CHAPTER 1

Introduction

In the course of its history, the English language has undergone an enormous expansion and has probably now reached the status of the world's leading language. It has become an important tool in international encounters, in both the digital and the analogue world, and it occupies a central role as an official or co-official language in many countries. One consequence of this unprecedented spread is the emergence of different Englishes all around the world, new varieties of the language, which are influenced by various factors, including language contact and change, national and international sociopolitical conditions, mechanisms of language acquisition, and speakers' attitudes and motivations. As a result, English has developed into a heterogeneous conglomerate of different national and international varieties – and this process is still ongoing. The number of English speakers is rising rapidly, particularly in Asian and African countries, where more and more people are acquiring English as a foreign or second language, both through formal education and in grassroots contexts (cf. Bolton & Graddol, 2012, on the spread of English in China; for an overview of research on grassroots Englishes cf. Meierkord, 2020). Moreover, English is also increasingly becoming the first language of many children whose parents are not necessarily native speakers of English themselves (cf. Buschfeld, 2020, on the situation in Singapore).

Research into 'World Englishes' – the cover term used for 'all or any of the [English] varieties spoken around the world' (Schneider, 2020, p. 31) – traditionally focused on modelling this variation and on describing the morphosyntactic and phonological structures of individual varieties, often with the help of large-scale electronic corpora. More recently, the field has opened up to include interdisciplinary research; for example, in areas such as construction grammar, computer-mediated communication, transnationalism, and variational pragmatics. Nevertheless, one central aspect has been neglected so far: World Englishes tends to regard conversational interaction mainly as a data source from which linguistic features – be they

1

contact-induced or linguistic universals – can be extracted. The conduct of conversation itself, that is, the effect which the linguistic makeup of individual varieties and the different cultural backgrounds of their speakers have on talk-in-interaction, has been largely ignored (cf. also Schneider, 2018, p. 97). In fact, the only studies explicitly dealing with conversational patterns in World Englishes seem to be four journal articles: Shields-Brodber's analysis of Jamaican radio talk shows (1992), Sidnell's papers on turn-taking and repair in Caribbean English Creoles (2001, 2008), and, more recently, Haselow's work on other-initiated repair (OIR) in varieties of English (2021).

This research gap is certainly surprising, especially because language and culture are inextricably linked with their use in social interaction (Watson, 1992, p. 2; cf. also Eglin, 2015). This relationship is essentially a dynamic one, as conversation both shapes and is shaped by its cultural and linguistic context. On the one hand, varieties of English are situated in culturally diverse contexts and have often developed nativised linguistic properties and structures. On the other hand, varieties are not entities *in vacuo*; they are used by specific speakers in specific contexts and specific constella-tions – with face-to-face interactions constituting the vast majority of settings. The fact that speakers are engaging in conversational interaction is therefore consequential for the speakers' use of language. In fact, as Couper-Kuhlen and Selting put it, '[t]he universals of language practice in talk-in-interaction can . . . be traced back to the interactional foundation of human society. In this very real sense then, *interaction shapes language*' (2018, p. 555, emphasis in original).

In this book, I analyse conversational patterns in Caribbean and Southeast Asian English face-to-face interactions. Specifically, I investigate whether variety-specific features and properties influence turn-taking in ordinary conversation. I focus on the question whether turn-taking conventions in Caribbean and Southeast Asian English inter-actions correspond to those that have been established by previous research on turn-taking, which has almost exclusively focused on British and American conversations, that is, traditional Inner Circle varieties, or on languages other than English. Furthermore, different varieties of English are also inseparably linked to different cultural backgrounds, which means that turn-taking patterns might be subject to cross-cultural variation. Indeed, this possibility has been emphasised early on; for example, by Moerman who stresses that conversations are essentially 'human events, events of meaning. Their description, explication, and analysis require a synthesis of ethnography – with its concern for context, meaning, history,

and intention – with the sometimes arid and always exacting techniques that conversation analyis offers for locating culture *in situ*' (1988, p. xi, emphasis in original). Similarly, D'souza understands 'the acceptable possibilities of behaviour within a particular culture' as '[t]he grammar of culture [which] affects and influences the use of language in very striking ways and [which] is in turn affected by language' (1988, p. 160). For conversational interaction this means that different realisations of turn-taking or speaker change might be due to cultural preferences rather than constitute deviations from a universal norm – yet, the latter has often been claimed in descriptions of interactions in non-Western languages and cultures (e.g. by Reisman, 1974, who describes Caribbean Creole English conversations as chaotic).

The primary focus of this study is on the set of strategies speakers from two culturally different groups have at their disposal in order to claim or hold a turn at talk. More specifically, I conduct an empirical study based on authetic audio data of Southeast Asian and Caribbean English conversations to investigate if the two speaker groups differ in their turn-taking routines. Using the theoretical and methodological framework developed by Conversation Analysis (CA), I transcribe and closely analyse five hours of unscripted face-to-face interactions taken from the Asian Corpus of English (ACE) and two Caribbean components of the International Corpus of English (ICE): ICE-Jamaica (ICE-JA) and ICE-Trinidad and Tobago (ICE-T&T). Based on the results from this first, qualitative, analysis, I then develop a formal coding system that allows me to conduct a second, quantitative, analysis of the data. I illustrate that World Englishes differ in the strategies they prefer to organise turn-taking and speaker change and show that these interactional preferences sometimes correlate with variety-specific patterns, such as a higher usage of topicalised constructions. Overall, four major research questions are addressed:

(1)   Does turn-taking in Southeast Asian and Caribbean English conversations generally follow the turn-taking framework described for Inner Circle Englishes?
(2)   What are the different forms, contexts, and frequencies of turn allocation in Southeast Asian and Caribbean English conversations?
(3)   Which strategies do Southeast Asian and Caribbean English speakers employ to claim or hold a turn at talk?
(4)   Do the findings from questions (1) to (3) correspond to previous descriptions of Southeast Asian and Caribbean English 'speaking styles'?

In Chapter 2, I introduce the theoretical framework for the study. As the analysis is located at the interface of comparative CA and World Englishes, I start by briefly sketching both research paradigms. I then address central epistemological differences between the two traditions and explain why and how they can be combined to provide new insights into the interactional patterns and variational development of Englishes around the world.

In Chapter 3, I start by describing the reasoning behind choosing the particular speaker groups under investigation in the present study. I explain how the notion of 'culture' is understood in this book and then introduce the data that form the basis for the analysis. As the corpora from which the data were extracted were originally compiled for studying the linguistic organisation of World Englishes but not necessarily their interactional patterns, I also comment on their suitability and potential shortcomings with respect to conversation analytic research. In a next step, I describe how I established a collection of unscripted natural conversations for the project and briefly comment on the transcription process involved. I illustrate how qualitative analysis can be successfully combined with subsequent quantification and explain why this is actually necessary in the field of comparative CA. The last part of the chapter provides a detailed description of the codification procedure and the formal coding system developed for the project, before summarising the steps involved in the quantitative section of the analysis.

In Chapter 4, I analyse the different types and scenarios of speaker change in Southeast Asian and Caribbean conversations. The three general types of turn allocation – next speaker selection, self-selection, and current speaker continuation (cf. Sacks et al., 1974) – are examined in great detail. I describe their concrete realisations in the data and illustrate nine major scenarios of speaker change and the specific sequential contexts they are situated in. The qualitative anlysis is then complemented by a quantitative comparision of the two speaker groups.

Chapter 5 provides an in-depth analysis of the resources Southeast Asian and Caribbean speakers of English use to claim or hold a turn at talk. Four larger strategy groups are described and compared: latches and overlaps, phonetic resources, lexical resources, and syntactic strategies. I investigate how these are realised by the individual speaker groups, compare my findings to previous research on Inner Circle Englishes, and identify techniques that seem to be variety-specific or culturally sensitive.

Chapter 6 complements Chapter 5 from a quantitative perspective. As conversationalists rarely rely on one type of resource when it comes to claim or hold a turn, I identify and compare strategy clusters used by the

interactants. In a second step, I focus on specific realisations of turn-taking strategies (e.g. direct address, volume upsteps, slowing down) and compare their usage across the two speaker groups.

Chapter 7 builds on the quantitative and qualitative analyses presented in Chapters 4–6 and answers the question why conversational interaction has been described as orderly and supportive in some contexts but is perceived as chaotic or interruptive in others. I start by scrutinising the data for signs of 'interruptiveness' and show that the concept is often confused with a preference for more direct turn-taking strategies. Based on this observation, I conclude the chapter with a discussion of the apparent dichotomy of cooperation and competition and suggest regarding turn-taking as an instance of coopetition instead.

Finally, Chapter 8 sums up the main findings of the study, answers the research questions, and gives an outlook on potential further research at the interface of CA and World Englishes.

CHAPTER 2

Investigating Talk-in-Interaction in Culture

This chapter introduces the theoretical framework for the subsequent analysis. As mentioned earlier, the analysis is situated at the interface of two scientific traditions that have not had much contact before: CA and World Englishes. Hence, to cater for the needs of scholars from both areas, a concise overview of each field's central tenets is necessary. Section 2.1 starts by outlining central theoretical concepts of CA; Section 2.2 briefly introduces the field of World Englishes. Finally, Section 2.3 addresses the theoretical and methodological differences between both paradigms before providing a rationale for why (and how) CA and World Englishes can be reconciled in a fruitful way.

2.1   What Is Conversation Analysis?

Conversation Analysis (CA) is a scientific approach concerned with the study of talk-in-interaction. Its primary interest is not investigating language per se; rather, the focus is on social interaction and its underlying organisation. This is hardly surprising – CA originally emerged from sociology and its sociological roots are still apparent, even though it can only be described as a highly interdisciplinary research field in its current form (Maynard, 2013, p. 13). Nevertheless, language holds a special position in the conversation analytic framework. A large amount of social action is carried out through talk, making interaction the 'basic and primordial environment' (Schegloff, 1996, p. 54) for language use in our everyday lives (Clift et al., 2009, p. 40; Sidnell, 2016). Quite naturally, scholars interested in language used in conversational and institutional interaction have therefore resorted to findings and methods from CA, which eventually led to the formation of a linguistically focused branch of the field, Interactional Linguistics (Couper-Kuhlen & Selting, 2018, p. 12).

One of CA's main tenets is the idea of 'order at all points' (Sacks, 1984, p. 22). Every detail of an interaction is treated as a potentially orderly

phenomenon and, thus, as worthy of investigation – no matter how small and seemingly mundane. Conversation analytic research has set out to identify this intricate set of rules underlying talk-in-interaction, starting with the organisation of turn-taking in conversation, which Schegloff describes as '[o]ne of the most fundamental organizations of practice for talk-in-interaction' (2007, p. 1). In everyday conversation, who speaks next and for how long is not fixed in advance. Rather, speakership or turn size is negotiated on a turn-by-turn basis as the interaction develops; it is locally and interactionally managed. In other words, the authority over turn-taking lies with the conversationalists themselves, making the system a party-administered one (Sacks et al., 1974, p. 725f). Despite this, everyday conversations are by no means chaotic and speakers do not start up randomly – on the contrary, there seems to be a tendency to limit both the number of overlaps and the periods of silence between speakers, and there is evidence that these preferences are universal (Stivers et al., 2009). In their seminal paper, Sacks et al. (1974) show that this close fine-tuning is possible, because turn-taking is a two-part mechanism, which consists of a turn-constructional and a turn-allocation component. The former deals with the construction of the 'smallest interactionally relevant complete linguistic units in their given context' (Selting, 2000, p. 512), the so-called turn-constructional units (TCUs). These can be conceptualised as gestalt-like but flexible schemata 'on which participants rely for their orientation in constructing and interpreting units; for example, the schema of a "possible sentence," a "possible clause," a "possible phrase," or a particular kind of "intonation contour" with a "possible unit or turn-ending pitch movement' (Selting, 2000, p. 492). That is, TCUs are not merely defined based on syntactic completion but constitute an 'interplay of syntactic, lexico-semantic, pragmatic, activity-type-specific, and prosodic devices in their sequential context' (Selting, 2000, p. 487). To date, most of these devices have been studied extensively, particularly the role of syntax and grammar (e.g. Schegloff, 1996; Lerner, 1996, 2004a) and prosody (Local et al., 1986, 1985; Couper-Kuhlen & Selting, 1996; Wells & Macfarlane, 1998; Local & Walker, 2012), revealing a finely tuned interrelatedness between the different elements (Selting, 2000, p. 491; cf. also Oreström, 1983; Tanaka, 1999). TCUs might hence best be described as 'multimodal packages … that make use of a range of different modalities' (Hayashi, 2005, p. 47), including verbal but also non-verbal resources. Still, to date the interplay of the individual devices remains an understudied aspect of the turn-construction component (Li, 2014, p. 6), particularly for languages other than English and for Outer and Expanding Circle Englishes.

The essential role TCUs fulfil in the turn-taking mechanism is
grounded in their projectability. As TCUs constitute logical units of talk,
they enable potential next speakers to project their first possible comple-
tion point, that is, the place where speaker changes becomes relevant (the
so-called transition-relevance place (TRP)). The notion of projectability is
central for the smooth progress of everyday interactions: interactants do
not merely wait for and react to signals indicating turn-completion (as
claimed, e.g., by Duncan, 1972; Duncan & Fiske, 1977) but are able to
predict upcoming TRPs before their actual occurrence. In consequence,
gaps and overlaps are typically mimimised in ordinary conversations –
participants monitor the TCU underway and project transfer from one
speaker to the next accurately enough to avoid starting up too early or too
late. Latches, that is, speaker changes in the moment of the last sounds of
an utterance, probably demonstrate this most impressively.

As soon as a TRP is reached, that is, a TCU has been completed by the
current speaker, the allocational component of the turn-taking mechanism
becomes relevant. This second component can be split up into three
hierarchically ordered rules (Sacks et al., 1974, p. 704):

(a)   The current speaker might have selected a next speaker, which means
      that speakership is transferred to this (and no other) interactant.
(b)   If no next speaker has been selected, any participant can self-select.
      In this case, speakership is transferred on a 'first come, first served'
      basis.
(c)   If no next speaker has been selected and none of the participants self-
      selects, the current speaker can choose to continue talking.

As turn-taking is locally managed, these options apply anew at each TRP.
Ordinary conversation is thus *inter*-action in the literal sense of the word:
Participants are actively negotiating turn allocation (Sacks et al., 1974,
p. 726; Sidnell, 2016) – either as '[c]o-participants [who] will properly be
oriented to possible completions as places where they may have rights
or obligations to talk, [or as] . . . speakers [who] accordingly will be
oriented to them as resources for drawing others in and exiting the turn
themselves, or holding others off so as to extend what is being said'
(Schegloff, 1996, p. 82).

With both the turn-constructional and the turn-allocational component
being directed towards them, TRPs can rightfully be described as the
'places of action' in an interaction (cf. Sidnell, 2010, p. 47). In the turn-
taking mechanism, speakers can only acquire the right for one TCU and
transfer becomes relevant again as soon as it is completed or its completion

is projectable. In other words, speaker change constitutes the default case
in ordinary conversation, and speakers who want to produce multi-unit
turns have to invest extra work to avoid other participants claiming the
floor (Schegloff, 1996, p. 61). Turns that include more than one TCU are
therefore always 'marked' (Selting, 2000, p. 510) to some extent: speakers
can, for instance, resort to using turn-holding strategies, thus reshaping the
TRP 'in such a way as to reduce the likelihood of turn transfer and increase
the likelihood of the current speaker extending his/her turn' (Clayman,
2013, p. 159). Multi-TCU turns are thus typically characterised by active
and observable manipulations of an upcoming TRP ('action formation'
(Levinson, 2013, p. 110)), which allow prospective next-speakers to draw
inferences about the speaker's probable intended action ('action ascription'
(Levinson, 2013, p. 104)). On the other hand, prospective next speakers
can also employ various interactional resources to strengthen their claim
for speakership at an upcoming TRP. As with turn-holding, these turn-
claiming strategies are oriented towards the TRP as their central point of
reference, that is, they typically cluster around the end of one and the
beginning of another TCU. According to Schegloff:

> the endings of TCUs live under the shadow of the incipient beginnings of
> next turns, and . . . [the] beginnings of turns can be thoroughly preoccupied
> with the ends of their preceeding turns. . . . [T]aken together, the two sets
> of practices – of turn and TCU beginnings and turn and TCU endings [*sic*]
> constitute the major factors shaping *the social and interactional organization
> of the transition space*. (1996, p. 96, emphasis in original)

There is some evidence that turn-holding and turn-claiming strategies
differ across languages and possibly even dialects. The typology of a
language will, for instance, affect the extent to which speakers can rely
on syntax as a clue for TCU projection. Languages in which core elements
of the clause are mentioned early on allow prospective next speakers to
project upcoming TRPs at a very early stage, because the trajectory of the
TCU underway is already determined. Accordingly, these languages have
been labelled 'early projection' languages in the literature. British/
American English or German constitute prototypical examples of typolo-
gies that enable early projection (Egbert, 1996; Ford et al., 1996). They
differ from languages such as Japanese, which have a more flexible word
order and make TCU projection difficult; for example, because the verb is
mentioned late in the clause, or because contextually inferable syntactic
elements can be omitted (Hayashi, 2004, p. 1344). In consequence,
'crucial information concerning the shape of turn being produced tends

to be concentrated towards the end of a turn. These features can make it difficult for participants in Japanese to project a possible completion point or the type of activity which will be performed by a turn until slightly before the end of a turn' (Tanaka, 1999, p. 143; cf. also Ford et al., 1996, p. 449). In 'delayed projection' languages such as Japanese, non-syntactic elements such as prosody or particles play a greater role for projecting upcoming TRPs (e.g. Tanaka, 1999; Iwasaki, 2009).

Still, research on turn-holding and turn-claiming so far has mainly focused on a handful of languages and many aspects are still unexplored. In particular, it is not clear at all whether varieties of English can automatically be classified as early projection languages and whether different speaker groups rely on the same set of turn-taking strategies, or if they do so to the same extent. Apart from that, the relationship of conversational interaction and culture is still under-researched, which is probably due to the fuzziness of the notion of 'culture' as such. Of course, there are anthropological reports of cultural differences in conversational style, such as a greater tolerance of long silences in Finnish or Australian Aboriginal interactions (Lehtonen & Sajavaara, 1985; Eades, 2007) or a larger amount of (multi-speaker) overlaps in public debates of the Xavante, a Brazilian tribe (Graham, 1995; see also Meyer, 2018, pp. 108–122 for more examples). However, despite these findings, the question whether the turn-taking system might be culturally sensitive is still unanswered. On the one hand, Sacks et al. have always conceptualised their turn-taking framework as 'context-free' (1974, p. 699), that is, the scaffolding of natural conversation was thought to be independent from external parameters such as location, time, and speakers' social identities (Sacks et al., 1974, p. 699, fn. 8). In other words, turn-taking is regarded as a 'candidate universal' (Schegloff, 2006, p. 83; cf. also Sidnell, 2006, p. 171), a view that is supported by a number of studies showing that Sacks et al.'s model not only holds for British and American English interactions but also for other languages (such as Thai-Lue (Moerman, 1977), Japanese (Tanaka, 1999), German (Selting, 2000), or Mandarin (Li, 2014)). However, none of these studies looked at cultural preferences with respect to turn-holding or turn-claiming, even though the possibility that some aspects of turn-taking might be shaped by the sociocultural context has been acknowledged: With respect to Japanese conversations, Tanaka explicitly mentions 'local socio-cultural orientations' (1999, p. 32) as a potential factor influencing the speakers' choice of turn-holding or turn-claiming resources but does not analyse them in greater detail. Ochs (1984) explicitly describes a link between specific repair strategies and the speakers' cultural