

English in Taiwan

SOFIA RÜDIGER , JAKOB R. E. LEIMGRUBER  AND
MING-I LYDIA TSENG 

Expanding the scope of corpus-based research on East
Asian Englishes

Introduction

While World Englishes scholarship has always been concerned with different types of English varieties, Expanding Circle (i.e., non-postcolonial) Englishes have had a ‘late start’ in being added to its research remit. As a result, much important work in this area remains to be done. Expanding Circle Englishes in general and Asian Expanding Circle Englishes in particular are still neglected in many handbooks of World Englishes (e.g., in *The Cambridge Handbook of World Englishes*; Schreier, Hundt & Schneider, 2020). Notable exceptions here are, for example, *The Routledge Handbook of World Englishes* (Kirkpatrick, 2020; including, among others, chapters on Japanese, Chinese, and Slavic Englishes) and *The Handbook of Asian Englishes* (Bolton, Botha & Kirkpatrick, 2020; including, among others, chapters on Taiwanese, Cambodian, and Indonesian Englishes). While traditionally much focus has been laid on matters of language policies, education, and attitudes, corpus linguistic approaches to Expanding Circle Englishes have become more and more relevant (see, e.g., Edwards, 2016 for the Netherlands; Rüdiger, 2019 for South Korea). In this article, we present the first results from a corpus-based study of Taiwanese English, drawing on the pilot version of a spoken Taiwanese English corpus.

Taiwan is located in East Asia and has been described as ‘an Expanding Circle society with ambitions to move into the Outer Circle’ (Kobayashi, 2020: 553). In this paper, we will compare our results on Taiwanese English to another East Asian English variety, that is, South Korean English. South Korea constitutes a great point of comparison in this regard as we find a number of similarities between both regional

contexts (beyond being located in the same broader geographical region). Both countries, for example, have been described as exhibiting an explicit orientation towards the United States (see Seilhamer, 2019: 188 for Taiwan; Grant & Lee, 2009 for South Korea) and have a keen demand for private English education (epitomized in a flourishing business of private English learning institutions, known in Taiwan as *buxiban* and in South Korea as *hagwon*). Furthermore, even though learned as an additional language in both contexts, English plays an important role in the form of social, cultural, symbolic, and economic capital in Taiwan and South Korea (see, e.g., Seilhamer, 2019: 173 for Taiwan; Park & Abelman, 2004; Park 2011 for South Korea). Last but not least, in both cases, different stakeholders have proposed the adoption of English as an official language (see Chen et al., 2018 for Taiwan; Yoo, 2005 for South Korea).¹

In this paper, we first give the relevant background information on the Taiwanese context, with a particular focus on language education policies. Next we introduce the pilot version of TASE, the Taiwanese Spoken English corpus. Using keyword analysis to compare TASE to the Spoken Korean English corpus (SPOKE), we discuss some first analytical starting points, before investigating two morpho-syntactic patterns, that is, plural marking on the noun and the general use of pronouns, in more detail.

English in Taiwan

The island of Taiwan, with a population of 23.5 million (National Statistics, 2021), is located to the southeast of mainland China across the

Taiwan Strait. Kobayashi (2020: 548) points out that Taiwan should be considered a multilingual society, as a number of languages are in use: various Formosan languages, Mandarin, Hokkien, Hakka, and Taiwan Sign Language (see Tai & Tsay, 2015). Due to the dominance of Mandarin Chinese, which is also the de facto official language, Taiwan has, however, been described as a ‘pseudo-monolingual’ society (Go, 2018, quoted

in Kobayashi, 2020: 548). Language education policy in Taiwan, particularly English education, consequently has been shaped by nationalism, modernization and economic growth, as well as globalization (cf. Tsao, 2000). The following sections present a brief description of language education policies in relation to the history, and the political and economic development in Taiwan before turning to a discussion of English education in the last two decades, especially Taiwan planning to implement bilingual education by 2030.



SOFIA RÜDIGER is a postdoctoral research associate at the University of Bayreuth, Germany, with a research focus on World Englishes, corpus linguistics, digital communication, culinary linguistics, and (historical) pragmatics. She is author of Morpho-Syntactic Patterns in Spoken Korean English (Varieties of English Around the World, 62, John Benjamins, 2019) and co-editor of Discourse Markers and World Englishes (special issue of World Englishes, 40/4, 2021). Email: sofia.ruediger@uni-bayreuth.de



JAKOB R. E. LEIMGRUBER is currently lecturer at the University of Basel (Switzerland) and substitute professor at the University of Freiburg (Germany). His research foci include World Englishes, language planning and policy, and linguistic landscapes. He is particularly interested in the English found in Singapore, Taiwan, and the United Arab Emirates. He is the author of Singapore English (CUP, 2013) and Language Planning and Policy in Quebec (Narr, 2019), and co-editor of Multilingual Global Cities (with Peter Siemund, Routledge, 2021). Email: jakob.leimgruber@unibas.ch



DR. MING-I LYDIA TSENG is an associate professor in the Department of English Language and Literature at Fu Jen Catholic University, Republic of China (Taiwan). Her research interests include academic literacy studies, second/foreign language learning and teaching, qualitative research methodology, and discourse analysis (particularly, critical discourse analysis and genre analysis). Email: 023148@mail.fju.edu.tw

Earlier language education policies in Taiwan: From the 17th century to 2000

Up to the 17th century, there are no written historical records of languages used in Taiwan. In the following 400 years, four major political regimes ruled Taiwan, all of which had an effect on language education; these four regimes include European colonization, the Qing Dynasty, Japanese colonization, and the Kuomintang (KMT) Nationalist government (Wu & Lau, 2019). European colonization, namely, the Dutch (1624–1662) and the Spanish (1626–1642), had limited impact on language policies in Taiwan. Religion, more specifically, the Dutch’s intent to convert the aboriginal people to Christianity, resulted in the creation of *Sinkang*, a Romanized written form for the language spoken by the aboriginal tribe of Siraya. Second, the Qing government extended its language policy of teaching Chinese as a lingua franca to the aboriginal population and to other residents who spoke local Taiwanese Hokkien or Hakka in their daily life. The Qing governance of Taiwan ended in 1895 after it lost the First Sino-Japanese War (1894–1895), which marked the beginning of a half-century of Japanese colonization. The Japanese regime’s ambitious scheme of Japanization of Taiwan became evident in its prohibition of speaking local vernaculars, including Hokkien and Hakka, in all private domains and the teaching of Japanese as the colonizer’s language in school. During the Japanese colonization, Japanese with its vocabulary and syntactic structure mixed with some forms of Dutch and aboriginal languages resulted in linguistic hybridity and creativity in Hokkien (Simpson, 2007). It is only with the end of the Japanese colonization of Taiwan after the Second World War and the Kuomintang Nationalist government (KMT) regime arriving and ruling Taiwan till 2000 that language education policy, particularly English education, was motivated by ‘status planning’ of Mandarin Chinese (Cooper, 1989) and the driving force to economic prosperity (Chen, S.-C., 2006). Before 1987,

Mandarin Chinese was privileged as a nationalist language over other local dialects (Tsao, 2000). After the lifting of martial law in 1987, the shift of political power from the KMT to the Democratic Progressive Party (DPP) in 2000 and the need to open the market internationally resulted in an open-minded attitude toward diversity in using languages. This most prominently affected language education policy in two ways: first, raising people's awareness of speaking local vernaculars for revealing their 'Taiwanese' identity; and second, favoring English as a dominant foreign language for promoting globalization (Wu & Lau, 2019).

Current language education policies in Taiwan: Localization and globalization

After the DPP gained political power, localization or so-called Taiwanization (Wu, 2011) reversed language education policy by the implementation of 'Local-Language-in-Education' (LLE), which refers to the inclusion of teaching local vernaculars, especially Taiwanese Hokkien, in primary education and reducing the hours of teaching Mandarin Chinese as the national language. According to the Ministry of Education (MOE) 2001 curriculum guidelines, local vernaculars as mother tongues are more than means of daily communication: they are an embodiment of Taiwanese people's cultural identities. In addition to localization, the interplay of economic and political development in tandem with globalization further reshape the language ecology and education landscape of Taiwan, determining English language education. The implementation of English language education in this context is to facilitate the 'simultaneous promotion of internationalization' in response to 'social change and national goals' (Chen, S.-C., 2006: 322).

English Education Policy and the 2030 Bilingual Nation Development Goal

More recently, the English Education policy (EE) has been enforced to embrace globalization, particularly in three aspects: (1) English being introduced as a compulsory subject in primary education in 2001 with two class periods of teaching per week, in most counties starting from Year 3, and in some cities, including the capital, Taipei, starting from Year 1, (2) an English exit requirement for college/university students to pass certain threshold scores before graduation, (3) English as a Medium of Instruction courses being promoted in higher education, with higher payment for university teachers (Tsou & Kao, 2017; Wu & Lau, 2019). On September 19th, 2018, then Premier Lai announced that for coping with globalization,

Taiwan has to boost people's English proficiency and develop into a bilingual nation by 2030. Since then, a blueprint for this plan has been published by the National Development Council. The schemes introduced for promoting the bilingualization of Taiwan's educational system include: extending bilingual education to pre-school caretaking activities in the kindergarten curriculum, requiring a particular number of obligatory English-medium courses as integral to higher education, establishing all-English television channels, increasing English broadcasting programs, as well as cultivating friendly bilingual tourism environments (National Development Council, Taiwan, 2018).

While the government makes every endeavor to enhance international competitiveness through the EE policy and bilingual education development goals, English teaching and learning has long been part of Taiwanese people's daily life. Before the EE policy of requiring English as a compulsory subject in primary education, thousands of school pupils had already started learning English in pre-schools and the majority of people have been using some English at work or in daily life (Chen & Tsai, 2012). The ideologies of English are shaped by two driving forces: (1) the goals of the English curriculum in K-12 since 2001, particularly developing English communicative skills, fostering learning motivation, and promoting foreign cultures (Chien, 2014); (2) an indispensable role of supplementary education, best known as cram schools or *buxibans* (Chou & Yuan, 2011). The latter does not come as a surprise, as cram schools have been prevalent in many East Asian countries such as China, Japan, and South Korea (Hu & McKay, 2012). Liu (2012) has shown that studying at cram school appeared to have a positive correlation with Taiwanese students' academic performance, yet scant attention has been paid to the overall impact of secondary English education in both formal and cram schools upon Taiwanese young learners' English competence (Chou, 2015).

The latest report by the National Development Council in Taiwan (2018) points out that the goal of developing a '2030 bilingual nation' has evoked enthusiasm in learning, teaching, and using English in different domains. Nevertheless, the emergence of a 'bilingual nation discourse' may not only affect the education in Taiwan but also national identity. Many challenges of implementing bilingual education remain to be resolved: the lack of English teachers instructing non-English subjects, missing resources and infrastructure, and the thorny issue of using Mandarin Chinese as the next promising global language, a powerful

regional language to demonstrate one's national identity (Ferrer & Lin, 2021). English education in Taiwan in the coming ten years may undergo a rapid transformation; the linguistic landscape of Taiwanese English as an East Asian English variety may become more dynamic and pluralistic, being spoken and utilized in various contexts of communication.

Outside of the education sector, English features in Taiwanese lives in various ways, for example, in advertising (Chen, C. W.-Y. 2006) and different media (such as the English-language newspaper *Taipei Times* and movies shown in cinemas in their original language with subtitles). We also find substantial uses of English in some workplaces, such as the medical sector, where diagnoses and communication between professionals often take place in English (Bosher & Stocker, 2015). In addition, Go (2018; quoted in Kobayashi, 2020: 548) reported uses of English by children with their parents and peers. Nevertheless, English in Taiwan remains largely underexplored, which provides an ideal backdrop for our study. Apart from studies on language planning and policies (e.g., Simpson, 2007; Price, 2014), research output on English in Taiwan remains scant, is usually anecdotal, and, additionally, often already outdated (Hsu, 1994; Chen, C. W.-Y. 2006). In one of the very few recent publications on English in Taiwan, Seilhamer (2015: 376) describes the English use by his six female participants as 'exud[ing] confidence, active agency, and indeed a sense of ownership [of English]'. This indicates that English is indeed actively employed by (at least) parts of the Taiwanese society, despite English having no official status within the country (see also Seilhamer, 2019).

Data

As we outlined before, previous research on English in Taiwan has largely focused on matters of language learning and teaching as well as language policy. To the best of our knowledge, no spoken corpus of English by Taiwanese speakers exists to date. In this article, we report the first

results from a pilot corpus project with the title 'English in Taiwan – Forms and Functions'. The project ultimately aims at compiling a corpus of English in Taiwan comparable to the SPOKE corpus (a 300,000-word corpus of Spoken Korean English; see Rüdiger, 2019). In the first stage of the project, a pilot corpus consisting of 19 interviews with 21 speakers² was collected. The interviews were conducted in October 2019 by the first author of this paper and followed the 'cuppa coffee' framework (Rüdiger, 2016), which ensured a relaxed atmosphere conducive to informal conversations. Participants were recruited with the help of academic staff at a major university and personal contacts, with subsequent snowballing to recruit additional speakers. All participants signed an informed consent sheet before the recording started. The overall recording time lies at a bit more than 13 hours (i.e., 795 minutes). In three cases, two speakers participated in the recording at the same time (triadic conversations), the rest of the interviews involved one Taiwanese speaker and the interviewer (dyadic conversations). The length of individual recording sessions ranged between 22 and 64 minutes (with the triadic conversations being on average longer than the dyadic ones). Most of the conversations were recorded in different cafés on or close to a major university campus in New Taipei City. In the following, we refer to this corpus as TASE (Taiwanese Spoken English corpus). An overview of the demographics of the corpus speakers is given in Table 1.

After basic orthographic transcription, the pilot corpus spans ~76,000 words produced by the Taiwanese speakers (the whole corpus, i.e., including interviewer speech and the international student [see footnote 2], amounts to 133,000 words). The corpus was subsequently tagged for two morpho-syntactic phenomena for which we also have results from the SPOKE corpus, that is, plural marking on the noun and the use of pronouns. This enabled us to draw on previously established tagging procedures and allows a comparison between the results for both corpora.

Table 1: Speaker demographics TASE

Sex	Age	Profession	Nationality	Stay abroad (English-speaking country)
female (n = 13)	ranging from 18-27 average: 20	students (n = 19) ³ other (n = 2)	Taiwanese (all) ⁴	no stay abroad (n = 11) 0-6 months (n = 8) 7-24 months (n = 1) more than 24 months (n = 1)
male (n = 8)				

Comparing SPOKE and TASE

Using Antconc's (Anthony, 2018) keyword analysis tool to compare the pilot version of TASE with SPOKE, we find 54 keyword types in SPOKE (top ten reproduced in Table 2) and 104 keyword types in TASE (top ten reproduced in Table 3). It is of course not surprising that words like *Korean*, *Korea*, and *Seoul* are keywords in SPOKE and *Taiwan* and *Taiwanese* in TASE. Some keywords seem to stem from different thematic choices (*beer* in SPOKE and *mom* in TASE), even though both corpora were collected by the same interviewer, applying the same data collection and interviewing method. However, some keywords could be indicative of more fundamental differences between the two English varieties. In the following, we want to mention only three of the observations that stand out in this preliminary inquiry, all of which would warrant further in-depth quantitative and qualitative analysis.

First, SPOKE, but not TASE, has a surprising number of keywords which potentially function as backchannel responses (cf. Jefferson, 1984; Peters & Wong, 2014) such as *mhm*, *ah*, *right*, and *yeah*. This might be an indication of different discourse-structuring strategies by South Korean and Taiwanese speakers of English; a preliminary observation definitely worthy of future research.

Second, Rüdiger (2021), based on the SPOKE corpus, has shown how *like* is a firm part of the South Korean English repertoire and is used across the item's functional range (including its discourse-pragmatic functions). Despite the Korean speakers' attested use of this lexical item in its various functions, *like* shows up as keyword for TASE (not

SPOKE), on rank 3 nonetheless – definitely an invitation to have a closer look at both corpora to find out what Taiwanese speakers are doing differently.

Last but not least, we also find some function words as keywords in TASE, for example, the indefinite article *a* (rank 7). Indeed, an analysis of SPOKE has shown a low rate of indefinite article use (when compared to corpora of spoken American English, British English, and various ICE-corpora; cf. Rüdiger, 2019: 114–115). At first glance, Taiwanese speakers of English do not seem to share this characteristic. Other function words which are key in TASE are the multifunctional *to* (infinitive marker, preposition; rank 9) and the 1st person plural pronoun *we* (rank 10).

We hope to have shown here how productive this kind of comparison can be, even in its preliminary form and despite the current limitations at play (i.e., TASE being in its pilot stage). This provides ample pointers for the kind of research which we want to take up once the full corpus has been compiled. In the following two sections, we now turn to the results of the manual coding of plural markers on the noun and overall pronoun usage.

Plural marking on the noun

The primary language spoken by our informants, Mandarin Chinese, has practically no inflectional plural marking on nouns. Plurality is typically achieved by means of quantifiers, numerals, and context. The only exception is the suffix *-men*, which can attach to some human nouns, so that *lāoshī* 'teacher' can become *lāoshīmen* 'teachers', for instance in vocative use. Plural pronouns (see next section) also end in *-men*. However, *-men*

Table 2: Top ten keywords in SPOKE (when compared to TASE; keyword analysis with Antcon; effect size measurement: difference coefficient)

Rank	Frequency	Keyness	Effect	Keyword
1	4,672	+574.68	0.4829	<i>mhm</i>
2	2,455	+541.21	0.687	<i>ah</i>
3	772	+245.33	0.8495	<i>Korean</i>
4	8,216	+241.17	0.2101	<i>uh</i>
5	721	+199.42	0.786	<i>Korea</i>
6	1,036	+196.65	0.6295	<i>right</i>
7	13,491	+103.61	0.1008	<i>yeah</i>
8	2,809	+90.76	0.2236	<i>was</i>
9	145	+65.82	1	<i>beer</i>
10	161	+57.88	0.907	<i>Seoul</i>

Table 3: Top ten keywords in TASE (when compared to SPOKE; keyword analysis with Antconc; effect size measurement: difference coefficient)

Rank	Frequency	Keyness	Effect	Keyword
1	192	+502.5	0.9559	<i>Taiwan</i>
2	1,061	+399.3	0.3806	<i>uhm</i>
3	1,927	+344.62	0.2544	<i>like</i>
4	39	+124.39	1	<i>Taiwanese</i>
5	55	+124.16	0.9115	<i>mom</i>
6	271	+115.09	0.4074	<i>will</i>
7	1,440	+113.44	0.1651	<i>a</i>
8	112	+109.42	0.2243	<i>because</i>
9	1,944	+99.43	0.1313	<i>to</i>
10	849	+99.11	0.2039	<i>we</i>

cannot be used for inanimates (*shū* ‘book’ cannot be pluralized to **shūmen* ‘books’) and cannot co-occur with numerals (**sān-ge lǎoshīmen* ‘three teachers’) or quantifiers (**hěn duō de lǎoshīmen* ‘many teachers’), thus entirely eschewing the obligatory plural redundancy of other standardized Englishes. This makes morphological plural marking a marginal phenomenon in Chinese.

We focus here on cases of plural redundancy reduction, that is, cases where no plural marking (i.e., minus-plural marking; cf. Rüdiger, 2019: 47–48 for details on our terminological choice here) is found on a noun as the plural is already semantically entailed, either via specific lexical triggers preceding the noun (quantifiers like *many*, *several*, and *all*; numbers above *one*) or the discursive context (e.g., interviewer speech, common sense). For quantifiers and numerals, all instances were automatically retrieved from TASE via AntConc. The concordance lines were then manually examined to exclude irrelevant hits (e.g., lexical trigger not part of a noun phrase, unclear cases, non-count head of the noun phrase). In general, we applied the same annotation criteria as used for the analysis of plural marking in SPOKE to the TASE material (Rüdiger, 2019: 77). This allows a comparison between the two corpora, but we need to keep in mind that TASE is still in its pilot stage. The numerical results thus have to be interpreted with caution, particularly in cases where the overall number of instances is very low.

Quantifiers

Altogether, we identified 84 cases of minus-plural marking on the noun after quantifiers in TASE (see Table 4). This involves instances like (1)–(3):

- (1) I need to do *many thing* (TASE-005)
- (2) I feel like *one of the reason* I like the movie is well because it’s unreal (TASE-012)
- (3) so we are (.) required to take *these class* of course (TASE-003)

The aggregated reduction rate for TASE across all examined quantifiers lies at 21% (based on 323 realized plural forms and 84 minus-plural markings after quantifiers).

Numerals

Minus-plural marking on the noun after numerals (above *one*) also does occur in TASE (see examples [4]–[6]) but is rather infrequent. The aggregated reduction rate after numerals in TASE lies at 7% (based on 222 realized plural forms and 17 minus-plural markings after numerals).

- (4) I go there about like *two time* per week (TASE-002)
- (5) so right we have like *three convenient store* per block (TASE-006)
- (6) we have *fifty-seven classmate* (TASE-018)

A similar phenomenon occurred in SPOKE, where the redundancy reduction rate is also lower after numerals than after quantifiers (SPOKE: 29% → 15%; TASE: 21% → 7%).

Discursive context

The TASE corpus material additionally contains 1,036 plural nouns and 146 nouns with minus-plural marking (not preceded by a quantifier or numeral). In those cases, the plurality of the noun was already determined by the discursive context (uttered by the Taiwanese speaker or the interviewer; see examples [7]–[10]); for instance, in

Table 4: Overview plural marking after quantifiers (TASE) and % plural redundancy reduction (TASE and SPOKE) (round brackets indicate that the percentage is based on less than 10 overall occurrences)

	Plural (raw frequency) (TASE)	Reduced (raw frequency) (TASE)	% plural redundancy reduction (TASE)	% plural redundancy reduction (SPOKE)
<i>one of</i>	15	18	55%	53%
<i>these</i>	5	5	50%	28%
<i>few</i>	3	2	(40%)	8%
<i>those</i>	8	4	33%	15%
<i>many</i>	40	11	22%	30%
<i>all</i>	33	8	20%	27%
<i>some</i>	65	14	18%	22%
<i>other</i>	40	8	17%	32%
<i>more</i>	17	3	15%	18%
<i>lot(s) of</i>	65	8	11%	36%
<i>most</i>	25	3	11%	38%
<i>both</i>	4	0	(0%)	0%
<i>several</i>	3	0	(0%)	12%
Σ	323	84	21%	29%

(7), the surrounding context clarifies that the speaker here talks about several cases of tasks being assigned spontaneously. In (9), the context clarifies that the speaker describes how infants in general produce sounds (and not *a* specific sound).

- (7) but he uh he gives us **assignment** like randomly like (TASE-003)
- (8) so we are facing like an overthrow [overhang] of **teacher** (.) and excess of **teacher** whereas we are lacking people in other fields (TASE-001)
- (9) so they pronounce their **sound** with the **lip** (TASE-019)

In some cases, common sense led to a minus-plural marking reading of the data. In (10), the speaker describes her general movie preferences. It is rather unlikely that she likes to watch the same romantic movie over and over again and thus the concept referenced here is plural.

- (10) I like romantic **movie** (TASE-012)

Based on the numbers given above, the plural redundancy reduction rate for nouns not preceded

by quantifier or numeral lies at 13% (cf. SPOKE 21%).

In 88 instances, the noun did not bear plural marking, but it could not be determined with certainty whether the noun refers to a plural or singular referent. An example for this can be found in (11): it remains unclear, and indeed ultimately irrelevant, whether the speaker's boss and one colleague or several colleagues are travelling together.

- (11) like my boss and my **colleague** sometimes they need to fly to uh Thailand or Vietnam or other country (TASE-019)

Looking at the reverse case of minus-plural marking, we find 86 'unexpected' plurals (i.e., plus-plural marking on the noun). For example, in (12), the speaker refers to the obligatory PE class that students at the university have to take with the plural 'courses'.

- (12) it's a required **courses** (TASE-002)

Taking all plural forms in TASE into account (i.e., 323 following a quantifier + 222 following a numeral + 1,036 plurals in all other noun phrases), this amounts to a plus-plural rate of 5% (cf. SPOKE 4%).

Minus-Pronouns

Chinese allows minus-pronouns in virtually any position (subject, object) as long as they are recoverable from context. Unlike in Korean, there are no politeness levels encoded in the Mandarin pronouns; neither are they subject to case marking. Their pragmatic and syntactic ‘value’ is, therefore, low. We find that the Taiwanese speakers, as represented in TASE, employ minus-pronouns when they are speaking English. The examples below show different kinds of minus-pronouns in subject position (first-person singular in [13] and [14], second-person in [15] and [16], and third-person neuter in [17] and [18]) as used by the Taiwanese English speakers in our data set. Altogether, 306 minus-pronouns occur in subject position in TASE.

- (13) but in Taiwan Ø don’t think we need that very much (TASE-009) (first-person singular *I*)
- (14) Ø got a chance to have a interview (TASE-010) (first-person singular *I*)
- (15) and you don’t need to talk much (.) Ø just do it by yourself (TASE-018) (second-person singular *you*)
- (16) oh Ø don’t have convenient stores? (TASE-006) (second-person plural *you*)
- (17) actually Ø can be even good for (TASE-009) (third-person singular *it*)
- (18) mainly Ø is these kind of stuff or hotpots (TASE-002) (third-person singular *it*)

Minus-pronouns also occur in object position in TASE (n = 14):

- (19) you have to do it otherwise it could curse Ø (TASE-007) (second-person singular *you*)
- (20) you don’t know Ø? (TASE-018) (third-person singular *her*)
- (21) so I booked Ø I write a email to you (TASE-019) (third-person singular *it*)

The distribution of minus-pronouns in TASE, as illustrated in Table 5, differs only in some instances from that in SPOKE. We note, firstly, that the inanimate third-person singular *it* is the most prominent minus-pronoun in both datasets, accounting for more than half of the cases. The first-person singular follows, more obviously so in TASE, but still accounting for 22% in SPOKE. Overall, singular minus-pronouns appear with much greater frequency than plural ones, with 90% vs 10% in TASE and a similar 85% vs 15% in SPOKE (note that *you* is counted as singular in this computation).⁵

Table 5: Minus-pronouns in TASE and SPOKE

Minus-pronouns	Raw frequency + percentage (TASE)	Raw frequency + percentage (SPOKE)
<i>I/me</i>	88 29%	79 22%
<i>you</i>	20 6%	14 4%
<i>she/her</i>	5 2%	13 4%
<i>he/him</i>	8 3%	12 3%
<i>it</i>	156 51%	184 52%
<i>we/us</i>	16 5%	21 6%
<i>they/them</i>	15 5%	33 9%
Σ singular minus-pronouns	277 90%	302 85%
Σ plural minus-pronouns	31 10%	54 15%

The data presented here needs to be qualified somewhat, though. Firstly, the numbers presented in Table 5 are raw frequencies of minus-pronouns, and do not yet take into account plus-pronouns. The percentages refer only to the types of minus-pronouns within each corpus. Secondly, TASE is still in its pilot phase and is thus, with ca. 76,000 words, less than a third the size of SPOKE (300,000 words). Planned future data collection efforts will make the corpora more comparable in size.

Conclusion

With the increasing spread of English into locales of Kachru’s Expanding Circle, the need for a better understanding of these varieties (both from a structural and sociolinguistic point of view) arises. To date, the bulk of large-scale English-language corpora has focused on Inner Circle and Outer Circle Englishes – the former benefitting from a variety of historical and synchronic corpora. In the Outer Circle, the International Corpus of English (better known as ICE) project has provided a usefully comparable set of data from a range of varieties (see Nelson, 2019). When it comes to

the Expanding Circle, however, there is a dearth of comparable databases, which has a negative impact on our understanding of the structural features of these varieties.⁶ The prominence of English in many such settings, especially in the larger East Asian context, suggests that the number of speakers of Expanding Circle Englishes is likely to continue growing rapidly. It is our conviction that in order to further our understanding of the English language system in all its variation, a broad selection of corpora from a variety of Englishes, regardless of their Kachruvian status, can only improve the state of research in World Englishes.

From a global sociolinguistic perspective, we also believe that scholarly attention given to Expanding Circle Englishes ('norm-dependent' varieties, in Kachru's terms) has the potential to legitimize these ways of speaking in a way that also legitimizes the speakers themselves. We here align with Seilhamer who notes that 'ideologies privileging North-American accented English' can most effectively be challenged by speakers 'proudly, and even audaciously, assert[ing] the legitimacy of their English usage rather than apologizing for it' (2019: 194–195). We hope that the scholarly recognition of Expanding Circle Englishes, also in the form of publicly available corpora, is a further supporting step into this direction.

Finally, while some might question the systemic relevance of data from Englishes that are spoken as a mere 'foreign' language in polities where it has few institutional roles, we submit that if the language is embedded enough within the speech community, then it deserves attention, especially so as increased opportunities for cross-varietal contact may well bring these varieties to international prominence. Research endeavors that account for the entire range of English variation are inherently valuable contributions to the entire field of World Englishes.

Notes

- 1 While this course of action seems ultimately unrealistic, the point here is that these kind of discourses exist in both contexts and have also been taken up by local media.
- 2 While 22 speakers were recorded, one speaker of a triadic conversation had to be excluded from the corpus as it later turned out that she was an international student and thus not part of the target group of speakers.
- 3 Due to the exploratory nature of the fieldwork and the initial contact between the researchers and the participants, most students (n = 13) were undertaking their degrees at the English department. For the corpus at large (i.e., beyond the pilot corpus described here), we aim to include a more diverse speaker population,

with a focus on non-English major students and early professionals (cf. SPOKE; Rüdiger, 2019).

4 One speaker indicated dual Taiwanese and Thai citizenship and one speaker had dual Taiwanese and American citizenship. All participants specified (Mandarin) Chinese as their L1, with the exception of the speaker with Thai/Taiwanese citizenship who specified Thai as L1.

5 The realized frequencies of pronouns (n = 9,623; 126.6 ptw [per thousand words]) in TASE are as follows: *I/me* (n = 4,204; 55.3 ptw), *you* (n = 1,214; 16.0 ptw), *she/her* (n = 359; 4.7 ptw), *he/him* (n = 431; 5.7 ptw), *it* (n = 1,718; 22.6 ptw), *we/us* (n = 935; 12.3 ptw), *they/them* (n = 762; 10.0 ptw).

6 The *International Corpus of Learner English* (ICLE) project (Granger et al., 2009) is sometimes mentioned as an Expanding Circle counterpart to ICE. ICLE, however, exclusively contains written language (i.e., written essays), and as reflected in its name, subsumes data collected in educational settings, framing corpus contributors as 'learners'. This is inherently different from our approach to both data collection (spoken, 'informal') and our conceptualization of speakers as users (and not learners) of English.

Acknowledgments

Jakob Leimgruber and Sofia Rüdiger gratefully acknowledge the DFG (German Research Foundation) funding for the pilot study (research grant LE 3136/4-1 and RU 2369/1-1). We are particularly thankful to our hosts and colleagues at Fu Jen University and our interview participants. We would also like to thank Franziska Ahlstick and Miriam Neuhausen for their help with transcription and tagging.

References

- Anthony, L. *AntConc* (3.5.7) [software]. 2018. Tokyo: Waseda University.
- Bolton, K., Botha, W. & Kirkpatrick, A. (eds.) 2020. *The Handbook of Asian Englishes*. Hoboken: Wiley.
- Bosher, S. & Stocker, J. 2015. 'Nurses' narratives on workplace English in Taiwan: Improving patient care and enhancing professionalism.' *English for Specific Purposes*, 38, 109–20.
- Chen, C. W.–Y. 2006. 'The mixing of English in magazine advertisements in Taiwan.' *World Englishes*, 25(3/4), 467–78.
- Chen, C. et al. 2018. 'Government offers vision of English as official language in Taiwan.' *Focus Taiwan News Channel*, January 1. Online at <<http://focustaiwan.tw/news/aedu/201801310015.aspx>> (Accessed November 15, 2018).
- Chen, S.–C. 2006. 'Simultaneous promotion of indigenization and internationalization: New language-in-education policy in Taiwan.' *Language and Education*, 20, 322–37.
- Chen, S.–C. & Tsai, Y. 2012. 'Research on English teaching and learning: Taiwan (2004–2009).' *Language Teaching*, 45(2), 180–201.

- Chien, C.-W. 2014. 'Integration of school features into Taiwanese elementary school new English curriculum.' *Education* 3-13, 42(6), 598-600.
- Chou, C. P. & Yuan, J. K. S. 2011. 'Buxiban in Taiwan.' *The Newsletter*, 56, 15.
- Chou, M.-H. 2015. 'Impacts of the Test of English Listening Comprehension (TELC) on students' English learning expectations in Taiwan.' *Language, Culture and Curriculum*, 28(2), 191-208.
- Cooper, R. L. 1989. *Language Planning and Social Change*. Cambridge: Cambridge University Press.
- Edwards, A. 2016. *English in the Netherlands – Functions, Forms and Attitudes*. Amsterdam/Philadelphia: John Benjamins.
- Ferrer, A. & Lin, T.-B. 2021. 'Official bilingualism in a multilingual nation: A study of the 2030 bilingual nation policy in Taiwan.' *Journal of Multilingual and Multicultural Development* DOI: 10.1080/01434632.2021.1909054
- Granger, S., Dagneaux, E., Meunier, F. & Paquot, M. 2009. *International Corpus of Learner English – Version 2*. Louvain-la-Neuve: Presses universitaires de Louvain.
- Grant, R. A. & Lee, I. 2009. 'The ideal English speaker: A juxtaposition of globalization and language policy in South Korea and racialized language attitudes in the United States.' In R. Kubota & A. Lin (eds.), *Race, Culture, and Identities in Second Language Education: Exploring Critically Engaged Practice*. London: Routledge, pp. 44-63.
- Hsu, J.-L. 1994. 'Englishization and language change in Modern Chinese in Taiwan.' *World Englishes*, 13(2), 167-84.
- Hu, G. & McKay, S. L. 2012. 'English language education in East Asia: Some recent developments.' *Journal of Multilingual and Multicultural Development*, 33(4), 345-62.
- Jefferson, G. 1984. 'Note on a systematic deployment of the acknowledgement tokens "yeah" and "mm hm".' *Research on Language and Social Interaction*, 17(2), 197-216.
- Kirkpatrick, A. (ed.) 2020. *The Routledge Handbook of World Englishes*. London/New York: Routledge.
- Kobayashi, P. I. 2020. 'English in Taiwan.' In K. Bolton, W. Botha & A. Kirkpatrick (eds.), *The Handbook of Asian Englishes*. Hoboken: Wiley, pp. 547-67.
- Liu, J. 2012. 'Does cram schooling matter? Who goes to cram schools? Evidence from Taiwan.' *International Journal of Educational Development*, 32(1), 46-52.
- National Development Council, Taiwan. 2018. 'Blueprint for developing Taiwan into a bilingual nation by 2030.' Online at <https://www.ndc.gov.tw/en/Content_List.aspx?n=D933E5569A87A91C> (Accessed July 16, 2021).
- National Statistics. 2021. 'Statistics from Statistical Bureau.' Online at <<https://eng.stat.gov.tw/point.asp?index=9>> (Accessed July 12, 2021).
- Nelson, G. 2019. 'World Englishes and corpora studies.' In C. L. Nelson, Z. G. Proshina, & D. R. Davis (eds.), *The Handbook of World Englishes* (2nd edn.) Hoboken, NJ: Wiley Blackwell, pp. 523-37.
- Park, J. S.-Y. 2011. 'The promise of English: Linguistic capital and the neoliberal worker in the South Korean job market.' *International Journal of Bilingual Education and Bilingualism*, 14(4), 443-55.
- Park, S. J. & Abelman, N. 2004. 'Class and cosmopolitan striving: Mothers' management of English education in South Korea.' *Anthropology Quarterly*, 77(4), 645-72.
- Peters, P. & Wong, D. 2014. 'Turn management and backchannels.' In K. Aijmer & C. Rühlemann (eds.), *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press, pp. 408-29.
- Price, G. 2014. 'English for all? Neoliberalism, globalization, and language policy in Taiwan.' *Language in Society*, 43(5), 567-89.
- Rüdiger, S. 2016. 'Cuppa coffee? Challenges and opportunities of compiling a conversational English corpus in an Expanding Circle setting.' In H. Christ, D. Klenovšak, L. Sönning & V. Werner (eds.), *A Blend of MaLT: Selected Contributions from the Methods and Linguistic Theories Symposium 2015*. Bamberg: University of Bamberg Press, pp. 49-71.
- Rüdiger, S. 2019. *Morpho-Syntactic Patterns in Spoken Korean English*. Amsterdam/Philadelphia: John Benjamins.
- Rüdiger, S. 2021. 'The use of *like* in Korean English speech.' *World Englishes*, 40(4), 548-61.
- Schreier, D., Hundt, M. & Schneider, E. (eds.) 2020. *The Cambridge Handbook of World Englishes*. Cambridge: Cambridge University Press.
- Seilhamer, M. F. 2015. 'The ownership of English in Taiwan.' *World Englishes*, 34(3), 370-88.
- Seilhamer, M. F. 2019. *Gender, Neoliberalism and Distinction through Linguistic Capital: Taiwanese Narratives of Struggle and Strategy*. Bristol/Blue Ridge Summit: Multilingual Matters.
- Simpson, A. 2007. 'Taiwan.' In A. Simpson (ed.), *Language and National Identity in Asia*. Oxford: Oxford University Press, pp. 235-59.
- Tai, J. & Tsay, J. 2015. 'Taiwan Sign Language.' In J. Bakken Jepsen, G. De Clerck, S. Lutalo-Kiingi & W. B. McGregor (eds.), *Sign Languages of the World: A Comparative Handbook*. Berlin: De Gruyter Mouton, pp. 771-809.
- Tsao, F. 2000. 'The language planning situation in Taiwan.' In R. B. Baldauf & R. B. Kaplan (eds.), *Language Planning in Nepal, Taiwan, and Sweden*. Clevedon: Multilingual Matters, pp. 60-106.
- Tsou, W.-L. & Kao, S.-M. (eds.) 2017. *English as a Medium of Instruction in Higher Education: Implementations & Classroom Practices in Taiwan*. Singapore: Springer.
- Yoo, O. K. 2005. 'Discourses of English as an official language in a monolingual society: The case of South Korea.' *Second Language Studies*, 23(2), 1-44.
- Wu, L.-Y. & Lau, K. 2019. 'Language education policy in Taiwan.' In A. Kirkpatrick & A. J. Liddicoat (eds.), *The Routledge International Handbook of Language Education Policy in Asia*. London: Routledge, pp. 151-61.
- Wu, M. 2011. 'Language planning and policy in Taiwan: Past, present and future.' *Language Problems & Language Planning*, 35, 15-34.