STEVE WALSH

USING CORPORA
FOR ENGLISH LANGUAGE TEACHER EDUCATION (ELTE)

INTRODUCTION

This article considers how corpora can be used in the training and professional development of English language teachers. In this context, corpora might include spoken interactions between trainers and trainees, small group teaching sessions, reflective commentaries, or written texts which use narrative modes of enquiry. The main advantage of using corpora in teacher education is that they provide an excellent point of access to complex phenomena related to teaching and learning and have the potential to trace a teacher's development over time. Video-recordings, for example, collected and constructed as a corpus, offer a very useful springboard for reflection, discussion and the co-construction of meaning. Given the complexity of teaching, corpus-based approaches allow users to gain closer, fine-grained understandings of professional practice, often with a view to making improvements.

The paper will present and exemplify the use of corpus-based approaches in ELTE, focusing on two corpora: LIBEL CASE (The Limerick and Belfast Corpus of Academic Spoken English) and SETTVEO (Self Evaluation of Teacher Talk and Video Enhanced Observation, Walsh, 2019). In discussing examples of corpus-based approaches to English Language Teacher Education, the aim is to demonstrate how corpora can be utilised to promote more dialogic reflection, encourage the sharing of best practice and establish evidence-based, data-led reflective practice procedures (c.f. Mann & Walsh, 2017; Walsh & Mann, 2015, 2019).

STEVE WALSH, Professor of Applied Linguistics at Newcastle University; e-mail: steve.walsh@ncl.ac.uk; ORCID: https://orcid.org/0000-0003-4524-7827.

Finally, I consider how future research in ELTE might adopt corpus-based approaches by harnessing recent developments in, for example, the use of technology such as video and social media.

1. CURRENT RESEARCH IN ELTE

A useful starting point for this chapter in order to gain insights into the current state of play in ELTE research, is to briefly review the main themes emerging from *The Routledge Handbook of English Language Teacher Education* (Walsh & Mann, 2019). Production of any handbook is always going to be an onerous and challenging task and this project was no exception; from inception to publication took over three years, culminating in an edited collection of 39 chapters and involving more than 60 contributors. In the words of the editors:

> The main aim was to provide a comprehensive overview of English Language Teacher Education (ELTE), drawing on perspectives from the related fields of Applied Linguistics, Education and TESOL. The handbook seeks to identify, discuss and raise awareness of key issues and practices in education, training and in teacher development. The 39 papers assembled here offer insights into the range and detail of current practices in ELTE and associated research. (Walsh & Mann, 2019, p. 5)

If we accept that one of the goals of any handbook is to offer a "snapshot" of a particular field at a given moment in time, which themes feature in the handbook and to what extent might they be investigated through corpus-based approaches? A number of key themes can be identified in contemporary studies of ELTE. Owing to limitations of space, I will focus here on those themes which have informed—and which continue to inform—the ways in which ELTE "gets done" in various contexts around the world. The discussion which follows is an attempt to capture the main practices and processes which contribute to our understandings of the field, while offering insights into the value of corpus-based approaches.

1.1 UNDERSTANDING CONTEXT

Much of the current research in ELTE is located in a sociocultural theoretical (SCT) framework, a useful perspective for developing understandings of context. In the handbook, Golombek and Johnson offer a Vygotskyan SCT (VSCT) perspective on ELTE which provides an overview of key principles and concepts to demonstrate how VSCT might be used to inform the practices and interactions of

ELTE programmes in systematic ways. Other researchers, for example van Compernolle (2014), use corpus-based approaches to exemplify key features of SCT in classroom interaction. And while most approaches to researching ELTE look at context from the perspective of pre-, in-service and CPD (continuing professional development), Morton looks across all three contexts at content-based approaches to language teaching, focusing on training implications for teachers who teach academic content through English in contexts where English is not the first language of the majority of the students.

1.2 DEVELOPING SOPHISTICATED UNDERSTANDINGS OF LANGUAGE

Clearly, the 'L' in ELTE merits serious attention; as educators we are concerned to ensure that English teachers have a sophisticated and current understanding of "language". Increasingly, this means far more than a knowledge of the basic skills and systems of English and requires a detailed understanding of the ways in which language is used to communicate. Strong advocates of this perspective are Fiona Farr and Anne O'Keeffe, whose concern is to provide teachers and teacher educators with appropriate tools with which to analyse language. They propose the use of corpora and corpus linguistics as a means of facilitating teacher development in terms of content, pedagogy, technology, and research. Based on many years of research, the authors argue that corpora continue to play a minor part in much teacher education but are a resource with much potential.

Continuing the focus on the "language" of ELTE, Sert shifts the focus to the study of classroom interaction. His chapter looks at the close relationship between classroom interaction and teacher development. Using research evidence from both initial and in-service teacher education, the chapter makes a case for a technology-enhanced, reflective, and micro-analytic teacher education framework.

A key aspect of ELTE, and, indeed, of language teaching more generally, is feedback. Teachers and learners—or educators and student teachers—depend very heavily on the feedback we give and receive. Feedback is almost certainly the most important element in both language learning and professional development; as such, it needs to be understood more fully. This is taken up by Jo-Ann Delaney who considers the role of assessment of teaching and giving feedback in promoting teacher development. Her chapter explores issues around assessment of teaching and providing feedback and shows how these processes can play an important role in teachers' learning.

Continuing in the same vein, Copland and Donaghue focus specifically on the nature of feedback in the post-observation conference. They discuss a number

of key issues inherent in such post observation feedback. As well as focusing on the linguistic characteristics, purposes and value of the feedback conference, they also suggest ways of helping participants better understand feedback.

1.3 REFLECTIVE PRACTICE

One of the key strands of contemporary approaches and research in ELTE is Reflective Practice (RP). In the third chapter, Farrell discusses how educators can encourage learner teachers to become reflective practitioners. The chapter outlines how educators can make use of technology, critical friendships, team-teaching, peer coaching, dialoguing, service learning, writing, action research and analysis of critical incidents to encourage learner teachers to engage in reflective practice. The theme of RP is to be found in most sections of the handbook as a key element of ELTE which the editors have discussed in other publications (see, e.g., Mann & Walsh, 2013; Walsh & Mann, 2015).

1.4 INTEGRATING TECHNOLOGY

Another key strand running through the handbook is the use of technology, something taken up by Lightfoot, who focuses on teachers and teacher educators using technology to facilitate professional development and explores the various benefits, issues and challenges associated with the use of ICT.

Other aspects of technology are dealt with in the handbook. Stannard and Sallı, for example, focus on screen capture technology (SCT), a tool that is increasingly being widely used in teacher education and language training. The chapter provides examples of how SCT is used in ELTE through a range of case studies. Huettner presents an overview of the way video resources can be integrated into ELTE. Video is increasingly being used to provide a record of teaching practice for feedback and reflection. It is also used in other important ways (e.g. providing input on alternative practices). Huettner provides guidance on how video can be used effectively as a resource for the development of "professional vision". The second case study presented below looks closely at how video-based corpora can be used in ELTE.

Gulzar and Barrett detail the affordances that ePortfolios can provide as a learning and assessment tool in ELTE, providing a guide to the process of working effectively with ePortfolios and showing that technology can help support reflective processes in teacher education. It is not difficult to imagine how corpus-based approaches could be adopted to study ePortfolios.

While there is not the space here to review all themes which are to be found in the handbook, the intention is to give readers a flavour of contemporary issues in ELTE. It is clear, as many of the contributors to the handbook emphasize, that the use of corpora and corpus-based approaches have done much to enhance our understandings of the field. In the next section, I present specific examples of how such approaches have been used in previous studies.

2. CORPUS-BASED RESEARCH IN ELTE

2.1 CLCA

In the first example, published originally in 2012 (Walsh et al., 2012), I highlight a particular methodological approach which was used to provide university teachers with an enhanced understanding of the complexities of small group teaching interactions. The aim of this study was to try out and evaluate a methodology which combined corpus linguistics (CL) with conversation analysis (CA). The methodology follows an iterative process: from CL to CA, back to CL and so on. This approach—henceforth CLCA—provides powerful insights into the ways in which interactants establish understandings in educational settings and, in particular, highlights the inter-dependency of words, utterances and text in the co-construction of meaning. The aim was to consider how corpus linguistics (CL) and conversation analysis (CA) could be used together to provide enhanced understandings of spoken interaction and its relationship with learning.

The 2012 study was based on data from the Limerick Belfast Corpus of Academic Spoken English (LI-BEL CASE), comprising around 1 million words of recorded lectures, small group seminars and tutorials, laboratories and presentations. These data were collected in two universities on the island of Ireland: Limerick and Belfast, across common disciplinary sites within the participating universities: Arts and Humanities, Social Sciences, Science, Engineering and Informatics, and Business. From the main corpus, a sub-corpus of 50,000 words was created by identifying all the instances of small group teaching sessions (SGT), comprising up to 25 students.

While CA and CL have both been used on their own to study spoken encounters, each has its limitations. CL, for example, largely ignores context and focuses on large scale analysis, whereas CA offers detailed descriptions but is unable to generalise to larger contexts. Using a combined CL and CA approach (CLCA), cumulatively gives a more "up-close" description of spoken interactions in an ed-

ucational setting than that offered by using either one on its own. From the analysis, powerful insights are gained into the ways in which interactants establish understandings and observe how words, utterances and text combine in the co-construction of meaning.

The methodology used in this study can be described as mixed methods, combining as it does both quantitative and qualitative research methods. As in any mixed methods research, the aim was to compensate the weaknesses of one approach by highlighting the strengths of the other. CL is seen as a *methodological tool* used to interrogate a corpus of small group interactions recorded in higher education. Using CL as a tool allows us to automatically search a large dataset, something which would have been impractical manually. However, while CL allows us to count frequencies and find key words in micro-seconds, thus revealing patterns that we could not otherwise find, it does not allow us to explain the dynamics of these interactions. This is where a closer look at the data, through the use of CA, prevails. The first layer of analysis, therefore, enabled the researchers to scope out and quantify recurring linguistic features.

The second layer of analysis (using CA) draws upon these contextual patterns in the quantitative analysis and investigates them more closely. For example, in the corpus exploration, there were interesting findings around the frequency and use of certain discourse markers, which clustered around specific contexts. This led to a closer CA led investigation which, in turn, produced interesting findings above the level of turn and in relation to specific interactional features. The process was non-linear in that CL operations were sometimes used within the CA layer of analysis to quantify CA insights; the analysis progresses in an iterative manner: from CL to CA, back to CL and so on. There is an interdependence between the two modes of analysis.

A brief summary of the analytic procedure is now presented. The broad goal of the CL analysis was to first identify the high frequency and key lexical features—both single words and multi-word units (MWUs)—of this type of interaction, in comparison with a reference corpus. These features were salient since they highlighted particular lexical features which were quantitatively distinctive about this type of interaction using frequency lists and keywords analyses. The "keyness" of these items was also explored qualitatively by going back to the data to find out what it was they were "doing" in the interactions. The CL analysis thus provided a broad picture of how participants were going about "doing SGT" sessions. A CA methodology was then used to explore sequences of interaction in greater depth where there was a clustering of these statistically significant, or salient, features.

In addition to frequency counts, the concordancing function of *Wordsmith Tools* (Scott, 2004) was used to offer a more qualitative perspective on the data. This operation revealed differences in the functioning of some key words. So for example, *if* was found to have three main functions:

- pedagogic illustration of 'general truths/facts' *if John Kerry takes Texas, …he takes every vote…*
- projecting, meaning 'when you find yourself in this situation' *if you are on TP and you have a class that…*
- demonstrating, *if you click the mouse and then click…* also showed that the relatively high frequency of *need* in comparison to the BNC is relate to the speech act of giving instructions (*what I need you to do, you need to* etc.) (Walsh et al, 2012).

The final layer of analysis under the CL methodology was to identify patterns by focusing more closely on MWUs. Two- to six-word multi-word units were generated with a cut off frequency of four occurrences. These were then examined through concordance searches, producing 128 items which were seen as most salient to the SGT context. At this stage, the analysis has produced word frequency lists, key word lists, concordances and a list of MWUs. These were then classified according to their broad functions. In order to gain a deeper understanding of spoken interaction in this context, the next stage was to determine how particular salient features actually operated in speakers' turns and in longer sequences of interaction. By looking at micro-contexts within a CA framework, it was possible to describe their interactional and pedagogical features at and above the level of turn. The dialectic between CA and CL thus provided a better understanding of why certain items were clustering at certain points.

In any classroom discourse context, pedagogic goals and the language used to achieve them are inextricably linked (Seedhouse, 2004; Walsh, 2006). This is equally true in SGT interaction, where the tutor's pedagogical goals of the moment are largely responsible for the shape of the ensuing interaction. In the dataset, four such speech exchange systems were identified, each with its own distinctive shape and fingerprint (Drew & Heritage, 1992). Each is now briefly described (for a fuller description the reader is referred to Walsh et al, 2012).

2.2 PROCEDURAL TALK

These stretches of interaction often involve very long turns by the tutor, with minimal verbal contributions from the students. For example, the tutor may perform the role of both questioner and answerer as he/she talks through a procedure.

Typically, procedural talk is used to manage learning, organise tasks and activities, move a session to the next stage, etc. In the CL analysis, procedural talk is indexed by the use of MWUs which have the function of indicating actions which the tutor wants the students to do, such as the four word MWU *I want you to* or, less frequently, the six-word MWU *what I want you to do*.

2.3 DIDACTIC TALK

In this speech exchange system, turn-taking is tightly controlled by tutors, with next turn allocation firmly their hands and questions addressed to individual participants. Interactions closely resemble classic descriptions of classroom discourse, with evidence of IRF exchanges, heavy use of display questions, student nomination, short utterances from students.

2.4 EMPATHIC TALK

Empathic talk is characterised by shared space for learning (Walsh & Li, 2013), where tutors and students assume more symmetrical roles, and in which students typically share their experiences, ideas, opinions through "tellings". The tutor accepts and builds on these accounts, converting them into pedagogical material in the form of reflective statements about appropriate behaviour, roles and identities in the professional practice of the discipline. Agreement to assessments is favoured (there is a lack of dispreferred responses) and there is frequent use of interpersonal discourse markers to provide supportive responses to the speaker (*yeah*) and to mark/monitor shared knowledge (*you know*; *you see*, see Carter & McCarthy, 2006; Andersen, 2001).

2.5 ARGUMENTATIONAL TALK

The fourth type of exchange structure identified in the corpus was one in which there was also a shared cognitive space, but this time it is disputed rather than used to build empathy. Argumentational talk was indexed in the corpus data by the appearance of *but* as a "tying" element at the beginning of turns (Tao, 2003), sometimes preceded by an agreeing or acknowledging token such as *yeah* or *okay*.

To recap, the reader is reminded that the aim of the 2012 study was to evaluate the usefulness of combining CL and CA to provide enhanced descriptions of spoken interaction in a small group teaching higher education context. From the data and subsequent CLCA analysis, four speech-exchange systems were identified

according to their distinctive interactional and linguistic features and their accompanying pedagogic goals, each with distinctive interactional, linguistic and pedagogic features. The four speech exchange systems we identified are robust throughout the data. That is, at any point one or other will be operating, whether for long spates of interaction or for shorter bursts.

From the perspective of English language teacher education, the findings from this study have considerable value in promoting understandings of how central interaction is to learning in SGT settings. From the data, it is possible to make comparisons both within and across these interactional contexts; for example, when we compare didactic and empathic talk, very different profiles or "fingerprints" (Drew & Heritage, 1992) emerge. The former is characterised by short learner turns, tightly controlled turn-taking, evidence of IRF exchange structures, extensive use of the MWUs *tell me* and *can you tell me* and the main pedagogic function of eliciting. The main focus of empathic talk, on the other hand, is "show and tell": the tutor's pedagogic goal is to promote debate and discussion and create a safe environment for that to take place.

By combining CL and CA, it is possible to provide fine-grained descriptions of classroom interaction, descriptions which are of enormous value in the preparation and professional development of teachers and tutors. In particular, when combined with reflection and dialogue, there are many opportunities to develop classroom interactional competence (CIC, Walsh, 2013), enabling educators to identify key moments in the interaction to create opportunities for learning.

Although there have been many attempts to characterise spoken interaction in educational settings by focusing on micro-contexts (see, for example Seedhouse, 2004; Walsh, 2006), none, to my knowledge, offer the same level of detail as the present study. From the CLCA analysis, it is possible to provide detailed descriptions of the interaction from three perspectives: linguistic (portraying the use of high frequency items, keywords, MWUs, discourse markers, question forms and so on), interactional (focusing on turn-taking and turn design, sequential organization, etc.) and pedagogic (looking at specific pedagogic functions at a given moment to include eliciting, explaining, instructing and so on).

The use of either framework alone, CL or CA, would have doubtless provided some interesting findings: CL would have provided interesting lists of high frequency items which could have been explained functionally; but it would not have offered the same depth of analysis as that offered by the use of CA. Similarly, by using CA on its own, the four main speech exchange systems could be identified, but without the support offered by the fact that the words and patterns they contain were actually high frequency items (that is, key words, high fre-

quency words and multi-word units). It is, then, fair to say that the CLCA methodology has much to offer and could easily be adapted to a range of ELTE contexts.

2.5 SETTVEO

The discussion now turns to the second example of a corpus-based approach used in ELTE. The original study (Walsh, 2019) reports findings from a British Council ELTRA-funded research project which looked at the use of technology-enhanced learning in a CPD (continuing professional development) context. The aim of the study was to provide English language teachers with appropriate tools and procedures to enable them to reflect on and improve their practice through the creation and use of an app: SETTVEO.

This app extended previous work, using the SETT (Self-Evaluation of Teacher Talk) framework (Walsh, 2006, 2011, 2013, 2017) and the VEO (Video Enhanced Observation) app (Miller, 2015). Through the use of SETTVEO, and subsequent collective dialogue and reflection, an online community of practice was established, enabling participants to share and comment on examples of English language teaching around the world. This corpus of reflective and interactional data, was then used by both participants and researchers to establish and evaluate a more dialogic, collaborative approach to RP and to help teachers develop their classroom interactional competence (CIC, Walsh, 2013). Findings suggest that the use of self-observation, with data and accompanied by some kind of dialogue, can promote up-close and detailed understandings of teaching and learning.

A brief description of the app and its uses is now provided. (Readers should visit veogroup.co.uk and see Walsh, 2019 for further details.) The VEO app was developed by teacher educators and allows users to record and tag videos which can be uploaded and saved into a portal. This procedure may be carried out synchronously or asynchronously. Teachers can then build online communities in order to share and reflect on their videos to improve the quality of their teaching in a sustainable manner. By enabling the live-tagging of video, the VEO app generates both quantitative and qualitative data which, taken together, provide a profile of a teacher's practice at a given moment in time. The tags allow the user to jump to the exact instance within the video, presenting a rich view of action, interaction and context that can be shared for further analysis and evaluation. This tagging functionality allows for systematic data collection over time, supported by illuminating video evidence that can be interpreted and analysed by multiple practitioners and researchers. The advantages that VEO brings to analysing complex situations

make it highly appropriate for studying interaction, where multiple perspectives are possible and where relevant frameworks can clarify and enhance its understanding.

The SETT framework has been used in a range of educational settings since its publication in 2006 and further development in 2011 and 2013. The framework comprises four classroom micro-contexts (called *modes*) and 14 interactional features (called *interactures*). Classroom discourse is portrayed as a series of complex and interrelated microcontexts or modes, where meanings are co-constructed by teachers and learners and where learning occurs through the ensuing talk of teachers and learners (Walsh, 2013). The key to developing good practice is for teachers to acquire detailed profiles of the interactions that take place in their classes as a means of understanding how learning opportunities are created and how "space for learning" (Walsh & Li, 2012) can be opened up, and to create the kind of dialogic, engaged learning environments which have been advocated in UK primary and secondary classrooms for more than ten years (see, e.g., Mercer, 2009; Alexander, 2008).

In the current study, SETT and VEO were combined to help teachers profile their interactions, improve their classroom interactional competence (CIC) and enhance learning and learning opportunity. By sharing their reflections in an international online community of practice (CoP), it was possible to develop a global network of reflections, with teachers identifying and talking about common problems and issues in a range of English language contexts. Online communities were formed between teachers who shared videos of their teaching which they then evaluated by using the SETTVEO app.

The present study involved 24 university English language teachers in four countries (Chile, Turkey, Spain, Thailand). The project was organised over three phases. In phase 1, baseline data were collected, including a short video-recorded lesson segment of each teacher, a short reflective commentary on that segment, and an online interview with the research team. In the second phase of the study, participating teachers each made four "snapshot" recordings of their teaching (10 to 15 minutes per recording). Online training in the use of SETTVEO was provided to all participants. Each recording was then reviewed and evaluated using SETTVEO. The recordings and reflective commentaries were uploaded to the VEO platform as an online community of practice. They were then able to comment, question, reflect on and share both the snapshot recordings of their teaching (96 in all) and the accompanying reflections. In the final phase of the study, participating teachers took part in online focus groups and individual interviews to evaluate the extent to which their RPs had changed and to consider any changes in CIC.

From the corpus—the online community of practice—a number of themes emerged, summarised briefly here (originally published in Walsh, 2019).

2.5.1 Theme 1: Developing CIC

One of the key questions posed in this study was if and by how much teachers could improve their CIC. Several teachers made comments on this in their reflections. Teacher A, below, for example, talks about her elicitation strategies and gives reasons for her particular use of display or referential questions, two of the interacture tags used in the SETTVEO app. Of interest in this extract is the ability of the teacher to use an appropriate metalanguage to describe her practice. She is also able to give reasons for her choice of language and both justify her interactional decisions and explain how level is an important determiner of these practices.

Extract 1

I asked mostly display questions to help them do brainstorming about the topics and to make most of the students be able to speak about the topics. At the beginning I needed to ask some referential questions (00.20) to refer to the exercises we did in our former lessons. After watching myself in the video I saw that I repeated the same phrases several times, but I think this helped them speak better and self-confidently because they were elementary-level students and needed to hear too many repetitions and examples. Firstly, I used content feedback to emphasise how they would find relevant supporting ideas for the topics we discussed, how they would agree/disagree with each other, how they would organise their ideas and list them. Secondly, I mostly preferred form focused feedback because they were in need of hearing correct forms and learning how to make correct sentences.

An important feature of the SETTVEO app is the tagging function, which allows users to "tag" (i.e. mark) specific features of their interactions and then review these features later. The software then prepares a statistical breakdown showing how features are used, which ones occur most frequently and how the use of certain interactional features influenced the interaction. In Extract 2 below, Teacher C is talking about her use of teacher echo, a feature which many teachers comment on. Here, Teacher C focuses on her excessive use of discourse markers (sometimes referred to as transition markers). In fact, these tokens perform a very important function in classroom discourse, acting almost like punctuation marks and helping students stay focused and avoid becoming lost in the interaction (see, e.g., Breen, 1998). Teacher C then goes on to look at the statistical breakdown of specific features, making a valid comparison between display and referential questions—the former dominate most classroom talk, while teachers often miss opportunities to ask genuine, or referential, questions. Again, she demonstrates a

high level of CIC through her ability to use an appropriate metalanguage and justify her actions, even being quite critical of her decisions at times.

Extract 2

Looking at the tagging session report, I observed that the teacher echo was excessive to my standards. It made me realise that I should make an effort to reduce it, because I found it annoying to watch myself saying so many 'OK's and 'all right's. The amount of display questions on the report was high at 40 per cent, compared to the referential questions at 15 per cent. The amount of content feedback and seeking clarification was quite similar at approximately 15 per cent. The lowest rate was form-focused feedback with only about three per cent. Although I was a bit disappointed at myself for making the tagging session a bit long because of the wait times for the reading and watching the videos, the extended wait time was quite low in the report, at only five per cent.

By producing a profile of a teacher's classes, SETTVEO offers teachers an opportunity to reflect on the kind of teacher they are, and, more importantly, the kind of teacher they would like to become. In Extract 3 below, Teacher E demonstrates a high degree of student-centredness, explaining and justifying her interactional decisions with reference to her students. The app allows fine-grained, nuanced observations to be made which might otherwise go unnoticed.

Extract 3

The recorded part of my lesson was focused on comprehension of their reading material, which students completed right before the recording. Therefore, mostly referential questions were asked during this part of the recording. Right after, students answered related questions in pairs, which was a bit long. I marked it as extended time, but I allowed it because it is very beneficial for my students. Usually, during this part of the lesson, I like to walk around the room and provide them with form-focused feedback and error correction. And, the last part of the recording, students focused on exercises from the book, where the key vocabulary was used in practice. This part as well was marked as extended time.

2.5.2 Theme 2: Changes to practice and self-awareness

One of the goals of the study was to note changes to practice. The extracts below (4 and 5) exemplify some of the changes which teachers reported. The message in Extract 4 could not be clearer: the main goal of ELE is to teach students to *speak* the language; again, the emphasis is very much on the students rather than on teacher performance. One of the main advantages of SETTVEO is that teachers, in focusing on their own interactional competence, cannot ignore the interactions and language used by their students. Many of the comments in the data referred to the actions or engagement of learners; a further stage would be to try to

explain these in relation to what was said by the teacher since one aspect of CIC is to understand how teacher and learner interaction are inextricably linked.

Extract 4

Speaking and helping them speak is the best way of creating self-sufficiency. The more they speak, the better they feel and learn.

In Extract 5 below, Teacher C demonstrates how their awareness developed through participation in the study, highlighting the importance of wait time and the need to give students space by not "stealing" their speaking time. She also notes the importance of allowing students to express their opinions more and take control of topic (cf. Slimani's (1989) work on "topicalisation": learners are more likely to learn effectively when they have control of the topic).

Extract 5

If I were to change anything in this session, what I would do would be to talk less. I observed that I was haste (sic), and I stole from my students' speaking time. I also asked them if they were ready during the extended wait time, more than I should have. Towards the end, my students were more active, and they engaged more, because they had a chance to give their opinion on the subject.

2.5.3 Theme 3: The advantages of video

It is clear when reading the next extract that, for some of the participants at least, the use of video in teacher learning is of great benefit. Previous studies (see, for example, Mann 2018) have highlighted the use of video in ELTE and pointed to this medium as an important and progressive tool in future CPD. Teacher B, with 16 years' experience, emphasises the value of self-observation as a means of understanding classroom dynamics and understanding student feelings and attitudes. Of importance in this observation is the fact that she seems to suggest that her own strengths and weaknesses can be gleaned by focusing on her learners, an acute and mature observation. Rather than "blaming the learner", this teacher takes responsibility for her professional practice, acknowledging that while her understanding of teaching and learning can be developed through self-observation, a focus on her students will clearly highlight her strengths and weaknesses as a teacher.

Extract 6

I had an opportunity to reflect on my teaching as a teacher having 16 years' experience considering my classroom video as a part of the SETTVEO project. During

this period, I found a chance to make a self-observation which enabled me to be deeply aware of my classroom dynamics, students' attitudes and feelings during the class and as a matter of course my strengths and weaknesses.

2.5.4 Theme 4: Project evaluation

The final section of analysis offers a brief evaluation of the study. In Extract 7, Teacher H comments on some of the issues raised by participants when viewing themselves on video for the first time. Of interest here is the value of video in affirming and strengthening beliefs about teaching. And this is a valid and important point: through self-observation, we are not always seeking to change or develop; rather, we are looking to be reassured that our practice is appropriate, that we are decent teachers and that our students actually learn something. To that end, the SETTVEO project achieved its goals.

Extract 7

SETTVEO has been an opportunity for participant instructors to raise awareness of what they have employed in their classes. The participants from time to time informed me about the issues which led to raise their awareness of their own experience inside a real classroom context. For example, one of them stated that she had not heard her voice like this in the classroom and it was so strange for her. Also, another participant told me that she had not thought she repeated the same information that much inside the classroom. As a conclusion, SETTVEO is both a challenge and a change for all of us. The project is also fun and a learning process for us and it contributes to our understanding of the teaching–learning process with genuine classroom-based data.

In this section, I have presented a second example of a corpus-based approach to teacher education. The present study was an attempt to facilitate reflection by providing teachers with something to reflect on (CIC) and something to reflect with (a tool in the form of SETTVEO). From the relatively limited evidence presented here, it is apparent that video has much to offer in any teacher development context. Recent studies (see, for example, Mann, 2018) confirm the value of video in mediating understandings of teaching and in unpacking the complexities of that process. One of the most important aspects of the use of video in teacher education is that it quickly, easily and inexpensively provides evidence on which to reflect.

The SETTVEO app enabled short recordings to be made, shared in a CoP and discussed. The potential from what was essentially a small-scale study is enormous; it would not be difficult, for example, to extend the present project to something much bigger, leading to an international corpus of professional practice comprising video recordings and reflections from every corner of the globe. Not only would such a resource promote greater understandings of teaching and

learning, it would result in closer and deeper understandings of context, surely the most important element in language education.

CONCLUSION

This article set out to explore and exemplify some of the advantages and main reasons for using corpora and corpus based approaches in ELTE. The main advantages of this approach can be summarised as follows. Corpus-based approaches provide real and tangible evidence to facilitate reflection and professional development. The use of "snapshot" video-recordings (short, 7–10-minute extracts) enables a more dialogic approach to reflection and might involve two or more teachers. The same dataset can then be used to compare progress and professional development over time, highlight key issues or "puzzles" in a specific context and allow teachers to both raise problems and discuss solutions.

A second advantage of the approach, presented in the CLCA example, is that corpora can be interrogated in a number of ways, thereby providing more robust and richer datasets. Using a combined CLCA methodology is only one way in which different perspective on the same corpus might be offered. Other researchers have successfully combined CL with discourse analysis and critical discourse analysis (refs). And even using a CL approach alone can offer both quantitative and qualitative interpretations of the same dataset. The point I am making is that we need up-close, fine-grained analyses of corpora to see what is really going on and to unpack the complexities of teaching and teacher development.

A third advantage of corpus-based approaches is that they lend themselves very well to technology integration. I have already highlighted the relevance of video-based corpora which provide reliable evidence and stimulate debate and discussion. Other online formats which are equally valuable include the use of e-portfolios, blogs and so on, plus the many opportunities offered to share space through social media such as WhatsApp, WeChat and so on. Sandra Morales, for example, working in teacher education in Chile was able to use online blogs with a group of teachers to facilitate technology integration (Morales, 2017).

Turning now to a consideration of the future of ELTE and how corpus-based approaches might play a role, I would like to open the discussion with a focus on technology. Other researchers, for example Brona Murphy (Mann & Walsh, 2017) highlight the relevance of technology and the use of corpora in initial teacher education. In particular, she stresses the value of technology in terms of its value in engaging student-teachers and promoting interactivity and discussion.

Other research has demonstrated that reflections are more likely to be genuine when shared via online formats (ref). As Murphy herself says (Mann & Walsh, 2017, p. 261):

> Working with TESOL postgraduate pre-service teachers, I find they seem more enthused about reflecting when technology is involved. They tend to enjoy blogging and producing e-portfolios which can be shared and used in an interactive way. I find that this also appears to produce more genuine reflective engagement than more solitary work where students, in the past, have said that they don't really mind what they write as they're usually the only ones to see it. Future directions for RP would benefit from allowing teachers to experiment with reflection online and in a more interactive way drawing on insights from social media and using photos, quotes and other tools to support collaboration and sharing around reflection.

There are many examples of both pre- and in-service ELTE where teachers' professional development is tracked over time; corpora are the ideal means of recording these developments and allowing both tutors and their students to comment on and share their thinking. Such longitudinal projects are becoming increasingly important and, again, corpora have much to offer.

Other researchers predict that there will be much more interdisciplinary research in the field of ELTE and, again, corpus-based approaches have a key role to play. Choi and Richards (2018) for example, using a corpus of interdisciplinary discourse, look at the ways in which interdisciplinary researchers learn to "talk the talk" of their subjects with colleagues who stand outside their discipline. Such an approach could easily be adapted to ELTE where, increasingly, English teachers are working with other subject specialists in contexts such as CLIL and EMI. Farr (2017) also predicts that ELTE has much to learn by working with other disciplines, especially in the areas of cognitive psychology and neuroscience to examine how patterns in the brain translate into reflections and changes to professional practice.

In terms of research methodologies which could benefit from corpus-based approaches, Linguistic Ethnography (LE) has much potential, as advocated by Fiona Copland (see, e.g., Copland & Creese, 2015). LE combines linguistic analysis of spoken or written texts with ethnographic approaches, such as observation and interview. A corpus of, for example, post-observation teaching sessions (see, e.g., Harris, 2012) could provide an extremely valuable resource for subsequent analysis involving interviews or focus groups. In her own words, Copland explains:

> [LE provides] a detailed and nuanced exploration of oral reflective practice in post observation feedback conferences (Copland, 2010; 2011; Donaghue, 2016). In this regard, LE provides particular affordances in terms of understanding the contextual fea-

tures of feedback conferences. (Mann & Walsh, 2017, p. 263)

In addition to providing opportunities for the kind of longitudinal work mentioned above, LE has the potential to offer deeper and more comprehensive understandings of an issue, especially when different datasets are used. So, for example, a corpus of student-teachers' reflections could be used to stimulate discussion and further reflection, which could then form a second dataset.

As this paper has hopefully shown, there are many innovative and interactive ways of using corpora in ELTE. There is no one "quick fix" to researching the field, nor is there one "best method". As a profession, we need multiple perspectives on the practice of ELTE, entailing the use of a wide range of procedures and methods. By collecting evidence — in the form of written or spoken corpora — we are, I would suggest, much more likely to capture more accurately the process of teacher education and develop closer understandings of its many complexities. As the opening section of this chapter has demonstrated, the current status of ELTE together with its future development will necessitate better understandings of context, closer integration of technology, finer-grained knowledge of reflective practice and more sophisticated understandings of language. In all these key areas, corpus-based approaches have a major role to play.

REFERENCES

Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.

Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation. A relevance-theoretic approach to the language of adolescents*. John Benjamins.

Carter, R., & McCarthy, M. J. (2006). *Cambridge grammar of English. A comprehensive guide to spoken and written grammar and usage*. Cambridge University Press.

Choi, S., & Richards, K. (2017). *Interdisciplinary discourse: Communicating across disciplines*. Springer.

Copland, F. (2010). Causes of tension in post-observation feedback in pre-service teacher training: An alternative view. *Teaching and Teacher Education*, *26*(3), 466–472.

Copland, F. (2011). Negotiating face in the feedback conference: A linguistic ethnographic approach. *Journal of Pragmatics*, *43*(15), 3832–3843.

Copland, F., & Creese, A. (2015). *Linguistic ethnography*. Sage.

Donaghue, H. (2016). *The construction and negotiation of identity and face in post observation feedback* [Unpublished doctoral thesis]. Aston University.

Drew, P., & Heritage, J. (1992). Analyzing talk at work: An introduction. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3–65). Cambridge University Press.

Johnson, K. E. (2009). *Second language teacher education: A sociocultural perspective*. Routledge.

Mann, S., & Walsh, S. (2013). RP or 'RIP': A critical perspective on reflective practice. *Applied Linguistics Review*, *4*(2), 291–315.

Mann, S., & Walsh, S. (2017). *Reflective practice in English language teaching: Research-based principles and practices*. Routledge.

Mercer, N. (2009). *Words and minds: How we use language to think together*. Routledge.

Miller, P. (2015). *Video-enhanced observation: A new way to develop teacher practice.* Optimus Education.   http://my.optimus-education.com/video-enhanced-observation-new-way-develop-teacher-practice

Morales, S. (2016). *How in-service language teachers become effective users of CALL for online teaching and learning: A case study of their development processes in a transformative online teacher training and development course* [Unpublished PhD thesis]. Newcastle University.

Scott, M. (2008). WordSmith Tools (Version 5). Lexical Analysis Software.

Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Blackwell.

Slimani, A. (1989). The role of topicalization in classroom language learning. *System*, *17*(2), 223–234.

Tao, H. (2003). Turn initiators in spoken English: A corpus-based approach to interaction and grammar. In P. Leistyna & C. F. Meyer (Eds.), *Corpus analysis: Language structure and language use* (pp. 187–207). Rodopi.

van Compernolle, R. A. (2014). *Sociocultural theory and L2 instructional pragmatics*. Multilingual Matters.

Walsh, S. (2006). *Investigating classroom discourse*. Routledge.

Walsh, S. (2011). *Exploring classroom discourse: Language in action*. Routledge.

Walsh, S. (2013). *Classroom discourse and teacher development*. Edinburgh University Press.

Walsh, S. (2019). *SETTVEO: Evidence-based reflection and teacher development.* British Council. https://www.teachingenglish.org.uk/publications/case-studies-insights-and-research/settveo-evidence-based-reflection-and-teacher

Walsh, S., Morton, T., & O'Keeffe, A. (2011). Analysing university spoken interaction: A CL/CA approach. *International Journal of Corpus Linguistics*, *16*(3), 325–344.

Walsh, S., & Mann, S. (2015). Doing reflective practice: A data-led way forward. *ELT Journal*, *69*(4), 351–362.

Walsh, S., & Mann, S. (2019). *The Routledge handbook of English language teacher education*. Taylor and Francis.

Walsh, S., & Li L. (2012). Conversations as space for learning. *International Journal of Applied Linguistics*, *23*(2), 247–266.

USING CORPORA
FOR ENGLISH LANGUAGE TEACHER EDUCATION (ELTE)

S u m m a r y

The article seeks to explore the use of corpora in the training and professional development of English language teachers. It will be argued that in ELTE (English Language Teacher Education), the use of corpus-based approaches offers a wide range of possibilities which enable and promote

reflection, encourage the sharing of best practice and foster data-driven practice procedures. The analytical core of the paper focuses on two examples of integrating corpora into ELTE—CLCA, an amalgamation of corpus linguistics and conversational analysis, and SETTVEO, a combination of the SETT (Self-Evaluation of Teacher Talk) framework and the VEO (Video Enhanced Observation) app. The former provides insights into enhanced understandings of spoken interaction and its relationship with learning. The latter aids teachers in developing their classroom interactional competence by collecting reflective and interactional data so as to establish and evaluate a dialogic and collaborative approach to reflective practice.

Keywords: corpora; English Language Teacher Education; conversational analysis; reflective practice; classroom interactional competence

ZASTOSOWANIE KORPUSÓW JĘZYKOWYCH
W SZKOLENIU ZAWODOWYM NAUCZYCIELI JĘZYKA ANGIELSKIEGO

S t r e s z c z e n i e

Niniejszy artykuł koncentruje się na użyciu korpusów językowych w szkoleniu i rozwoju zawodowym nauczycieli języka angielskiego. Autor zakłada, że wykorzystanie podejść korpusowych w rozwoju nauczyciela oferuje szeroki wachlarz możliwości, które pozwalają na refleksję, wspomagają dzielenie się dobrymi praktykami i wdrażanie procedur opartych na tych danych. Analityczna część artykułu koncentruje się na dwóch przykładach integracji korpusów z edukacją nauczyciela – CLCA, połączenia językoznawstwa korpusowego i analizy konwersacyjnej, oraz SETTVEO, czyli kombinacji modelu SETT (autoewaluacja języka nauczyciela) z aplikacją VEO (obserwacja wspierana materiałem wideo). Pierwsze rozwiązanie pozwala na lepsze zrozumienie interakcji i jej związku z uczeniem się, drugie zaś wspomaga nauczycieli w rozwijaniu kompetencji interakcyjnej w klasie poprzez gromadzenie danych o interakcji umożliwiających autorefleksję w celu stworzenia i oceny dialogicznego i kolaboracyjnego podejścia do pracy opartej na refleksji.

Słowa kluczowe: korpusy; szkolenia zawodowe nauczycieli języka angielskiego; analiza pod kątem konwersacji; praca oparta na refleksji; kompetencja interakcyjna