

# Using Corpus Linguistics in Materials Development and Teacher Education

Niall Curry<sup>a</sup> and Geraldine Mark<sup>b</sup>

## Abstract

*In published ELT materials, norms relating to written language often dominate, at the expense of spoken language. This dominance of writing typically arises as the features characteristic of everyday spoken language do not reflect the rather neat syllabi expected by teachers and students, globally. Furthermore, there are deeply held opinions relating to the usage and acceptability of features of spoken language. As a result, features of spoken language are often not represented in mainstream materials, despite a perennial request from learners for more 'conversation'. One potential means to meet such a request is by employing corpus linguistics and spoken language research to inform ELT materials development and teacher education. Responding to this issue, this paper draws on findings from a corpus-based conversation analysis of spontaneous spoken British English to conduct workshops on materials development. The goal of these workshops is to gain a comprehensive perspective on the affordances of corpus linguistics for ELT materials development from core stakeholders: teachers. Through a grounded theory-based thematic analysis of lesson plans, recorded discussions, and survey responses, the findings of this study demonstrate that teachers find value in corpus linguistics research. Specifically, the teachers' insights offer recourse for developing strategies for exploiting corpora better in future ELT materials development. Moreover, teachers signalled effective means through which corpus linguistics can be embedded in future teacher education.*

KEYWORDS: CORPUS LINGUISTICS, COURSEBOOKS, MATERIALS DEVELOPMENT, SPOKEN LANGUAGE, TEACHER EDUCATION, TESOL

---

Manchester Metropolitan University, United Kingdom  
n.curry@mmu.ac.uk

Cardiff University, United Kingdom  
markg2@cardiff.ac.uk

Submitted: 2023-02-21 Accepted: 2023-10-03

## 1 Introduction

Despite a shift in language pedagogy over the last 50 years towards a communicative language teaching model, norms relating to written language often dominate, at the expense of spoken language (Carter & McCarthy, 2017; Timmis, 2016). In fact, as Carter and McCarthy (2015) state, “we are still struggling under the burden of a [grammatical] metalanguage inherited from writing that does not seem always to work for speaking, and many teaching resources have yet to reflect what everyday speaking is really like.”

Against this backdrop, there is a perennial request from learners for more conversation in classroom materials, often owing to general feelings of anxiety surrounding spoken performance (Tosun & Cinkara, 2019; Young, 1992). However, emulating conversation practices in ELT coursebooks is a challenge, as the features characteristic of everyday spoken language do not reflect the neat syllabi expected by teachers and students, globally (Gilmore, 2007; Timmis, 2016). Moreover, there are deeply held opinions relating to the usage and acceptability of the linguistic features of spoken language, which are often seen as ungrammatical or overly complex (Carter & McCarthy, 2017; Timmis, 2016). Compounding this challenge, contextual implications of limited space on a page, issues of data availability and language representation (Curry, Love, & Goodman, 2022), and the mutual shaping of coursebook syllabi and international assessments (Bailey & Masuhara, 2013) mean that deciding what goes into an ELT coursebook is a complex task. As a result of this complexity, features of spoken conversation are often missing from mainstream materials (Curry, 2023a; Curry et al., 2022; Gilmore, 2007; Timmis, 2016).

The absence of the representation of the characteristics of spoken conversation from ELT coursebooks is noteworthy. Owing to advances in a range of contexts, including within corpus linguistics (Boulton & Vyatkina, 2021; Chambers, 2019; Rundell & Stock, 1992) and usage-based theory (e.g., Ellis, 2002), for example, a canon of evidence exists to support the view that exposing learners to attested examples of language use can support language acquisition. Yet, despite this evidence, reticence remains on the part of materials developers regarding the use of corpus linguistics and examples of attested language use (Burton, 2012; Curry et al., 2022; Ur, 2017).

While the findings of corpus linguistics have played a role in informing coursebook materials development, e.g., *Touchstone* (McCarthy, McCarten, & Sandiford, 2004–2006) and *Evolve* (e.g., Goldstein & Jones, 2019), the use of such linguistic research is often peripheral. For example, corpus linguistics research has been used to develop feature boxes in coursebooks, displaying comparative frequencies of usage of lexical or grammatical items.

Similarly, it has been used to inform the writing of video scripts (Curry et al., 2022). This restricted uptake of corpus linguistics research for materials development is explicable, as there are often competing demands for publishers. As such, to address learners' needs, it is important to include spoken language in a way that:

- does not require much space on a page;
- can meet teacher expectations;
- can help learners improve their conversational competencies; and
- does not undermine learner success in language assessments if any features included may be perceived prescriptively as ungrammatical.

To respond to this challenge, this paper draws on the findings from a corpus-based conversation analysis of spontaneous spoken British English (Curry, 2023a) to inform qualitative workshops with teachers. The workshops were designed to gain a comprehensive perspective on the affordances of corpus linguistics for ELT materials development from a core stakeholder in ELT – teachers. Specifically, participants were asked to design lessons, to reflect on insights regarding the presence (or lack) of identified features of spoken language in the ELT materials, and to consider the affordances of language research and corpus linguistics for ELT materials development, lesson design, and teacher education. McCarthy (2008) notes the power of teachers in shaping and reshaping the ELT industry. Recognising this view, this paper argues that, by drawing on teachers' perspectives, it is possible to deliver guidance to ELT materials and assessment developers as to how they can incorporate corpus linguistics research on spoken language into ELT materials development. Furthermore, by accessing teachers' perspectives in this way, it is possible to outline how teachers can effectively engage with corpus linguistics research as part of their teacher education.

This study is therefore guided by the following research questions:

- (1) To what degree are language research and corpus linguistics seen by English language teachers to be valuable resources for improving the representation of spoken language in ELT materials?
- (2) How might these perceptions help inform the design of appropriate corpus-based pedagogic content on teacher education programmes?

To support this analysis, the paper presents a literature review in Section 2 that discusses spoken language representation in ELT materials, and teachers' perceptions of the affordances of corpora for materials development. Subsequently, in Section 3, the data and methods are presented,

outlining the workshops with and surveys of ELT practitioners that have informed this study. Next, in Section 4, the results are presented and discussed. Finally, Section 5 offers a brief conclusion, focusing on the implications for materials development, teacher education and future research directions.

## 2 ELT Coursebooks, Spoken Language Representation, and Teachers' Perspectives

As a type of instructional resource designed to support the teaching and learning of the English language, ELT materials range from published coursebooks (e.g., Goldstein & Jones, 2019), to other materials such as newspapers (Viana, 2022), and teacher-made resources (e.g., supplementary materials; Gray, 2016). ELT coursebooks are arguably the most widely used published materials in the ELT context, accounting for a global, multi-billion pound industry (Jordan & Gray, 2019). Though not without critique (e.g., Jordan & Gray, 2019), ELT coursebooks are largely perceived as serving to help learners develop their language skills (Gray, 2016), and are typically organised around grammar, vocabulary, and the four main language skills: listening, speaking, reading, and writing (Curry et al., 2022; Skela & Burazer, 2021; Timmis, 2016). Generally, ELT coursebooks include activities that enable learners to practise and develop their ability to understand and produce English in a range of contexts. However, the prioritisation of written language norms (Timmis, 2016) and consensus-driven syllabi of grammatical structures (Burton, 2018; Murphy, 2018) are among some of the key barriers to the adequate representation of spoken language in ELT coursebooks.

In terms of the focus on grammar in ELT coursebook development, grammar syllabi have been seen as vehicles to help students develop their language skills, with learners regularly asking for more opportunities to engage with grammar in their courses (Masuhara & Tomlinson, 2008). Such a view contends that grammar is an essential aspect of language learning, and that the ability to use it accurately is a prerequisite for effective communication (Masuhara & Tomlinson, 2008). It is reasonable to assert that those involved in creating and using ELT coursebooks see grammar syllabi as an integral facet of their development and utility (Başar, 2020; Burton, 2020; Masuhara & Tomlinson, 2008). For example, teachers have argued that coursebooks can act as an important tool for providing grammar instruction, as they offer a systematic and structured approach to teaching grammar (Başar, 2020), often following the presentation, practice, and production model (Masuhara & Tomlinson, 2008).

While there is an evident value attributed to grammar instruction in ELT coursebooks by their users, it is notable that the grammar that is taught in conjunction with the four main skills typically draws on written language norms (Skela & Burazer, 2021; Timmis, 2016). In fact, despite the canon of influential research on spoken grammar (e.g., Biber, Johansson, Leech, Conrad, & Finegan, 2021; Carter & McCarthy, 2006), advances in the area have made little impact on ELT grammar syllabi (Burton, 2020; Timmis, 2022). This is possibly due, in part, to the deeply held opinions relating to the usage and acceptability of features of spoken language, which are often considered ungrammatical (Carter & McCarthy, 2017; Timmis, 2016), ‘messy’ on the page, or too challenging to learn (Timmis, 2005). As such, despite the generally accepted view that developing learners’ ability to understand and communicate in spoken English is of critical importance in the ELT classroom (Başar, 2020), research has shown that ELT materials often do not provide an adequate representation of the characteristics of spoken language (Cullen & Kuo, 2007; Curry, 2023a; Timmis, 2022).

As much of the research discussed thus far demonstrates, corpus linguistics can play, and has played, a critical role in developing a contemporary understanding of spoken English language, building on the emerging role of corpus linguistics in ELT as part of the corpus revolution (Rundell & Stock 1992); notably, that revolution is ongoing (Boulton & Vyatkina, 2021; Chambers, 2019). Using both quantitative (e.g., Curry et al., 2022) and qualitative (e.g., Curry & Chambers, 2017; Curry, 2023b; O’Keeffe, McCarthy, & Carter, 2007) corpus linguistics approaches, researchers can conduct form-based analyses of, for example, word or phrase frequency or collocation across speaker’s age, gender, social background, accents, and regional dialect (McCarthy, 2020). Likewise, qualitatively, researchers can use corpora to understand the contextualised usage of spoken language, drawing on function-to-form and corpus pragmatic approaches (e.g., Curry, 2021; 2023b; O’Keeffe, 2018). Both quantitative and qualitative approaches can identify key features of spoken language that are shared across or differentiate demographic variables; such research has been used to inform pedagogic grammars (e.g., Carter & McCarthy, 2006), which describe contemporary spoken language independent of written-language norms.

In the context of ELT, such research has allowed for the identification of frequent patterns of language use as well as context- and domain-specific language use, and can prove useful for developing teaching materials (O’Keeffe et al., 2007). Frequency information for the language to be taught is embedded in coursebooks like *Touchstone* (McCarthy et al., 2004–2006), and research on language change (Curry et al., 2022) has been used to inform the language features sections of *Evolve* (Goldstein & Jones, 2019),

for example. Similarly, learner errors, based on learner corpus research, have been used to identify areas of challenge for learners from specific language backgrounds, exemplified also in *Evolve's* (Goldstein & Jones, 2019) 'accuracy check' feature. Corpus linguistics and ELT have been further linked in supporting the development of intercultural competence and 21<sup>st</sup> century skills (Curry, 2022). Recognising the growing importance of such competency development in ELT (Bedir, 2019; Nurkamto & Saleh, 2013), corpus research on spoken language can serve to represent the characteristics of spoken language within coursebooks, and expose learners to different accents, informal or spontaneous speech, and different varieties and domains of use (Sung, 2016).

Publishers and their editorial teams appear to see a value in corpus linguistics, recognising its affordances, for example, in signalling frequent language usage, or in the development of listening scripts (Curry et al., 2022). Likewise, assessment developers see the value of corpus linguistics for revealing patterns of difficulty in learner data with different L1 backgrounds, which can guide the development of localised training and support materials for teachers and learners in specific regions (Curry & Clarke, 2020). With regards to teachers, corpus linguistic research has been argued to have the potential to revolutionise English language teaching by providing teachers with evidence-based insights into the patterns of language use, enabling them to tailor their teaching materials to their learners' needs (Farr & O'Keeffe, 2019; Naismith, 2017). However, despite the growing interest in corpus-based language teaching, the extent to which teachers perceive the value of corpus linguistic research for language teaching remains unclear (Poole, 2022). Reticence on the part of teachers originates in issues including access, time, interest, and teacher education and development.

Finally, from a materials development perspective, Curry and Mark (2024) presents a study comparing corpus findings from a corpus-based conversation analysis of spoken language with spoken language representation in three ELT coursebooks. The study finds many opportunities throughout the book for 'speaking'; however, while activities often involve the discussion of opinions or facts, or responding to some kind of input, speaking activities often pertain to the kinds of tasks that involve 'speaking out loud', such as taking part in role-plays or semi-scripted dialogues. As such, speaking tasks do not necessarily focus on conversational features or strategies for engaging in conversations. In this view, dialogues often appear as vehicles for presenting grammatical or lexical content, and there is a lack of engagement with notions of co-construction in spoken conversation, the affordances of 'small words' (Carter & McCarthy, 2017), the use of response tokens, repairs, vagueness, hedging, and yes/no questions. The study proposes that ELT materials should include the likes of *I don't*

*know* as a turn-initiator, and questions as turn-ending features. Likewise, the study argues for the exploitation of the likes of *yeah* and *yes* as interactive devices with multiple functions, e.g., they act as floor-holding devices, as well as signalling both minimal and emphatic agreement.

Overall, the representation of characteristics of spoken language in ELT materials, or the lack thereof, is an issue that has been discussed extensively in the field. This paper argues that there remains a need to embed research on spoken language within ELT materials to better support teachers and learners. To do so, however, teachers need to be the driving force for change, and researchers must amplify teachers' voices and perspectives. The argument is based on four key rationales.

First, perennial requests from learners for more support with conversation skills (e.g., Tosun & Cinkara, 2019; Young, 1992) derive from a critical challenge faced by learners and teachers, i.e., learners who primarily learn from written-based materials may struggle to understand and communicate in spoken English situations (Timmis, 2005).

Second, the move to a focus on communication competencies (e.g. speaking, listening, and gesture; Bedir, 2019; Nurkamto & Saleh, 2013) offers an opportunity to revisit the notion of spoken language as a means of better equipping learners with the requisite competencies to manage and navigate spoken interactions (Timmis, 2016), and not simply to learn to speak out loud.

Third, the binary focus on British and American Englishes that pervades global ELT can limit learners' exposure to different accents and dialects, which they may encounter in the real world (Sung, 2016). Recognising the move in language education towards developing learners as global, socially responsible citizens (Bedir, 2019; Nurkamto & Saleh, 2013), there is a need to widen perspectives about language, language varieties, and language cultures (Sung, 2016). Embedding characteristics of spoken language within coursebooks, reflecting challenging listening situations, such as in noisy environments, different accents, and informal or spontaneous speech, can create opportunities to raise learners' awareness of language varieties and their capacity for engaging with valid, international varieties beyond the standard British and American English (Sung, 2016).

Fourth, as ELT materials are often based on British or American varieties of English, cultural references in ELT materials may be specific to one region or country, which can lead to misunderstandings among learners from different backgrounds (Young & Walsh, 2010). Representing characteristics of spoken language from a range of contexts and speakers could address emerging issues of representation in ELT coursebooks (Curry et al., 2023).

Evidently, there is an opportunity to use corpus linguistics to address extant issues of spoken language representation in ELT. This paper

recognises the role of teachers as key stakeholders in shaping the ELT industry (McCarthy, 2008), and reports on teachers' perspectives on the value of corpus linguistics research for materials development and teacher education.

### 3 Data and Methodology

The study is based on two workshops with ELT teachers. Section 3.1 describes the participants, the workshop structure and content, and the data taken from the workshops. Section 3.2 gives details on how these data are analysed.

#### 3.1 Workshop Data

The two workshops lasted 90 minutes each, and there were four participants in each workshop. The teachers reflected a range backgrounds and lengths of experience, including experienced teachers with over 15 years of experience as well as newly qualified teachers. Four of the teachers worked in the ESL context, in the UK, and four in the EFL context, in Malta. All those from the UK context were experienced teachers, all with at least 10 years of teaching in the UK, Europe, Central America, and Asia. All had extensive experience in using a wide range of published materials. Two of them were also teacher educators. The four teachers from the Maltese context ranged from a newly qualified teacher to a teacher with over 5 years' experience. The Maltese teachers had experience of teaching in Malta as well as in other international teaching contexts. None of the teachers had previous experience of corpus linguistics approaches in any context.

The two workshops were conducted on *Zoom* and were recorded. The workshops were designed to identify how the participants would design a speaking lesson, before and after being exposed to information on spoken language derived from corpus research (Curry, 2023a; Curry & Mark, 2024). Both workshops had the following eight stages:

- (1) Introduction to session
- (2) Task 1 – Review and lesson plan for speaking task materials
- (3) Submit Task 1 and debrief
- (4) Presentation on corpus findings
- (5) Task 2 – Review and lesson plan for speaking task materials
- (6) Submit Task 2 and debrief
- (7) Group reflection
- (8) Closing/Thanks



For (1), the introduction, the participants were introduced to the main aims of the study, the format of the workshop, and issues of ethics and consent were discussed to ensure participants understood their rights. For (2), Task 1, participants were given a two-page speaking section from a B1-level, adult, general English coursebook (Kay, Jones, Maggs, & Smith, 2009). They were asked to design a lesson plan, following the instructions presented in Figure 1.

## Lesson Plan Instructions

Imagine you are teaching a lesson using the page from the coursebook shared with you. Develop a lesson outline based on this page and fill in the template below. Please note:

- Lesson aims/objectives: Feel free to include details on the class make-up, level, reasons for studying.
- Lesson outline/steps/stages: Aim for the length of lesson you would normally do.
- Changes you would make to the material
- Things you would do outside of the material
- You have 15 minutes to complete this task.

Figure 1: Lesson plan instructions.

The teachers were also given a template into which they could write their lesson plans. A copy of this template is shown in Figure 2.

Participant:  
 Date:  
 Task:

Instruction	Lesson Notes
Please state your lesson aims and objectives.	
Please list the stages of your lesson here. Please also state the length of the lesson.	
Please outline the changes you would make to the materials shared with you. Include notes/rationales for doing so.	
Please share ideas for activities you would do outside of using this coursebook extract.	

Figure 2: Template for lesson planning.

In (3), each participant submitted the lesson plan by email, and as a group discussed their approaches to designing the lesson and their impressions of the speaking task. Part (4), the presentation of corpus research, consisted of a short presentation based on Curry (2023a) and Curry and Mark (2024), as discussed in Section 2. The presentation focused on the findings from corpus research on talk-in-interaction, as well as the review of the identified characteristics of spoken language in the ELT materials. For (5), participants repeated the activity for (2) with a different course-book extract. At this stage, they were particularly encouraged to consider incorporating any of the findings discussed in (4) that they considered relevant to the lesson. To avoid leading questions (Arsel, 2017), the participants were told to not feel pressured to use these findings, and they were reminded that the main aim of the study was to see whether this kind of research appeared useful for teachers. Next, (6), a second debrief, mirrored (3), and participants submitted their second lesson plan and discussed them together. Finally, before closing the session and thanking participants in (8), there was a group discussion (7), in which participants reflected on the affordances of language research and the affordances of corpus linguistics for language teaching materials development and teacher education.

To counteract the effects of leading questions (Arsel, 2017) and group-think as part of the discussion (George, 2013), participants were sent a post-workshop survey, to give them an opportunity for individual responses. The surveys asked the following five questions:

- Have your thoughts about using research to inform your teaching changed after completing this workshop? Please explain your answer.
- What role do you see corpus linguistics research playing in language teaching materials development (if any)? Please explain your answer.
- What changes would you like to see in language teaching materials going forward (if any)? Please explain your answer.
- In your view, what roles, if any, could linguistic research play in informing teacher education and development courses? Please explain your answer.
- Do you have any other reflections you would like to share based on the workshop?

Overall, the data collected included:

- Two 90-minute recordings of workshop discussions with eight participants.
- Sixteen lessons plans, with eight for Task 1 and eight for Task 2.
- Eight post-workshop surveys.

### 3.2 Workshop Analysis

In terms of analysis, the debriefs and discussions, lesson plans, and post-workshop surveys were analysed from a bottom-up perspective, drawing on critical grounded theory approaches (Curry & Pérez-Paredes, 2023; Hadley, 2017). The review of the lesson plans indicates the impact and influence of the workshop on teacher development, and the key themes that emerged based on the analysis of debriefs, discussions, and post-workshop surveys are:

- Research and corpus linguistics for ELT materials development;
- Future needs in ELT materials development;
- Teacher development and corpus linguistics.

## 4 Results and Discussion

This section presents and discusses the results of the study, beginning with a review of the lesson plans, in Section 4.1. Section 4.2 discusses research and corpus linguistics for ELT materials development, which is followed by a discussion of future needs in ELT materials development in Section 4.3. Finally, in Section 4.4, the topic of teacher development and corpus linguistics is discussed.

### 4.1 Lesson Plans

In reviewing the lesson plans for Task 1 and Task 2, the teachers largely followed a similar trajectory, working initially on topics more closely related to the content on the page in Task 1 and moving towards more spoken language modifications in Task 2. For example, in Task 1, the teachers focused on:

- Language and lexis.
- Creating interactivity and a space for speaking.
- Concept-checking and monitoring speaking.

In Task 1, the teachers were led by the coursebook (Bedir, 2019), focusing their initiatives on managing the classes and delivering the content. There is a notable difference in the content of lesson plans in Task 2, where teachers focused on:

- Notions of frequent language.
- Conversational practices, e.g., interrupting.
- The value of small words.
- Language awareness raising.

There is a clear movement in Task 2 towards a focus on spoken language. The teachers saw an opportunity to draw students' attention to small words, register variation, and spoken grammar, reflecting an advanced engagement with spoken language research for language teaching and learning. One could argue that the teachers were led to this thinking (Arsel, 2017) by the structure of the workshop. Therefore, the discussions that followed the lesson planning sought to unpack these decisions. These discussions, and the post-workshop survey responses, were analysed, and the findings are discussed in the next section.

#### 4.2 Research and Corpus Linguistics for ELT Materials Development

In both the group discussions and the post-workshop survey, all the teachers saw a value in using research and corpus linguistics to inform materials development. For example, one teacher noted that engagement with research was a form of professional development (Farr & O'Keeffe, 2019), developing their research literacy:

My participation in the workshop confirmed how important research is; however, my idea of research was limited mainly to teachers carrying it out in order to improve their skills, teaching and becoming more efficient in their profession. Thanks to the workshop, my idea of research has expanded.

Drawing on the notions of frequency as indicative of relevance for learners (Ellis, 2002), the teachers shared the view that frequency information was valuable for making choices, as the following extracts demonstrate:

I had not thought of checking the frequency in a corpus to see if an activity is relevant so it opened my eyes. Thank you.

It was interesting in general, but it was eye-opening to see how much more frequent certain phrases are over others and that made me consider whether coursebooks and teachers are giving enough importance to these more frequent phrases over others.

These same ideas were reported as key factors guiding the lesson design in Task 2. Teachers reported that knowing which words are common is useful to share in classes.

Interestingly, the role of frequency for determining relevance is not often reflected in ELT materials, where authorial intuitions and levelled grammar syllabi would limit lower-level students from engaging with such spoken language materials (Timmis, 2016). Notably, key stakeholders responsible for materials development, such as publishers, have argued that including frequency information in materials would not be useful to teachers (Curry et al., 2022). Reflecting on their lesson design, teachers' views in this study

conflict with those of publishers, as reported in Curry et al. (2022), as the teachers argue that ‘harder words’ at higher levels are usually uncommon words, and that materials would be better if they focused on very common, seemingly simple words that can do a range of useful things in a conversation. For example, reflecting on Task 2, one teacher argued that the inclusion of words like *well* in materials could be staggered in order to unpack their polysemy and rhetorical value at higher levels.

Echoing the long-recognised value of research in education (Frankenberg-Garcia, 2012; Tyne, 2012; Woore, Mutton, & Molway, 2020), some teachers see the value of research in equipping them with the knowledge to falsify their intuitions:

It has made me realise that I should use research to check how common phrases are and how they are used in real conversations. I think I rely too much on my own opinion at the moment.

[Using research is] definitely more authentic and more creative than blindly following a coursebook.

Focusing specifically on corpus linguistics, teachers echoed positive views on research, noting, for example, that:

If the corpus is wide enough or relevant to the course being designed, it should be a cornerstone of the materials and language used within the course.

These perspectives also guided the lesson design in Task 2, as teachers argued that corpus insights had made them more aware of how language works, and gave them a means of selecting words based on utility and relevance. Such views are shared by Frankenberg-Garcia (2012), O’Keeffe et al. (2007) and Tyne (2012), for example, leading to arguments for corpus linguistics to be central to language teaching materials development.

Further supporting this view, one teacher also noted that corpus linguistics “can facilitate learning and make teaching more natural.” This view was echoed throughout the discussions, as teachers saw an opportunity to modify existing materials using corpus linguistics:

By using corpus linguistics, the material can be adapted to reflect how language is really being used and this will help students in their understanding of natural conversation outside the classroom, and it will help to improve their fluency.

Filtering through the discussion of research and corpus linguistics for ELT development, some more complex perspectives emerged. Though all the teachers saw a value in research and corpus linguistics, notions of native-speaker norms emerged among the less experienced EFL teachers. Non-native speakers were positioned from a deficit perspective, as the

following extract from a second language (L2) English speaking teacher shows:

It would be useful to see course books that used language that is more frequently used in native speakers conversations.

These views echo a sustained perspective among English language teachers, in which first language (L1) English speakers are privileged over L2 teachers (Bolstad & Zenuk-Nishide, 2015). Such a view challenges the role of English as an international language in ELT (Sung, 2016) and, notably, is not shared by other key stakeholders, e.g., publishers, who argue that corpora are not sufficiently representative to support their move toward international English (Curry et al., 2022).

Interestingly, teachers argued that it is their responsibility to engage with such research, as they argue that coursebooks (and arguably their producers) do not:

It has highlighted the need for research into the evolving nature of spoken language and how as teachers, we cannot rely on coursebooks to provide us with the necessary resources.

This view of materials providers reflects the findings of Burton (2012) and Ur (2017), who similarly argue that materials producers do not have time to engage with research. This notion of pros or cons for teachers' direct engagement with corpus research pervades the literature (e.g., Pérez-Paredes, 2022), and is encapsulated by this remark from a teacher:

I read a bit about conducting activities involving the learners in using the corpus itself. It comes with benefits and limitations.

### 4.3 Future Needs in ELT Materials Development

When discussing ELT materials development, teachers discussed changes and developments they would like to see in the industry. Expressing contrasting view to those of publishers (Curry et al., 2023), teachers saw a value in explicitly engaging with frequency information, noting that:

All too often, the language [presented in ELT materials] is unnatural and low frequency.

As such, a value is attributed to frequency, which is a view shared by teachers in Tyne (2012), for example.

Teachers also call for a greater focus on speaking, that can:

...create real speaking opportunities within classrooms rather than just role-plays, and for such activities not to be at the end of a vocab/grammar heavy input, which often means the learner is tired and overloaded at that point.

This call is noteworthy, as it reflects a clear, sustained need from teachers for higher quality spoken-language teaching materials that can support learner engagement with the characteristics of spoken conversation (Tosun & Cinkara, 2019).

In seeking a means to address this issue of developing speaking lessons, one idea that emerged in the discussion with teachers related to contextualisation. Teachers argued that grammar and vocabulary should be presented “how it is actually used rather than its traditional use.” To do so, one teacher noted that:

I would like language to be more contextualised in course books and I would like them to include more natural conversations as models.

Carrying this thought forward, they noted that before listening activities, coursebooks could include a small paragraph contextualising the speaking environment. Such a paragraph could be used as a vehicle for exposing learners to relevant vocabulary, and the dialogues could then include more examples of characteristics of spoken language, such as small words (Carter & McCarthy, 2017).

Teachers also discussed the importance of representation in materials (Gray, 2016), arguing that:

More real-life contexts should be provided before controlled practice of the language and listening materials need to be more authentic.

Moreover, offering a critique of the current status quo of ELT materials development, teachers encouraged materials of the future to be less “Eurocentric” and to include more “real-world language and tasks”. However, in terms of language variety representation, issues of native-speakerism re-emerge, with teachers noting:

Learners should be trained to deal with native speakers who do not use learner-speak with them.

I still believe that it's important to teach Standard English and that students should know how to speak proper English.

Not only do such views conflict with global publisher agendas (Curry et al., 2023), they also defy advances in English as an international language (Sung, 2016). While there is expansive literature challenging this deficit representation of L2 speakers (Bolstad & Zenuk-Nishide, 2015), such perspectives evidently remain among practitioners.

#### 4.4 Teacher Development and Corpus Linguistics

Finally, in terms of teacher development, all teachers saw a value in engaging with research and corpus linguistics, reflecting a well-recognised affordance of corpus linguistics in the literature (Farr & O’Keeffe, 2019):

The sooner the future educators realise its benefits, the better. It will make them more aware as teachers and has beneficial effects on their students in the future.

Linguistic research should definitely be used to inform teacher training and development as we need to be familiar with the changing nature of English usage.

Interestingly, teachers saw a power in publishers, seeing them as agents for change and development, noting that they can create change via training programmes, as the following extract attests:

Some teachers tell students that what they hear outside the classroom is grammatically incorrect or wrong and I think this would help to soften the rigid attitudes some teachers have if training was provided from reputable sources then it would be more likely to be received as an acceptable teaching approach.

While McCarthy (2008) positions teachers as the driving force for change in the profession, teachers appeal to the likes of materials developers, i.e., publishers, in this role. Therefore, unlocking teachers’ understanding of their own role and agency in effecting change may still be necessary. Arguably, bridging teachers and publishers and creating dialogue between them may be an effective way to do so.

In terms of audiences for such education and professional development, teachers saw a value for both early career and advanced teacher education programmes:

It could be useful to reference spoken grammar and corpus data for more than just a passing mention in pre-service courses, along with providing novice teachers with some awareness of these tools and how they can be incorporated in teaching.

It would be good especially for more advanced teacher training so that teachers don’t become rigid in their planning.

These comments reflect the view that engagement with corpus linguistics is potentially of value for teachers at any stage of their development and career (Callies, 2019), and, in the case of this study, there is a volition on the part of the teachers to engage with ‘corpus thinking’ (Mauranen, 2004; Tyne, 2012).



## 5 Conclusion and Implications

Overall, this study aimed to investigate the degree to which language research and corpus linguistics are seen to be valuable resources for improving the representation of spoken language in ELT materials by English language teachers. Practitioner reflections made clear: (1) the perceived value of frequency information for ELT materials development; (2) the potential role of small words as opposed to infrequent words, for generating more challenging content for higher proficiency levels; and (3) the potential value of corpus linguistics for teacher education and development.

Notably, while the teachers who participated in the workshops varied in background and profile, a saturation of themes emerged, spanning research and corpus linguistics for ELT materials development, the future of ELT materials development, and corpus linguistics for ELT teacher development. While the general view of the affordances of research on spoken language is positive, reflected in the responses from the teachers, challenges also emerged with regards to issues of representation, language standards, and deficit perspectives of L2 speakers.

Following McCarthy (2008), this paper sought to investigate teachers' perspectives, recognising their roles as key agents shaping the ELT industry. Based on these perspectives, it is reasonable to argue that corpus linguistics can be used to address deficits in evidence-based representation of spoken language in ELT materials, and develop much-needed content for informing future teacher education programmes (Boulton & Vyatkina, 2021; Chambers, 2019; Rundell & Stock, 1992). Contrasting their views with those of publishers (Curry et al., 2022), for example, there is a shared value held for corpus linguistics. However, while teachers focus on the importance of frequency and embedding research within speaking dialogues, for publishers, research on spoken language appears to largely support the signalling of useful language or the development of listening scripts. Furthermore, while publishers see challenges with representation and the need for the development of pedagogical corpora that reflect English as an international language, teachers appear concerned with standard varieties and L1 norms. Given that both teachers and publishers are key agents in the shaping of the ELT profession, this disparity in their approach to spoken language representation is noteworthy.

Likewise, when compared to work with assessment developers on the affordances of corpus linguistics (Curry & Clark, 2020), differences in perspectives emerge. For assessment developers, a clear value of corpora is their capacity to reveal patterns in learner data that may disadvantage learners in different regions when engaging in global assessments. Issues of assessment rarely emerged among the teachers' discussions, and

their focus on language in the real world did not engage with assessment expectations. Nonetheless, points of synergy occurred with regards to the affordances of corpora for developing training resources for teachers, as assessment developers could use corpora to develop localised training and support materials for teachers and learners in specific regions.

Evidently, there are potential areas of conflict across and between the three key stakeholders, and the teachers welcomed the opportunity to engage with this research project, recognising its aims to feedback to publishers and assessment developers. As one teacher notes:

I especially liked the idea of consulting the outcomes with publishers. Hopefully, they will be willing to listen and consider teachers' point of view.

To conclude this paper, there is evidently a clear value in using language research to inform both ELT materials and teacher development. Challenges with spoken language teaching and learning can be addressed through corpus research, and there is a need to unpack perceptions of spoken language and coursebook syllabi to further develop the field. As such, this paper argues for greater engagement with teachers' perspectives for developing materials and teacher education resources.

## Acknowledgements

We would like to thank Cambridge University Press and Assessment for supporting this research through the Cambridge Funded Research Programme scheme.

## References

- Arsel, Z. (2017). Asking questions with reflexive focus: A tutorial on designing and conducting interviews. *Journal of Consumer Research*, 44(4), 939–948. <https://doi.org/10.1093/jcr/ucx096>
- Bailey, K. M., & Masuhara, H. (2013). Language testing washback: The role of materials. In B. Tomlinson (Ed.), *Applied linguistics and materials development* (pp. 303–318). Bloomsbury. <https://doi.org/10.5040/9781472541567.ch-019>
- Başar, S. (2020). Teacher cognition in grammar teaching: A case study in a Turkish EFL context. *International Journal of Modern Education Studies*, 4(2), 57–80. <https://doi.org/10.51383/ijonmes.2020.51>
- Bedir, H. (2019). Pre-service ELT teachers' beliefs and perceptions on 21st century learning and innovation skills (4Cs). *Journal of Language and Linguistic Studies*, 15(1), 231–246. <https://doi.org/10.17263/jlls.547718>
- Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of spoken and written English*. John Benjamins.

- Bolstad, F., & Zenuk-Nishide, L. (2015). Beyond the deficit model: Co-constructing team teaching to address learner goals and needs. In A. Tajino, T. Stewart, & D. Dalsky (Eds.), *Team teaching and team learning in the language classroom* (pp. 78–93). Routledge.
- Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning and Technology*, 25(3), 66–89. <https://doi.org/10.1257/73450>
- Burton, G. (2012). Corpora and coursebooks: Destined to be strangers forever? *Corpora*, 7(1), 91–108. <https://doi.org/10.3366/cor.2012.0019>
- Burton, G. (2018). Grammar syllabus. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–6). Wiley. <https://doi.org/10.1002/9781118784235.eelt1013>
- Burton, G. (2020). Grammar. *ELT Journal*, 74(2), 198–201. <https://doi.org/10.1093/elt/ccaa004>
- Callies, M. (2019). Integrating corpus literacy into language teacher education. In S. Götz & J. Mukherjee (Eds.), *Learner corpora and language teaching* (pp. 245–263). John Benjamins.
- Carter, R., & McCarthy, M. (2006). *Cambridge English grammar*. Cambridge University Press.
- Carter, R., & McCarthy, M. (2017). Spoken grammar: Where are we and where are we going? *Applied Linguistics*, 38(1), 1–20. <https://doi.org/10.1093/applin/amu080>
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 52(4), 460–475. <https://doi.org/10.1017/S0261444819000089>
- Cullen, R., & Kuo, I. V. (2007). Spoken grammar and ELT course materials: A missing link?. *TESOL Quarterly*, 41, 361–386. <https://doi.org/10.1002/j.1545-7249.2007.tb00063.x>
- Curry, N. (2021). *Academic writing and reader engagement: Contrasting questions in English, French and Spanish corpora*. Routledge.
- Curry, N. (2022). On contrastive analysis and language pedagogy: Reimagining applications for contemporary English language teaching. In L. McCallum (Ed.), *English language teaching in the European Union: Theory and practice across the region* (pp. 239–256). Springer.
- Curry, N. (2023a). *Investigating publisher, assessment developer, and teacher application of corpus research to ELT materials and assessment development* [Conference presentation]. International Joint Conference on ELT, Korea.
- Curry, N. (2023b). Question illocutionary force indicating devices in academic writing: A corpus-pragmatic and contrastive approach to identifying and analysing direct and indirect questions in English, French, and Spanish. *International Journal of Corpus Linguistics*, 28(1), 91–119. <https://doi.org/10.1075/ijcl.20065.cur>
- Curry, N., & Chambers, A. (2017). Questions in English and French research articles in linguistics: A corpus-based contrastive analysis. *Corpus Pragmatics*, 1(4), 327–350. <https://doi.org/10.1007/s41701-017-0012-0>

- Curry, N., & Clark, T. (2020). Corpus-informed evaluation of examination criteria for MENA. In L. McCallum & C. Coombe (Eds.), *The assessment of L2 written English across the MENA Region: A synthesis of practice* (pp. 359–392). Palgrave Macmillan.
- Curry, N., & Mark, G. (2024). Applications of corpus linguistics in language education: Teacher, editor, and assessment developer perspectives. *Research Notes*.
- Curry, N., Love, R., & Goodman, O. (2022). Adverbs on the move: Investigating publisher application of corpus research on recent language change to ELT coursebook development. *Corpora*, 17(1), 1–38. <https://doi.org/10.3366/cor.2022.0233>
- Curry, N., & Pérez-Paredes, P. (2023). Using corpus linguistics and grounded theory to explore EMI stakeholders' discourse. In S. Curle & J. K. H. Pun (Eds.), *Qualitative research methods in English medium instruction for emerging researchers: Theory and case studies of contemporary research* (pp. 45–61). Routledge.
- Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188. <https://doi.org/10.1017/S0272263102002024>
- Farr, F., & O'Keeffe, A. (2019). Using corpora to analyse language. In S. Walsh & S. Mann (Eds.), *Routledge handbook of English language teacher education* (pp. 268–282). Routledge.
- Frankenberg-Garcia, A. (2012). Integrating corpora with everyday language teaching. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 36–53). Masarykova univerzita nakladatelství.
- George, M. (2013). Teaching focus group interviewing: Benefits and challenges. *Teaching Sociology*, 41(3), 257–270. <https://doi.org/10.1177/0092055X1246529>
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97–118.
- Goldstein, B., & Jones, C. (2019). *Evolve level 6*. Cambridge University Press.
- Gray, J. (2016). ELT materials: Claims, critiques and controversies. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 95–108). Routledge.
- Hadley, G. (2017). *Grounded theory in applied linguistics research: A practical guide*. Routledge.
- Jordan, G., & Gray, H. (2019). We need to talk about coursebooks. *ELT Journal*, 73(4), 438–446. <https://doi.org/10.1093/elt/ccz038>
- Kay, S., Jones, V., Maggs, P., & Smith, C. (2009). *New inside out: Upper intermediate student's book*. Macmillan Education.
- Masuhara, H., & Tomlinson, B. (2008). Materials for general English. In B. Tomlinson (Ed.), *English language learning materials: A critical review* (pp. 17–37). Bloomsbury.
- Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 89–105). John Benjamins.

- McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, 41(4), 563–574. <https://doi.org/10.1017/S0261444808005247>
- McCarthy, M. (2020). *Innovations and challenges in grammar*. Routledge.
- McCarthy, M., McCarten, J., & Sandiford, H. (2004–2006). *Touchstone series*. Cambridge University Press.
- Murphy, R. S. (2018). The concept of syllabus design and curriculum development: A look at five major syllabus designs. In A. Faravani (Ed.), *Issues in syllabus design* (pp. 1–23). Brill.
- Naismith, B. (2017). Integrating corpus tools on intensive CELTA courses. *ELT Journal*, 71(3), 273–283. <https://doi.org/10.1093/elt/ccw076>
- Nurkamto, J., & Saleh, M. (2013). Integrating life skills in English language teaching (ELT): Strategies, problems and their possible solutions. *Language Circle: Journal of Language and Literature*, 7(2), 75–91. <https://doi.org/10.15294/lc.v7i2.2600>
- O’Keeffe, A. (2018). Corpus-based function-to-form approaches. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in pragmatics* (pp. 587–618). Mouton de Gruyter.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning*, 35(1–2), 36–61. <https://doi.org/10.1080/09588221.2019.1667832>
- Poole, R. (2022). “Corpus can be tricky”: Revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641. <https://doi.org/10.1080/09588221.2020.1825095>
- Rundell, M., & Stock, P. (1992). The corpus revolution. *English Today*, 8(3), 21–32. <https://doi.org/10.1017/S0266078400006751>
- Skela, J., & Burazer, L. (2021). The architecture of ELT coursebooks: The internal organization of coursebook units. *Vestnik za tuje jezike*, 13(1), 383–407. <https://doi.org/10.4312/vestnik.13.383-407>
- Sung, C. C. M. (2016). Exposure to multiple accents of English in the English language teaching classroom: From second language learners’ perspectives. *Innovation in Language Learning and Teaching*, 10(3), 190–205. <https://doi.org/10.1080/17501229.2014.936869>
- Timmis, I. (2005). Towards a framework for teaching spoken grammar. *ELT Journal*, 59(2), 117–125. <https://doi.org/10.1093/eltj/ccj025>
- Timmis, I. (2016). Humanising coursebook dialogues. *Innovation in Language Learning and Teaching*, 10(2), 144–153. <https://doi.org/10.1080/17501229.2015.1090998>
- Timmis, I. (2022). Theory and practice in materials development. In J. Norton & H. Buchanan (Eds.), *The Routledge handbook of materials development for language teaching* (pp. 30–46). Routledge.
- Tosun, Ö. Ö., & Cinkara, E. (2019). Coursebook dependency in secondary and tertiary-level EFL teachers. *How*, 26(1), 81–105. <https://doi.org/10.19183/how.26.1.445>

- Tyne, H. (2012). Corpus work with ordinary teachers: Data-driven learning activities. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 136–151). Masarykova univerzita nakladatelství.
- Ur, P. (2017). Applications of research to materials design. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. III, pp. 132–143). Routledge. <https://doi.org/10.4324/9781315716893-10>
- Viana, V. (2022). Register variation in newspapers: Working with multidimensional analysis in English language teacher education. In V. Viana (Ed.), *Teaching English with corpora* (pp. 203–214). Routledge.
- Woore, R., Mutton, T., & Molway, L. (2020). ‘It’s definitely part of who I am in the role’: Developing teachers’ research engagement through a subject-specific Master’s programme. *Teacher Development*, 24(1), 88–107. <https://doi.org/10.1080/13664530.2019.1693421>
- Young, D. J. (1992). Language anxiety from the foreign language specialist’s perspective: Interviews with Krashen, Omaggio Hadley, Terrell, and Rardin. *Foreign Language Annals*, 25(2), 157–172. <https://doi.org/10.1111/j.1944-9720.1992.tb00524.x>
- Young, T., & Walsh, S. (2010). Which English? Whose English? An investigation of ‘non-native’ teachers’ beliefs about target varieties. *Language, Culture and Curriculum*, 23(2), 123–137. <https://doi.org/10.1080/07908311003797627>