Asia Pacific Journal of Corpus Research
Vol. 2, No. 1, 13-22

http://www.icr.or.kr/apjcr
e-ISSN 2733-8096
https://doi.org/10.22925/apjcr.2021.2.1.13

Fillers in the Hong Kong Corpus of Spoken English (HKCSE)

Andy Seto
(The Hong Kong Institute of Vocational Education)

Seto, A. (2021). Fillers in the Hong Kong Corpus of Spoken English (HKCSE). ***Asia Pacific Journal of Corpus Research, 2*****(1), 13-22.**

The present study employed an analytical framework that is characterised by a synthesis of quantitative and qualitative analyses with a specially designed computer software SpeechActConc to examine speech acts in business communication. The naturally occurring data from the audio recordings and the prosodic transcriptions of the business sub-corpora of the HKCSE (prosodic) are manually annotated with a speech act taxonomy for finding out the frequency of fillers, the co-occurring patterns of fillers with other speech acts, and the linguistic realisations of fillers. The discoursal function of fillers to sustain the discourse or to hold the floor has diverse linguistic realisations, ranging from a sound (e.g. 'uhuh') and a word (e.g. 'well') to sounds (e.g. 'um er') and words, namely phrase ('sort of') and clause (e.g. 'you know'). Some are even combinations of sound(s) and word(s) (e.g. 'and um', 'yes er um', 'sort of erm'). Among the top five frequent linguistic realisations of fillers, 'er' and 'um' are the most common ones found in all the six genres with relatively higher percentages of occurrence. The remaining more frequent realisations consist of clause ('you know'), word ('yeah') and sound ('erm'). These common forms are syntactically simpler than the less frequent realisations found in the genres. The co-occurring patterns of fillers and other speech acts are diverse. The more common co-occurring speech acts with fillers include informing and answering. The findings show that fillers are not only frequently used by speakers in spontaneous conversation but also mostly represented in sounds or non-linguistic realisations.

Keywords: Corpus Linguistics, Fillers, Linguistic Realisation, Speech Act Annotation, Spoken Business Discourse

1. Introduction

Language communication involves not only the production of sounds and lexicogrammar but also the use to which utterances are conventionally put in a community. The speaker normally expects that his or her communicative function in an utterance determined by the local context will be recognised by the hearer. This communicative function is in general referred to as speech act. Since the 1960s, empirical studies of speech acts have examined spoken or written language in different contexts of communication, including business and a range of speech acts, including fillers. A variety of speech events with specific objectives have been investigated in contexts of intracultural, intercultural, and cross-cultural communication. These speech act studies have addressed different research focuses, including the linguistic realisations of a speech act and the cultural differences in the use of a speech act. These studies attempt to examine the linguistic features and functions of certain speech acts to find patterns, similarities and differences. However, a majority of these studies are not conducted in the context of business communication.

Regarding methodology, many previous studies are not grounded in naturally occurring spoken data but based on data obtained from role-play interactions, discourse completion tests or written responses to prepared situations. It is argued that these elicited data neither reveal the range of strategies used nor the actual rate of occurrence of a speech act in real interactions. Though more and more studies are conducted with reference to naturally occurring data in which the participants are

engaging in a real-world interaction in a particular situational setting or a speech event with the assistance of computer technology, the studies on fillers in business communication remain limited.

Regarding fillers, some studies investigate the impact of different factors, including gender, age and socio-economic class, on the use of different filler words with naturally occurring spoken data from the British National Corpus (BNC) and the London-Lund Corpus (LLC) (Laserna, Seih, & Pennebaker, 2014). Others look into the common types of fillers in different languages such as German, Danish, Korean, and Japanese and examine the use of fillers cross-culturally. Some focus on the frequency and the functions of fillers in academic presentation or public speech (Kharismawan, 2017). However, these recent studies only focused on the features of a limited number of fillers found in the data.

The present paper aims to investigate, both quantitatively and qualitatively, the speech acts of fillers and their linguistic realisations in the business sub-corpus of the Hong Kong Corpus of Spoken English (prosodic) (hereafter HKCSE (prosodic)). The spoken business genres are meetings, telephone calls and conference calls, informal office talk, service encounters (airport and hotel), question and answer (Q&A) sessions, and interview (job and placement). The study also attempts to offer a new perspective for exploring the implementation of manual speech acts annotation with the aid of a specially designed software SpeechActConc in identifying all fillers in the sub-corpus and scrutinising their linguistic features in different genres.

2. Background

2.1. Speech Act

Speech act theory, first discussed among a number of ordinary language philosophers involved in analytical philosophy in the 1960s, focuses not on what language is but what language does. J. L. Austin and J. R. Searle are two prominent advocates who come up with different yet related classifications of speech acts. Searle revises and modifies Austin's (1975) five categories of speech acts, namely 'verdictives', 'exercitives', 'commissive', 'expositives', and 'behabitives', and proposes another five categories of speech acts that emphasise on the functional classification of illocutionary acts, namely 'assertives', 'directives', 'commissives', 'expressives', and 'declarations'.

Speech act is understood in three 'senses' or 'classes', namely 'locutionary act', 'illocutionary act', and 'perlocutionary act' (Austin, 1975, p. 94). A locutionary act, the first sense of an utterance, is 'the performance of an act of saying something' (Austin, 1975, pp. 99-100), the formal literal meaning of the words uttered. It is expressed as a meaningful linguistic utterance with reference to some object in the world. A locutionary act consists of three distinctive acts, which are the 'phonic act' of uttering some sounds, the 'phatic act' of uttering some words according to grammar rules, and the 'rhetic act' of reporting the phatic act with similar vocabulary and grammar. An illocutionary act, the second sense of an utterance, is 'the performance of an act in saying something' (Austin, 1975, p. 94), referring to the force or the intention behind the utterance, which is the communicative function of performing an utterance. A perlocutionary act, the third sense of an utterance, refers to certain consequential results or effects on the feelings, thoughts or action of the listener, the speaker, or other people (Austin, 1975, p.101). It is the performance of an act 'by saying something' (Austin, 1975, p.109), the effect the utterance might have.

2.2. Empirical Genre Studies in Business Communication

Genre analysis in the field of discourse and communication studies has become more important among academics since the 1990s (Bhatia, 1996). One orientation to genre theory is 'to develop a grounded description of language uses in professional and academic settings' (Bhatia, 1996, p.46). Not only the communicative purposes but also the structural features together with particular practices

and strategies are considered to be relevant factors for defining a genre (Koester, 2010). The main concern of this orientation is to apply genre analysis to specialist language teaching and learning and to focus on consistency of communicative purposes that control both lexicogrammatical and discursive choices (Bhatia, 1996).

Business communication is a domain, among the others such as education, healthcare, social welfare, media, etc., in the context of professional and workplace communication (Evans, 2010). The analysis of spoken and written workplace discourse has grown considerably in the last decade (Koester, 2010). A number of research studies have been done in some of the contexts found in the business sub-corpus of the HKCSE (prosodic). However, most of these studies take a conversation analysis approach or a sociolinguistics approach to analyse the interaction in the context of intracultural or intercultural business communication. There is in general a lack of speech act studies in business communication with focuses on the annotation of speech acts, the frequency and co-occurrence of speech acts, or the linguistic realisations of a particular speech act in different genres of business communication.

2.3. Speech Acts Annotation

Speech acts annotation can be broadly divided into four orientations, namely task-oriented (Kallen & Kirk, 2012), non-task-oriented (De Felice, Darby, Fisher, & Peplow, 2013), middle-ground-oriented (Weisser, 2003), and discourse-/conversation-oriented (Stenström, 1994; Tsui, 1994). The applicability of the speech act categories in these orientations is restrictive to the specific domains from which the categories are generated. The varied annotation frameworks or schemes show that there is no universal annotation framework or scheme available; individual research study has its specific focuses and objectives.

Given the advancement of information and communication technology in the 1990s, researchers have been developing software to identify speech acts from corpus data in an automated way and to propose different classifications of speech acts, such as Leech & Weisser's Speech Act Annotation Corpus (SPAAC) and Rayson's Wmatrix. Some programs aim at automatically identifying and classifying the real intent of the speaker in an utterance, for example, Weisser's Dialogue Annotation and Research Tool (DART).

Recent studies also focus on the semiautomatic labeling of speech acts such as filled pauses that can reduce the workload of manual transcription. However, in principle, these programs cannot precisely annotate the pragmatic meaning of an utterance that conveys the speaker's intention as it is dependent on the discourse context and the background knowledge of each speech event.

2.4. Fillers

Fillers are commonly used in spoken conversation. They are also referred to as 'discourse markers' (Schiffrin, 1987), 'filled pauses' (Stenström, 1994), or 'filler words' (Laserna et al., 2014). They can be used by the speakers to organise the discourse, such as starting a conversation, introducing and marking the end of a topic, introducing a digression and marking the resumption of an old topic, signally the end of a conversation, and of course, serving as a filler or delaying tactic to sustain discourse or hold the floor (Stenström, 1994).

Tsui (1994) proposes that fillers are a 'dispreferred' response used to postpone the decision making, containing linguistic features of delay (p. 59). Stenström (1994) argues that fillers, as a type of hesitation phenomenon commonly observed in spontaneous speech, are problematic as they are 'fairly unpredictable' (p.8). They can be used by speakers in any places in their utterances, not only at the beginning of a turn, when more time is needed to plan and prepare what is to be said next. In short, fillers are a type of speech disfluencies that can be realised by a set of lexical items such as 'well', 'okay', 'you know', 'I mean' (Schiffrin, 1987; Stenström, 1994) or 'uh', 'um' (Clark & Fox Tree, 2002).

Most fillers work on the discourse level and do not carry a communicative function or purpose. Besides, they are syntactically independent from the environment or the speech event and convey collateral or interactional messages in a communication (Stenström, 1994). Their high frequency of occurrence in the spoken discourse is in consistency with the findings from the present study.

3. Methodology and Results

The present work aims to answer three research questions: (1) What are the frequencies of most frequent fillers in the business sub-corpus? (2) What are the common co-occurring patterns of fillers and other speech acts? (3) What are the featured linguistic realisations of fillers?

3.1. Data Collection

The corpus analysed in the study is the HKCSE (prosodic) compiled in the English Department of the Hong Kong Polytechnic University from the mid-1990s to the early 2000s (Cheng & Warren, 1999). It comprises four sub-corpora that represent the main overarching spoken genres found in the Hong Kong context, namely academic discourses, business discourses, conversations, and public discourses. Each sub-corpus consists of a variety of discourse types and participants. The composition of the business discourse (259,484 words in 112 files) has Q&A, conference call / video conferencing, informal office talk, interview, meeting, service encounter, workplace telephone call (Cheng, Greaves, & Warren, 2008).

Most of the data were audio-recorded with a MD-recorder by the researchers involved in the project, and audio recordings of presentations on the topics of business and financial services were downloaded from the websites of different organisations. The physical or institutional contexts of recording include meeting rooms in business organisations, government, and university offices; hotel and airport reception or information desks; and convention or conference rooms where business or financial service presentations took place (Cheng et al., 2008).

The audio recordings were orthographically transcribed by research assistants and cross-checked by experienced researchers working with spoken data. Each orthographic transcription was coded with background and contextual information about the participants and the genre, including the speaker's gender, age, L1, occupation group, and the relationship between or among the speakers. Paralinguistic features such as throat-clearing and coughing are indicated. Other discursive details such as truncated words, overlaps, and inaudible speech are identified. Pauses are marked with a differentiation between a brief pause and a unit pause that generally lasts for a few seconds.

The HKCSE (prosodic) was further enriched as a research, learning and teaching resource by adding a prosodic transcription (i.e. indicating speakers' intonation in the transcript) (Cheng et al., 2008). The HKCSE (prosodic) is believed to be the largest prosodically transcribed corpus currently in existence. The prosodic transcription of the orthographic transcription was made by adopting discourse intonation systems (tone unit, tone, prominence, key and termination) (Brazil, 1997). It was carried out by a research associate. The prosodic features were determined after repetitive listening to the recordings. Sample transcriptions were cross-checked regularly by the project consultant (Cheng et al., 2008).

3.2. Data Annotation

Central to the paper is a detailed manually annotated speech act corpus with reference to a speech act taxonomy. The manual annotation of the speech acts in the spoken business corpus was carried out by making reference to the audio recordings, the prosodic transcription of the data (Cheng et al., 2008), and a taxonomy of 69 speech acts (Stenström, 1994; Tsui, 1994; Stolcke, Ries, Coccaro, Shriberg,

Bates, Jurafsky, Taylor, Martin, Van Ess-dykema, & Meteer, 2000). Speech acts were identified and annotated in terms of their forms or linguistic realisations (locutionary act) and the particular communicative functions they represent (illocutionary act) in the corpus data. Consequential results or effects on the feelings, thoughts or actions of the listener or other people (perlocutionary act) of the locutionary act were also considered as they can give more contextual information to inform the annotation of the utterances in the business sub-corpus.

The speech act annotation involved an iterative process of utterance-by-utterance annotation and revision, making reference to both the audio-recordings and the prosodic transcriptions of the data (De Felice, 2013). The annotation had been checked with the supervisors and a fellow researcher on a regular basis until a consensus was reached. The study illustrates that the speech acts in the taxonomy are sufficient and relevant in vindicating the moment-by-moment speech act choices made by discourse participants to achieve their specific transactional and relational goals as the spoken discourse unfolds.

3.3. Data Analysis

The corpus data were analysed in terms of speech acts; in addition, they had to be marked in a specific way for the computer software SpeechActConc to analyse it quantitatively, in terms of the frequency of unique and co-occurring speech acts. SpeechActConc is a software written by Chris Greaves in 2010, specially designed for concordancing the speech acts and automatically finding 2, 3, or 4 speech act co-occurrence even when they occur in different positions relative to one another (i.e. positional variation) and when one or more speech acts occur in between the other speech acts (i.e. constituency variation).

These functions are similar to ConcGram, another software written by Chris Greaves for concgraming and concordancing corpus texts, which can perform automated as well as user specified concgram searches to find 2, 3, 4 or 5-word patterns of all the word associations in a text, both grammatical and semantic, and patterns are listed by frequency of occurrence (Cheng, Greaves, & Warren, 2006). In SpeectActConc, a basic search for a single speech act can be done in concordance search. After the speech act to search for in the corpus file has been entered, all of the instances of the searched speech acts will be listed. A co-occurring speech acts search can be carried out by setting the preferences and entering the search words (2 to 4 words). The co-occurring speech acts are then sorted based on their position relative to the centred string.

4. Research Findings

4.1. Frequency

The discoursal function of fillers to fill a gap in the discourse in the six genres has shown similarities and differences regarding their linguistic realisations. These filled pauses are used by speakers to sustain the discourse or to hold the floor. It is found that there is a wide range of linguistic realisations of fillers for this discoursal function with diverse frequencies and percentages of occurrence. Analysis of the filled pauses in the six genres of business discourse has produced 116 linguistic realisations of fillers. Among them the top three most frequent ones are shown in Table 1.

Table 1. Top three Most Frequent Fillers in the Six Genres

	er	um	you know
Meeting	1,042 (61.80%)	108 (6.41%)	69 (4.09%)
Telephone & Conference Call	212 (40.77%)	40 (7.69%)	110 (21.15%)
Informal Office Talk	410 (35.65%)	196 (17.04%)	313 (27.22%)
Service Encounter: Airport	204 (61.08%)	25 (7.49%)	2 (0.60%)
Service Encounter: Hotel	90 (45.45%)	11 (5.56%)	3 (1.52%)
Q&A Session	630 (51.72%)	72 (5.91%)	117 (9.61%)
Interview: Job	1,136 (39.72%)	567 (19.83%)	74 (2.59%)
Interview: Placement	1,292 (51.85%)	790 (31.70%)	49 (1.97%)

Table 1 shows the top three most frequently occurring linguistic realisations found in all the six genres. Both the first and the second ones are non-linguistic signals (Clark & Fox Tree, 2002). 'Er' occurs most frequently with an average of 48.51%, ranging from 35.65% in informal office talks to 61.80% in meetings. 'Um', the second most frequent filler, has an average of 12.70%. The percentages range from 5.56 in hotel service encounter to 31.70 in placement interview, which are obviously lower than those of 'er'. The third most frequent filler is 'you know'. Different from the first two, it is a linguistic signal (Clark & Fox Tree, 2002). The average percentage is 8.59, with percentages ranging from 0.60% in airport service encounter to 27.22 in informal office talk.

Other than these top three frequent fillers, commonly known fillers such as 'yeah' (6.70%), 'erm' (6.25%), 'okay' (3.49%), 'mm' (1.64%), 'right' (1.50%), 'well' (1.24%), 'mhm' (1.12%), 'yes' (1.10%), 'so' (0.6%), 'oh' (0.41%) are also found across the six genres. These filled pauses are either sounds or words, some with frequencies far lower than those of the top three fillers mentioned.

4.2. Co-occurrence

It is observed that the identification of different associations of fillers with other speech acts can describe the relationship between them as performed by different speakers in different genres, though the collocational patterns are not as discursively representative as expected. As said before, fillers are one of the most frequently co-occurring speech acts across the six genres. The top three co-occurred speech acts in alphabetical order with fillers are described in Table 2:

Table 2. Top Three Most Frequent Co-occurring Speech Acts with Fillers in the Six Genres

	Answer	Expand	Inform	Justify	Opine	Precede	Request	Thank
Meeting			29.79%	2.91%	5.95%			
Telephone & Conference Call		7.00%	19.55%			10.08%		
Informal Office Talk			18.67%	6.44%		18.33%		
Service Encounter: Airport	5.13%		12.70%				4.40%	
Service Encounter: Hotel	2.29%		10.01%					2.58%
Q&A Session			19.89%		14.57%	5.97%		
Interview: Job	12.14%	7.21%	20.51%					
Interview: Placement	19.71%		9.05%	10.00%				

In meetings, fillers are very commonly used by speakers when they inform [provides information] (29.79%), opine [expresses one's personal opinion] (5.95%) or justify [defends what was just said] (2.91%). In telephone and conference calls, fillers are used by speakers when they inform (19.55%), precede [links up what was said before and gives information, or comments on something in the preceding dialogue] (10.08%), or expand [gives complementary information] (7.00%). In informal office talks, fillers are used together with inform (18.67%), precede (18.33%), or justify (6.44%) In service encounters at airport check-in counters and information counters, fillers are frequently associated with the speech acts of informing (12.70%), answering [gives adequate information explicitly] (5.13%), or requesting [asks somebody to do something] (4.40%) while at the hotel

concierges and retail outlets, fillers frequently go with informing (10.01%), thanking [expresses gratitude] (2.58%), or answering (2.29%). In Q&A sessions, speakers mostly use fillers when they inform (19.89%), opine (14.57%), or precede (5.97%). In job interviews for a research assistant, fillers are used when speakers inform (20.51%), answer (12.14%), or expand (7.21%) whereas in placement interview for a hotel trainee, fillers are usually used when they answer (19.71%), justify (10.00%), or inform (9.05%). It can be seen from Table 1 that among the top three most frequent co-occurring speech acts, 'inform' appears in every genre, followed by 'answer' and 'opine'.

4.3. Linguistic Features

It is found that the more complicated the realisations are, the less frequent they are. In other words, the more common forms are syntactically simpler than the less frequent realisations found in the genres. The 116 linguistic realisations are diverse, ranging from a sound (e.g. 'uhuh') and a word (e.g. 'well') to sounds (e.g. 'um er') and words (e.g. 'you know') to combinations of sound(s) and word(s) (e.g. 'and um', 'yes er um', 'sort of erm').

Apart from the top three most frequent fillers described in 4.1, the linguistic realisations of other fillers are diverse. Besides, a number of them appear only once in the genres, for example, 'alright' is found once in telephone & conference call, airport service encounter, Q&A session, and 'but er' in meeting, informal office talk, Q&A session. Some even occurs once in one particular genre, for example, 'okay great' in meeting, 'alright then' in telephone & conference call, 'aghh' in informal office talk, 'la' in airport service encounter, 'mhm yeah' in hotel service encounter, 'but erm' in Q&A session, 'mhm mhm' in job interview, and 'er you know' in placement interview.

5. Discussion

The aim of this article is to investigate the use of fillers in a number of business settings in Hong Kong. In particular, it uses corpus linguistics as a perspective to examine how different realisations of fillers are pragmatically presented by themselves and together with other speech acts in a spontaneous spoken discourse. The examples in Section 4 show that fillers are not only diverse in terms of linguistic realisations (Table 1) but also associated with different speech acts in a speech event (Table 2). To further understand how a filler is realised with other speech acts in a communicative context, the most frequent filler 'er' from the genre of meeting is selected for discussion.

Extract 1 illustrates how a Hong Kong Chinese speaker uses a filler to plan and prepare what is to be said next. In the extract, speaker b6 is a hotel staff member. He is reporting the progress of having a visit to a company (lines 1, 3, 5) with 'er' as a filler in between (lines 2, 4). After mentioning that he was rejected for the visit, he made a boundary in a separate tone unit with 'so' (line 6). Then he went on informing the colleagues of his plan to visit the company the following week (lines 7, 9, 11) with fillers (lines 8, 10):

Extract 1: B017

Location: Hotel

Participant: b6: male Hong Kong Chinese

1. b6: <SA063 [statement: inform] in fact I've asked to make a >
2. <SA032 [filler] er >
3. <SA063 [statement: inform] personal visit to make the >

4. <SA032 [filler] er er >
5. <SA063 [statement: inform] decision make but I I was being rejected >
6. <SA033 [frame] so >
7. <SA063 [statement: inform] anyway I will have a visit >
8. <SA032 [filler] er >
9. <SA063 [statement: inform] on next Monday to see if their >
10. <SA032 [filler] er >
11. <SA063 [statement: inform] to see the booking handler KCRC >

It can be seen from Extract 1 that fillers as 'hesitations are fairly unpredictable' (Stenström, 1994, p.8). These linguistic features of delay (Tsui, 1994) can be found anywhere in the utterances for before the speakers have made up their minds (Clark & Fox Tree, 2002).

In Extract 2, b1 and b4 are hotel staff members. They are talking about the new door handle. After b1 points out the problem about the handles (lines 1-14), b4 responds with an 'er' (line 15) as a filler and a 'yeah' (line 16) as an acknowledgement. It is noted that the filler 'er' is frequently used by b1 (lines 2, 4, 6, 8, 10, 13):

Extract 2: B016

Location: Hotel

Participants: b1, b4: male Hong Kong Chinese

1. b1: <SA063 [statement: inform] the problem is you have touch up you have >
2. <SA032 [filler] er >
3. <SA063 [statement: inform] you have >
4. <SA032 [filler] er >
5. <SA063 [statement: inform] polish them we have touch >
6. <SA032 [filler] er >
7. <SA063 [statement: inform] touch up the >
8. <SA032 [filler] er >
9. <SA063 [statement: inform] the door but then you still have the >
10. <SA032 [filler] er >
11. <SA063 [statement: inform] old handle then >
12. <SA063 [statement: inform] stretch again the >
13. <SA032 [filler] er >
14. <SA063 [statement: inform] back door >
15. b4: <SA032 [filler] er >
16. <SA053 [reply to statement: acknowledge] yeah >

The filler 'er' in line 15 is used to illustrate the importance of understanding the context of the speech act association. Though 'er yeah' in meeting and 'er yes' in airport service encounter and placement interview are also found to be a filled pause with one occurrence, 'er' in line 15 and 'yeah' in line 16 are annotated with a different speech act with reference to a holistic consideration of both the contextual and the prosodic information given.

6. Conclusions

Fillers are frequently used in naturally occurring spoken interaction with different forms and functions for helping the speakers not only mark the boundaries between topics but also negotiate the way of thinking. As demonstrated in the analysis, to let the speaker hold the conversation or to give the speaker more time to consider what to say next is found to be a very common function of fillers. Speakers use fillers in utterances when they hesitate and think during the speech so as to keep the turn in a conversation. Also, the co-occurred speech acts with fillers typically do not display a particular co-occurrence pattern, but instead represent content- and context-specific realisations. The diverse co-occurrence patterns during the process of expressing ideas in speaking also reflect the common function of giving more time to speakers to organize their thought.

The analysis highlights that the use of real-world, naturally occurring data is salient for the investigation of different aspects of spoken business communication, including typical lexicogrammatical and rhetorical features of spoken business discourse. Different from those collected from a controlled setting or environment that have a relatively limited variety of patterns or formulaic sequence categories, fillers in the naturally occurring data collected in the business sub-corpus have displayed specific frequencies, co-occurring patterns, and linguistics features.

Throughout the analysis section, it has been demonstrated that speech act realisation is a complex linguistic phenomenon that comprises the unity of form, content, and function. A large variety of pragmatic functions can be realised by unmarked and marked linguistic expressions. The findings of the study show that these expressions need to be interpreted in the specific contexts of situation in which they occur. Though it would have further complicated the pictures of grammatical patterns and lexical choices for the realisation of speech acts, these linguistic features are notable as they bring to light a variety of patterns of linguistic realisation of speech acts which can help to enrich the overall understanding of the speech act performance in each genre of business communication in Hong Kong.

References

Austin, J. L. (1975). *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955*. Oxford: Oxford University Press.

Bhatia, V. K. (2017). Methodological issues in genre analysis. *HERMES - Journal of Language and Communication in Business, 9*(16), 39-59.

Brazil, D. (1997). *The Communicative Role of Intonation*. Cambridge: Cambridge University Press.

Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics, 11*(4), 411-433.

Cheng, W., Greaves, C., & Warren, M. (2008). *A Corpus-driven Study of Discourse Intonation*. Amsterdam: John Benjamins.

Cheng, W., & Warren, M. (1999). Facilitating a description of intercultural conversations: The Hong Kong Corpus of Conversational English. *ICAME Journal, 20*, 5-20.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*, 73-111.

De Felice, R. (2013). A corpus-based classification of commitments in business English. In Romero-Trillo, J. (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2013* (pp. 153-171). Dordrecht: Springer.

De Felice, R., Darby, J., Fisher, A., & Peplow, D. (2013). A classification scheme for annotating speech acts in a business email corpus. *ICAME Journal, 37*, 71-105.

Evans, S. (2010). Business as usual: The use of English in the professional world in Hong Kong. *English for Specific Purposes, 29*(3), 153-167.

Kallen, J., & Kirk, J. (2012). *SPICE-Ireland: A User's Guide*. Belfast: Cló Ollscoil na Banríona.

Kharismawan, P. Y. (2017). The types and the functions of the fillers used in Barack Obama's speeches. *International Journal of Humanity Studies, 1*(1), 111-119.

Koester, A. (2010). *Workplace Discourse*. London: Continuum.

Laserna, C. M., Seih, Y.-T., & Pennebaker, J. W. (2014). Um … Who like says you know: Filler word use as a function of age, gender, and personality. *Journal of Language and Social Psychology, 33*(3), 328-338.

Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.

Stenström, A.-B. (1994). *An Introduction to Spoken Interaction*. London: Longman.

Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Van Ess-dykema, C., & Meteer, M. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics, 26*(3), 339-373.

Tsui, A. B. M. (1994). *English Conversation*. Oxford: Oxford University Press.

Weisser, M. (2003). SPAACy - A semi-automated tool for annotating dialogue acts. *International Journal of Corpus Linguistics, 8*(1), 63-74.

THE AUTHOR

Andy Seto is a Part-time Lecturer at the Language Centre of the Hong Kong Institute of Vocational Education (Haking Wong). He studied and conducted research in the field of speech acts and corpus linguistics at Department of English, the Hong Kong Polytechnic University. He is the author of Speech Acts Annotation for Business Meetings (2013) and the co-author of 'Thanks a Bunch': Cross-cultural Comparison of the Speech Act of Thanking (2015). His research interests include Business Spoken Discourse, Corpus Linguistics, English for Specific Purposes (ESP), Pragmatics, Speech Acts.

THE AUTHOR'S ADDRESS

First and Corresponding Author
Andy Seto
Lecturer
Language Centre
The Hong Kong Institute of Vocational Education (Haking Wong)
702 Lai Chi Kok Road, Cheung Sha Wan, Kowloon, CHINA
E-mail: andyswh18@vtc.edu.hk

Received: 13 August 2020
Received in Revised Form: 30 May 2021
Accepted: 27 July 2021