredes and Purificación Sánchez. 2012. Exploring the use of multidimensional analysis of learner language to promote register awareness. *System* 40/1: 90–103.
- Aijmer, Karin. 2018. Intensification with very, really and so in selected varieties of English. In Sebastian Hoffmann, Andrea Sand, Sabine Arndt-Lappe and Lisa Marie Dillmann eds. *Corpora and Lexis*. Leiden. Rodopi, 106–139.
- Allwood, Jens. 2000. An activity-based approach to pragmatics. In Harry Bunt and William Black eds. *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam: John Benjamins, 47–80.
- Anderwald, Lieselotte and Susanne Wagner. 2007. The Freiburg English Dialect Corpus: Applying corpus-linguistic research tools to the analysis of dialect data. In John C. Beal, Karen P. Corrigan and Hermann L. Moisl eds. *Creating and Digitizing Language Corpora Volume 1: Synchronic Databases*. London: Palgrave Macmillan, 35–53.
- Beeching, Kate. 2016. *Pragmatic Markers in British English: Meaning in Social Interaction*. Cambridge: Cambridge University Press.
- Bell, Philippa, Laura Collins and Emma Marsden. 2021. Building an oral and written learner corpus of a school programme: Methodological issues. In Bert Le Bryn and Magali Paquot eds. *Learner Corpus Research Meets Second Language Acquisition*. Cambridge: Cambridge University Press, 214–242.
- Bell, Philippa and Caroline Payant. 2021. Designing learner corpora: Collection, transcription, and annotation. In Nicole Tracy-Ventura and Magali Paquot eds, 53–67.
- Biber, Douglas. 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.
- Biber, Douglas, Susan Conrad, Randi Reppen, Pat Byrd, Marie Helt, Victoria Clark, Viviana Cortes, Eniko Csomay and Alfredo Urzua. 2004. *Representing Language Use in the University: Analysis of the TOEFL 2000 Spoken and Written Academic Language Corpus*. Princeton: Educational Testing Service.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 2021. *Grammar of Spoken and Written English*. Amsterdam: John Benjamins.

- Bunt, Harry. 2022. The multifunctionality of utterances in interactive discourse. In Zihan Yin and Elaine Vine eds. *Multifunctionality in English: Corpora, Language and Academic Literacy Pedagogy*. London: Routledge, 11–29.
- Carter, Ronald and Michael McCarthy. 2006. *Cambridge Grammar of English: A Comprehensive Guide*. Cambridge: Cambridge University Press.
- Carter, Ronald and Michael McCarthy. 2017. Spoken grammar: Where are we and where are we going? *Applied Linguistics* 38/1: 1–20.
- Castello, Erik. 2023. Stance adverbials in spoken English interactions: Insights from corpora of L1 and L2 elicited conversations. *Contrastive Pragmatics* 4/2: 243–273
- Cohen, Louis, Lawrence Manion and Keith Morrison. 2017. *Research Methods in Education*. New York: Routledge.
- Crawford, William J. 2022. Corpora and speaking skills. In Reka R. Jablonkai and Eniko Csomay eds. *The Routledge Handbook of Corpora and English Language Teaching and Learning*. New York: Routledge, 89–101.
- Curry, Niall, Robbie Love and Olivia Goodman. 2022. Adverbs on the move: Investigating publisher application of corpus research on recent language change to ELT coursebook development. *Corpora* 17/1: 1–38.
- Curry, Niall and Geraldine Mark. 2023. Using corpus linguistics in materials development and teacher education. *Second Language Teacher Education* 22: 187–208.
- De Cock, Sylvie. 2004. Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures* 2: 225–246.
- Friginal, Erik, Joseph J. Lee, Brittany Polat and Audrey Roberson. 2017. *Exploring Spoken English Learner Language Using Corpora: Learner Talk*. London: Springer.
- Friginal, Eric and Brittany Polat. 2015. Linguistic dimensions of learner speech in English interviews. *Corpus Linguistics Research* 1: 53–82.
- Fung, Loretta and Ronald Carter. 2007. Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics* 28/3: 410–439.
- Gablasova, Dana, Vaclav Brezina and Tony McEnery. 2019. The Trinity Lancaster Corpus: development, description and application. *International Journal of Learner Corpus Research* 5/2: 126–158.
- Gilquin, Gaëtanelle. 2021. Learner corpora. In Magali Paquot and Stefan Th. Gries eds, 283–303.
- Gilquin, Gaëtanelle, Sylvie De Cock and Sylviane Granger. 2010. *The Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-La-Neuve: Presses universitaires de Louvain.
- Gráf, Tomáš. 2017. *The Story of the Learner Corpus LINDSEI CZ*. Karlova: Univerzita Karlova, Filozofická fakulta.
https://dspace.cuni.cz/bitstream/handle/20.500.11956/97524/1541592_tomas_graf_22-35.pdf?sequence=1&isAllowed=y
- Gut, Ulrike. 2012. The LeaP corpus: A multilingual corpus of spoken learner German and learner English. In Thomas Schmidt and Kai Wörmer eds. *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam: John Benjamins, 3–24.
- Hanks, Elizabeth, Tony McEnery, Jesse Egbert, Tove Larsson, Douglas Biber, Randi Reppen, Paul Baker, Vaclav Brezina, Gavin Brookes, Isobelle Clarke and Raffaella Bottini. 2024. Building LANA-CASE, a spoken corpus of American English

- conversation: Challenges and innovations in corpus compilation. *Research in Corpus Linguistics* 12/2: 24–44.
- Hasselgreen, Angela. 2004. *Testing the Spoken English of Young Norwegians: A study of Test Validity and the Role of 'Smallwords' in Contributing to Pupils' Fluency*. Cambridge: Cambridge University Press.
- Ishikawa, Shin'ichi. 2019. The ICNALE spoken dialogue: A new dataset for the study of Asian learners' performance in L2 English interviews. *English Teaching* 74/4: 153–177.
- Jones, Christian. 2022. What are the basics of analysing a corpus? In Anne O'Keeffe and Michael McCarthy eds, 126–139.
- Knight, Dawn and Svenja Adolphs. 2022. Building a spoken corpus? In Anne O'Keeffe and Michael McCarthy eds, 21–34.
- Knight, Dawn, Fernando Loizides, Steven Neale, Laurence Anthony and Irena Spasić. 2021. Developing computational infrastructure for the CorCenCC corpus: The national corpus of contemporary Welsh. *Language Resources and Evaluation* 55/1: 789–816.
- Koester, Almut. 2022. Building small specialised corpora. In Anne O'Keeffe and Michael McCarthy eds, 48–61.
- Larsson, Tove, Tony Berber Sardinha, Bettany Gray and Douglas Biber. 2023. Exploring early L2 writing development through the lens of grammatical complexity. *Applied Corpus Linguistics* 3/3: 100077. <https://doi.org/10.1016/j.acorp.2023.100077>
- Lee, David. 2002. Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. In Bernhard Kettemann and Georg Marko eds. *Teaching and Learning by Doing Corpus Analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora*. Leiden: Rodopi, 245–292.
- Love, Robbie. 2020. *Overcoming Challenges in Corpus Construction: The Spoken British National Corpus 2014*. London: Routledge.
- Mann, Steve. 2011. A critical review of qualitative interviews in applied linguistics. *Applied Linguistics* 32/1: 6–24.
- McCarthy, Michael. 2010. Spoken fluency revisited. *English Profile Journal* 1. <https://doi.org/10.1017/S2041536210000012>.
- McCarthy, Michael. 2020. *Innovations and Challenges in Grammar*. London: Routledge.
- McCarthy, Michael and Ronald Carter. 1994. *Language as Discourse: Perspectives for Language Teaching*. Routledge: London.
- McEnery, Tony, Robbie Love and Vaclav Brezina. 2017. Compiling and analysing the Spoken British National Corpus 2014. *International Journal of Corpus Linguistics* 22/3: 311–318.
- McKinley, Jim and Heath Rose eds. 2017. *Doing Research in Applied Linguistics: Realities, Dilemmas and Solutions*. London: Routledge.
- Mukherjee, Joybrato. 2009. The grammar of conversation in advanced spoken learner English. In Karin Aijmer ed. *Corpora and Language Teaching*. Amsterdam: John Benjamins, 203–230.
- O'Keeffe, Anne and Svenja Adolphs. 2008. Response tokens in British and Irish discourse. In Klaus P. Schneider and Anne Barron eds. *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins, 69–98.

- O’Keeffe, Anne and Michael McCarthy eds. 2022. *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- Myers, Greg. 2010. Stance-taking and public discussion in blogs. *Critical Discourse Studies* 7/4: 263–275.
- Paquot, Magali and Stefan Th. Gries eds. 2021. *A Practical Handbook of Corpus Linguistics*. New York: Springer International Publishing.
- Paquot, Magali and Luke Plonsky. 2017. Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research* 3/1: 61–94.
- Pérez-Paredes, Pascual. 2019. The pedagogic advantage of teenage corpora for secondary school learners. In Peter Crosthwaite ed. *Data Driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners*. London: Routledge, 67–87.
- Pérez-Paredes, Pascual and Geraldine Mark. 2022. What can corpora tell us about language learning? In Anne O’Keeffe and Michael McCarthy eds, 312–327.
- Pérez-Paredes, Pascual and María Sánchez-Torne. 2019. The linguistic dimension of L2 interviews: A multidimensional analysis of native speaker language. *Focus on ELT Journal* 1/1: 4-26.
- Stubbs, Michael. 2007. On texts, corpora and models of language. In Michael Hoey, Michaela Malhberg, Michael Stubbs and Wolfgang Teubert eds. *Text, Discourse and Corpora: Theory and Analysis*. London: Bloomsbury, 127–161.
- Tracy-Ventura, Nicole and Florence Myles. 2015. The importance of task variability in the design of learner corpora for SLA research. *International Journal of Learner Corpus Research* 1/1: 58–95.
- Tracy-Ventura, Nicole and Magali Paquot eds. 2021. *The Routledge Handbook of Second Language Acquisition and Corpora*. London: Routledge.
- Tracy-Ventura, Nicole, Magali Paquot and Florence Myles. 2021. The future of corpora in SLA. In Nicole Tracy-Ventura and Magali Paquot eds, 409–424.
- Tyler, Andrea and Lourdes Ortega. 2018. Usage-inspired L2 instruction: An emergent, researched pedagogy. In Andrea Tyler, Lourdes Ortega, Mariko Uno and Hae In Park eds. *Usage-Inspired L2 Instruction: Researched Pedagogy*. Amsterdam: John Benjamins, 3–26.
- Waters, Cathleen. 2013. Transatlantic variation in English adverb placement. *Language Variation and Change* 25/2: 179–200.

Corresponding author

Pascual Pérez-Paredes
 University of Murcia
 Department of English Philology
 Plaza de la Universidad s/n
 30001 Murcia
 Spain
 E-mail: pascualf@um.es

received: August 2023

accepted: June 2024