23

World Englishes from the Perspective of Dialect Typology

Benedikt Szmrecsanyi and Melanie Röthlisberger

23.1 Introduction

LINGUISTIC TYPOLOGY is concerned with classifying human languages and with identifying structural similarities and differences between these languages. DIALECTOLOGY is the study of typically vernacular and regionally restricted and/or distinctive forms of language. The intersection between typology and dialectology has received considerable attention in recent years (see, e.g., the papers in Kortmann 2004; Szmrecsanyi and Wälchli 2014): DIALECT TYPOLOGY (also known as SOCIOLINGUISTIC TYPOLOGY) is interested in the "extent to which differences of linguistic structure, whether within or between languages, can be ascribed to or explained in terms of features of the society in which the dialects in question are spoken" (Trudgill 1996:3; see also Trudgill 2004; Trudgill 2009a; Trudgill 2011). The aim of this contribution is to survey work on World Englishes that takes a dialect typology perspective, the remit of which we define rather generously as also including, for example, areal patterns.

This chapter is structured as follows. In Section 23.2, we set the scene by reviewing the set of language-external factors (variety type, world region, exposure to contact) that has been used to categorize World Englishes. Section 23.3 summarizes what we know about (vernacular) universals, angloversals, and related notions in World Englishes. In Section 23.4, we synthesize work on parameters of structural diversity in World Englishes (analyticity versus syntheticity, complexity versus simplicity). Section 23.5 offers some concluding remarks.

23.2 Language-External Factors

23.2.1 Variety Type

Needless to say, there are many fine-grained distinctions to be made regarding different types and subtypes of varieties of English, and these distinctions are meticulously covered in this handbook (see Hickey, Chapter 2, this volume). From a dialect typology perspective, the most basic typology customary in the literature, which is also used in reference works such as Crystal (2004) and Kortmann and Lunkenheimer (2013), distinguishes the following variety types:[1]

- **Native L1 (or "English as a native language," ENL) varieties of English**, such as Canadian English or New Zealand English. This type roughly corresponds to the Inner Circle in Kachru (1992).
- **Indigenized L2 (or "English as a second language," ESL) varieties of English**, such as educated Jamaican English or Malaysian English. This type roughly corresponds to the Outer Circle in Kachru (1992).
- An inclusive typology will also recognize **English-based pidgin and creole languages**, such as Tok Pisin and Hawai'i Creole, as a third type.

This typology is primarily defined based on language-external facts – in terms of how and when English is acquired (first-language acquisition versus second-language acquisition) and with regard to whether or not we are dealing with a contact language. Against this backdrop, dialect typologists have been primarily concerned with establishing those linguistic features that are particularly diagnostic of specific types. Szmrecsanyi and Kortmann (2009a) refer to such features as "varioversals" (see also Section 23.3), that is, "features recurrent in language varieties with a similar socio-history, historical depth, and mode of acquisition" (p. 33).

How are varioversals identified empirically? A common way is to utilize survey databases, similar to the World Atlas of Language Structures (WALS) (Dryer and Haspelmath 2013), which have become an indispensable tool in the field of cross-linguistic typology. One popular survey in the World Englishes literature is the morphosyntax survey[2] that accompanies the *Handbook of Varieties of English* (Kortmann et al. 2004). This survey of nonstandard English morphosyntax was conducted by compiling a catalogue of seventy-six nonstandard features. The authors of the chapters in the morphosyntax volume of the *Handbook* were then asked to rate the features in the relevant variety according to the following categories:

[1] We acknowledge that an argument could be made to include Learner Englishes in this typology; however, in keeping with much of the dialect typology literature we take the liberty to not consider Learner Englishes and refer the reader to Edwards and Seargeant (Chapter 15, this volume) instead.

[2] See www.varieties.mouton-content.com/

A pervasive (possibly obligatory) or at least very frequent
B exists but a (possibly receding) feature used only rarely, at least not frequently
C does not exist or is not documented

The survey covers forty-six varieties of English around the world, all of which are vernacular (see Kortmann and Szmrecsanyi 2004 for discussion).[3] The survey reveals that the top L1 varioversals (i.e. morphosyntactic features that are particularly characteristic of L1 varieties of English) include existential/presentational *there's*, *there is*, *there was* with plural subjects (e.g. *There's two men waiting in the hall*); *me* instead of *I* in coordinate subjects (e.g. *Me and my brother*); and adverbs having the same form as adjectives (e.g. *Come quick!*). The top L2 varioversals are lack of inversion in main clause yes/no questions (e.g. *You get the point?*); irregular use of articles (e.g. *Take them to market*, *I had nice garden*, *about a three fields*, *I had the toothache*); and leveling of the difference between the present perfect and the simple past (e.g. *Were you ever in London?*, *Some of us have been to New York years ago*). Finally, distinctive varioversals for English-based pidgin and creole languages include lack of inversion/lack of auxiliaries in *wh*-questions (e.g. *What you doing?*); lack of inversion in main clause yes/no questions (e.g. *You get the point?*); and special forms or phrases for the second-person plural pronoun (e.g. *youse, y'all, aay', yufela, you . . . together, all of you, you ones/'uns, you guys, you people*). It is clear that feature profiles like this can be interpreted in terms of language simplicity and complexity in connection with the mode of acquisition (see, e.g., Szmrecsanyi and Kortmann 2009b and Section 23.4).

Beside this rather feature-centric view of variety types, dialect-typological work on World Englishes has also occasionally adopted a more holistic, variety-centric perspective: What is the extent to which variety type shapes overall similarities and differences between varieties of English? Szmrecsanyi and Kortmann (2009c) propose to use a statistical technique called Multidimensional Scaling (MDS) to address this issue. MDS (Kruskal and Wish 1978) is a well-known dimension-reduction technique that translates distances between objects (in our case, varieties of English) in high-dimensional space into a lower-dimensional representation. To establish aggregate distances between varieties, we may use the well-known squared Euclidean distance measure, which calculates the distance between any two varieties as the number of feature classifications with regard to which the varieties differ. Applying the technique to the morphosyntax survey that accompanies the *Handbook of Varieties of English* and merging the A and B ratings into an "attested" category (while

[3] There is also an updated version, the Electronic World Atlas of Varieties of English (eWAVE) (see Kortmann and Lunkenheimer 2013).

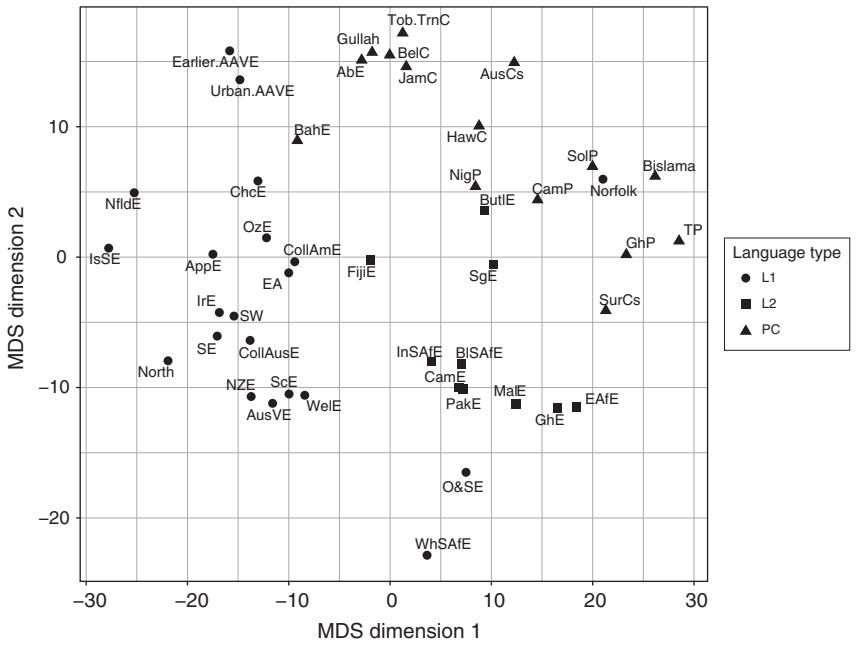

Figure 23.1 Metric Multidimensional Scaling (MDS) map, based on the morphosyntax survey from the *Handbook of Varieties of English* (see Kortmann and Szmrecsanyi 2004) Proximity between varieties in the plot is proportional to their aggregate morphosyntactic similarity.
Note. Abbreviations: E/R/UAAVE: Earlier/Rural/Urban African American Vernacular English; AbE/AborE: (Australian) Aboriginal English; AppE: Appalachian English; AusCs: Australian Creoles; AusVE: Australian Vernacular English; BahE: Bahamian English; BelC: Belizean Creole; BlSAfE: Black South African English; ButlE: Butler English; CamP/E: Cameroon Pidgin/English; ChcE: Chicano English; CollAmE: Colloquial American English; CollAusE: Colloquial Australian English; EA: East Anglian English; EAfE: East African English; FijiE: Fiji English; GhP/E: Ghanaian Pidgin/English; HawC/E: Hawai'ian Creole/English; InSAfE: Indian South African English; IrE: Irish English; IsSE/SEAmE: Isolated South Eastern American English; JamC/E: Jamaican Creole/English; MalE: Malaysian English; ManxE: Manx English; NfldE: Newfoundland English; NigP/E: Nigerian Pidgin/English; North: English dialects in the north of England; NZE: New Zealand English; O&SE: Orkney & Shetland English; OzE: Ozarks English; PakE: Pakistani English; ScE: Scottish English, Scots; SE: English dialects in the southeast of England; SgE: Singapore English; SolP: Solomon Islands Pidgin; SurC(s): Suriname Creoles; SW: English dialects in the southwest of England; TP: Tok Pisin, New Guinea Pidgin, Neomelanesian; Tob.TrnC: Creoles of Trinidad & Tobago; WelE: Welsh English; WhSAfE: White South African English.

C counts as "not attested") yields the MDS plot in Figure 23.1. In this plot, a nice three-way split emerges: The English-based pidgin and creole languages (e.g. Tok Pisin, Ghanaian Pidgin) are located in the right half of the plot, while native L1 varieties (e.g. Newfoundland English, dialects in the north of England) are located in the left half. Indigenized L2 varieties of English (e.g. Bahamian English, Fiji English, Malaysian English) are sandwiched in between, as it were. Therefore, in the big picture, variety type is clearly a major determinant of overall grammatical similarities and differences between varieties of English.

23.2.2 Areality

Areal patterns of linguistic similarity, possibly thanks to contact, is what takes center stage in both areal typology and in classical dialectology (see Murelli and Kortmann 2011), and so it is not surprising that areal effects in World Englishes have received attention as well: Do varieties of English spoken in, say, the British Isles share particular features that tend to be absent in varieties of English spoken in, say, North America? With regard to sound systems we know, for example, that British varieties of English tend to have an extensive system of diphthongs (Schneider 2004:1127), that West African varieties tend to have five-vowel systems, and that there is more generally speaking a list of sound features that are particularly diagnostic of regional accents, such as the TRAP vowel or yod-dropping (Schneider 2004:1129). In the realm of grammar, surveys such as Kortmann and Szmrecsanyi's (2004) show that British varieties tend to have, e.g., existential/presentational *there's*, *there is*, *there was* with plural subjects; that American varieties tend to have, e.g., special forms or phrases for the second-person plural pronoun; that Caribbean varieties often have, e.g., multiple negation; that Asian varieties regularly exhibit, e.g., irregular (from a standard English perspective) use of articles; and that African varieties more often than not attest a wider range of uses of the progressive.

More generally speaking, the areal null hypothesis is that geographical proximity between dialects or varieties should predict linguistic similarity between these dialects and varieties (Nerbonne and Kleiweg 2007:154 refer to this as the "Fundamental Dialectology Principle"). Yet precisely how important are such areal patterns, compared to other factors such as variety type? Consider Szmrecsanyi (2012), a study that is concerned with similarities and differences between thirty L1 varieties sampled in the World Atlas of Varieties of English (WAVE) (see Kortmann and Lunkenheimer 2012), a database that covers 235 morphosyntactic features. To shed light on relatedness patterns between those L1 varieties, the paper presents a NeighborNet diagram, which is reproduced in Figure 23.2. Originally developed in biometry and bioinformatics to map phylogenies and reticulate effects such as genetic recombination, NeighborNets are now quite popular in dialectology (e.g. McMahon et al. 2007) as well as in historical linguistics and in cross-linguistic typology (e.g. Dunn et al. 2008). Without insisting on a strictly phylogenetic interpretation, Figure 23.2 visually depicts – like the MDS map in Figure 23.1 – aggregate similarities and distances between L1 varieties of English. The diagram can be read like a family tree that is not rooted. Branch lengths are proportional to linguistic distance: Proximity in the plot broadly indicates morphosyntactic similarity. The most crucial split in Figure 23.2 we observe is between what Trudgill (2009b) would call "high-contact" L1 varieties (at the bottom of the diagram) and the other L1 varieties in the sample (more

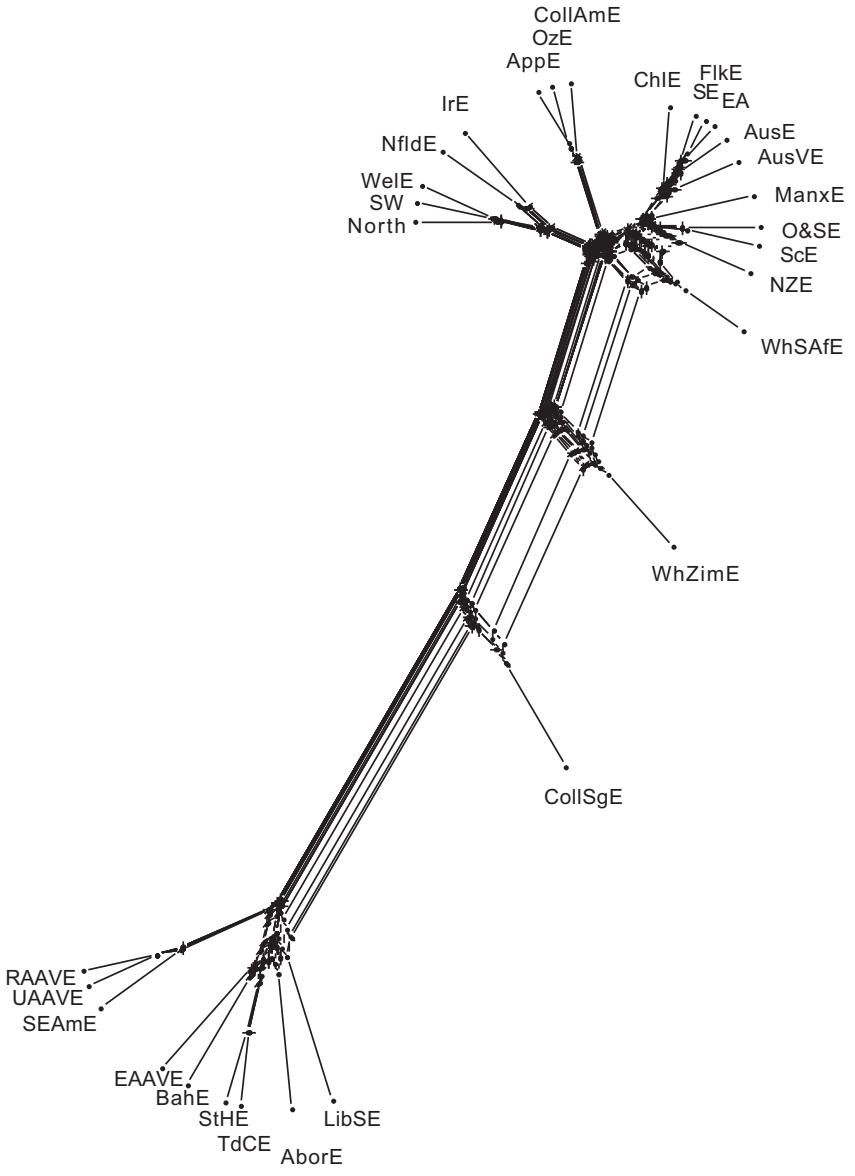

Figure 23.2 Visualizing aggregate similarities: NeighborNet diagram (after Szmrecsanyi 2012: figure 4, used with permission)
Distances (branch lengths) are proportional to cophenetic linguistic distances.
Note. Abbreviations: E/R/UAAVE: Earlier/Rural/Urban African American Vernacular English; AbE/AborE: (Australian) Aboriginal English; AppE: Appalachian English; AusCs:; AusVE: Australian Vernacular English; BahE: Bahamian English; ChIE: Channel Island English; CollAmE: Colloquial American English; CollSgE: Colloquial Singapore English; EA: East Anglian English; FijiE: Fiji English; IrE: Irish English; IsSE/SEAmE: Isolated South Eastern American English; LibSE: Liberian Settler English; ManxE: Manx English; NfldE: Newfoundland English; North: English dialects in the north of England; NZE: New Zealand English; O&SE: Orkney & Shetland English; OzE: Ozarks English; ScE: Scottish English, Scots; SE: English dialects in the southeast of England; SEAmE / IsSE: Southeastern US enclave dialects; StHE: St Helena English; SW: English dialects in the southwest of England; TdCE: Tristan da Cunha English; WelE: Welsh English; WhSAfE: White South African English; WhZimE: White Zimbabwean English.

on the distinction between high-contact and low-contact varieties in Section 23.2.3). In the "other" cluster to the top of the diagram, we find some areal subgroupings, such as Colloquial American/Ozarks/Appalachian English (an American cluster) and Welsh/Southwest/North English (a British cluster). At the bottom of the diagram, we find another areal cluster consisting of Rural AAVE, Urban AAVE, and Southeastern American English. The upshot is that there are areal effects in World Englishes, but these are less important than other factors.

We note that the importance of geography and areality can be precisely quantified by correlating pairwise linguistic distances (as calculated by, for example, the number of feature classifications in WAVE with regard to which two varieties differ) with pairwise geographic distances (as the crow flies, in kilometers, calculated using a standard trigonometry formula). For the set of L1 varieties depicted in Figure 23.2, the correlation coefficient between morphosyntactic and geographical distances comes out as $r = 0.226$ ($p < 0.001$). In plain English, this means that there is a significant areal relationship, but geographical distance explains no more than 5.1 percent ($R^2 = 0.051$) of the morphosyntactic variability in the dataset. This is a comparatively modest share compared to dialectometry measurements in traditional dialects. Shackleton (2007), for example, finds that geographical distance explains about 49 percent of phonetic variation in traditional English dialects in England. We conclude that areal patterns and geography are not particularly powerful explanatory factors in the realm of World Englishes, and so we turn to alternative factors, such as exposure to language and dialect contact, next.

23.2.3 Exposure to Language and Dialect Contact

Typological work aiming at cross-linguistic comparisons has often taken an interest in the impact that language contact has on the systematic distribution of structural features across the world's languages (see, e.g., Aikhenvald and Dixon 2001; Siemund and Kintana 2008). Such research has focused on disentangling the mechanisms and principles of contact-induced change, specifically the constraints (linguistic as well as social) that influence the outcome of language contact (Siemund 2008:3). For instance, factors such as the degree of bilingualism, architecture, and prestige of the languages involved; the number of speakers; the length of contact; and numerous other parameters (see Siemund 2008:4) have been scrutinized with regard to their influence on contact-based structural similarities between languages. Research in areal typology (see previous section) has further shown that these structural similarities can also arise between genetically unrelated languages in what has been called "*Sprachbund*" or "linguistic areas" (see Matras 2009: 236, 266).

The typological literature on the mechanisms and principles involved in language contact has stimulated dialectological work on World Englishes.

Spearheading this line of research was Peter Trudgill, who argues that language and dialect contact was and is the driving force for the diversification and emergence of new varieties (see Trudgill et al. 2000; Trudgill 2006; Trudgill 2008). Crucially, Trudgill (2009b: 320) proposes a typological split between "low-contact" varieties – long-established mother-tongue dialects – and "high-contact" varieties including the following:

- Non-native indigenized L2 varieties: e.g. Indian English, Hong Kong English, Philippine English, or Jamaican English
- Transplanted L1 Englishes or (post)colonial standards: e. g. New Zealand English, White South African English, or Maltese English
- Language-shift Englishes: e. g. Irish and Welsh English
- Standard L1 varieties: e. g. British and American English
- Creoles: e.g. Hawai'i Creole or Tok Pisin (see also Kortmann and Szmrecsanyi 2011: 15–16)

Needless to say, it is "high-contact" varieties that are of particular interest when it comes to language and dialect contact. During colonization, transplanted varieties of English were in contact with indigenous languages and a range of European languages. As a consequence, newly emerging varieties of English borrowed linguistic elements from other languages, for example *cookie* from Dutch or *mana* from Maori (Trudgill 2006: 267). Most importantly, however, colonization inevitably led to contact between various speakers of different (British) English dialects.

While it is agreed that new varieties of English emerge out of contact with other dialects and languages, predicting the structural outcome of such contact is a challenging task (Siemund 2008: 3). Researchers have provided evidence for both complexification and simplification as results of contact-induced change (see Section 23.4.2). Schreier (2016) has challenged the binary distinction between high-contact and low-contact contact varieties. He argues that the degree of similarity between the languages/dialects in contact is more important than the degree of contact (Schreier 2016: 145), as simplification only seems to occur in high-contact situations if the two linguistic systems are maximally different from each other. This difference between linguistic systems in high-contact scenarios is a crucial one if one distinguishes between dialect (same language) and language (different languages) contact settings. According to Schreier, it is language contact settings that lead to simplification while dialect contact settings do not.

The difference between language and dialect contact varieties becomes even more evident in the acquisition process: In dialect contact scenarios, the majority of language learners are children who select various features from the heterogeneous input of the feature pool (Trudgill 2010). Childhood language acquisition results in an increase of linguistic variants and complexification, as in the case of New Zealand English, Canadian English, or American English. In language contact scenarios, adults acquire English as a second language, which inevitably leads to

simplification due to the limited language-acquisition abilities of adult learners (e.g. India, Singapore, or Hong Kong English) (Trudgill 2010). Since we expect to find simplification in those varieties where English is spoken by adult learners, and to find signs of complexification in those varieties where English has been acquired by (bilingual) children, the prediction is that there is more simplification in indigenized L2 varieties of English compared to colonial L1 varieties. Whether this hypothesis matches the facts will be discussed in Section 23.4.2.

23.3 Dialect Universals, Implications, and Related Notions

The quest for generalizations, also known as universals – what is it that human languages in general, or languages belonging to particular types, tend to have in common? – is an important topic in cross-linguistic typology and so has also inspired work on the dialect typology of World Englishes. Against this backdrop, Szmrecsanyi and Kortmann (2009a) present the following typology of universals:

(i) GENUINE UNIVERSALS (e.g. *all languages have vowels*);
(ii) TYPOVERSALS, i.e. features that are common to languages of a specific typological type (e.g. *SOV languages tend to have postpositions*);
(iii) PHYLOVERSALS, i.e. features that are shared by a family of genetically related languages (e.g. *languages belonging to the Indo-European language family distinguish between masculine and feminine gender*);
(iv) AREOVERSALS, i.e. features common to languages which are in geographical proximity to each other (e.g. *languages belonging to the Balkan Sprachbund have finite complement clauses*);
(v) VERNACULAR UNIVERSALS, i.e. features that are common to spoken vernaculars (e.g. *spoken vernaculars tend to have double negation*);
(vi) features that tend to recur in vernacular varieties of a specific language: ANGLOVERSALS, FRANCOVERSALS, etc. (e.g. *in English vernaculars, adverbs tend to have the same morphological form as adjectives*);
(vii) VARIOVERSALS, i.e. features recurrent in language varieties with a similar socio-history, historical depth, and mode of acquisition (e.g. *L2 varieties of English tend to use resumptive pronouns in relative clauses*). (Szmrecsanyi and Kortmann 2009a: 33)

Orthodox typologists tend to be concerned with (i) to (iv), while (v) to (vii) are the realm of dialect typology. In what follows, we take the liberty to discuss the relevant notions in more detail.

According to Jack Chambers (e.g. Chambers 2004), VERNACULAR UNIVERSALS comprise "a small number of phonological and grammatical processes [that] recur in vernaculars wherever they are spoken . . . not only

in working class and rural vernaculars, but also in . . . pidgins, creoles and interlanguage varieties" (p. 128). Chambers has specifically suggested the following features as candidates for vernacular universalhood:

- (ng) or alveolar substitution in final unstressed –ing, as in *walkin'*, *talkin'* and *runnin'*.
- (CC) or morpheme-final consonant cluster simplification, as in *pos' office*, *han'ful*.
- final obstruent devoicing, as in *hundret* (for hundred), *cubbert* (for cupboard).
- conjugation regularization, or leveling of irregular verb forms, as in *Yesterday John seen the eclipse* and *Mary heared the good news*.
- default singulars, or subject-verb nonconcord, as in *They was the last ones*.
- multiple negation, or negative concord, as in *He didn't see nothing*.
- copula absence, or copula deletion, as in *She smart* or *We going as soon as possible*. (Chambers 2004: 129)

Chambers himself notes that the examples might be from English but since they are "primitive features, not learned" (and thus part of the language faculty), they cannot be restricted to English only (Chambers 2004: 129).

The putative ubiquity of such features is claimed to be unlikely to be due to sociolinguistic diffusion. Therefore, they must be "primitive features of vernacular dialects" (Chambers 2003: 243), unlearned and thus innate. For more discussion, we refer the reader to the papers in Filppula et al. (2009).

A related notion is that of ANGLOVERSAL(S), a term that is used in two different ways in the literature. While Mair (2003) uses the notion to refer to universals of postcolonial Englishes, Szmrecsanyi and Kortmann (2009a) – whose usage we follow here – define angloversals as being recurrent features in *all* varieties and types of English. Angloversals in this sense include, for example, lack of inversion in main clause *yes/no* questions (e.g. *You get the point?*); *me* instead of *I* in coordinate subjects (e.g. *Me and my brother were late for school*); and *never* as preverbal past tense negator (e.g. *he never came*) (Kortmann and Szmrecsanyi 2004: table 3).

In this connection, we should mention the related notion of "rara," that is, features that are rather infrequent in the languages of the world (e.g. Wohlgemuth and Cysouw 2010). Numerous reasons have been suggested for the infrequency of rara, such as their increased processing difficulty, the rara's malfunction in communication, and the low probability of rara arising in the first place (Harris 2010). In the realm of World Englishes, rara include the *after*-Perfect (e.g. *She's after selling the boat*) and the relative particle *at* (e.g. *This is the man at painted my house*) (Kortmann and Szmrecsanyi 2004: table 2).

Beyond the quest for universals, cross-linguistic typology also often takes an interest in co-occurrence patterns of linguistic features, for the sake of learning more about how features evolve. The relevant patterns

can be biconditional implications (for instance, "if in a language the genitive follows the noun, then the complement follows the adposition, and vice versa"; Greenberg 1963) or one-way implications (for instance, "if a language has a marked singular, it has also a marked plural, but not necessarily vice versa"; Greenberg 1966). Implications can be extended to what is known as "implicational hierarchies," which predict that if a language has a property at some point in the hierarchy it will also possess all properties further down the hierarchy (see Siemund 2013: 17). Again, these notions can be, and have been, fruitfully transferred to the study of World Englishes. To illustrate, Szmrecsanyi and Kortmann (2009c) study the morphosyntax survey coming with the *Handbook of Varieties of English* and find that 94 percent of the varieties covered in the survey either have both *ain't* as the negated form of *be* (e.g. *They're all in there, ain't they?*) as well as *ain't* as the negated form of *have* (e.g. *I ain't had a look at them yet*) or have neither. This biconditional implication is, needless to say, in line with the dialectological literature (Anderwald 2003: 149–150). Building on such observations, Szmrecsanyi (2017) puts the quest for co-occurrence patterns in World Englishes on a more solid quantitative footing and marshals multiple correspondence analysis (MCA) (Lê, Josse, and Husson 2008; Levshina 2015: 375–376) to study co-occurrence patters in the morphosyntax survey of the *Handbook of Varieties of English*. The technique explores how categorical variables (in our case, features) are associated with each other, and provides information about the behavior of individual observations (in our case, varieties): A particular variety will appear in the same part of the plot as the values of the features by which the variety is characterized. The World Englishes MCA plot is shown in Figure 23.3.

The picture that emerges from Figure 23.3 can be summarized as follows. In the upper left-hand quadrant, we find features such as serial verbs and *no* as preverbal negator, which are demonstrably characteristic of English-based pidgin and creole languages; and, indeed, the varieties that the MCA plot identifies as particularly attracted to these features (e.g. Australian Creoles) are all pidgins and creoles. In the upper right-hand quadrant three features are identified as particularly distinctive: *ain't* as generic negator before a main verb, *ain't* as the negated form of *have*, and *ain't* as the negated form of *be*. This co-occurrence pattern ties in with what was said already. Note additionally that the varieties located in this quadrant are all North American, indicating that – as is well-known – *ain't* is particularly characteristic of North American Englishes. In the lower right-hand quadrant the plot locates some British varieties as well as New Zealand English, a variety that is known to be fairly close to British English, at least in terms of grammar. Note that corpus-based studies can offer slightly different results in that they find New Zealand and Australian English often wedged between British and American English with diverging orientation toward one or the other depending on the part of grammar that one looks at (e.g. Hundt 1998). Distinctive features in this corner

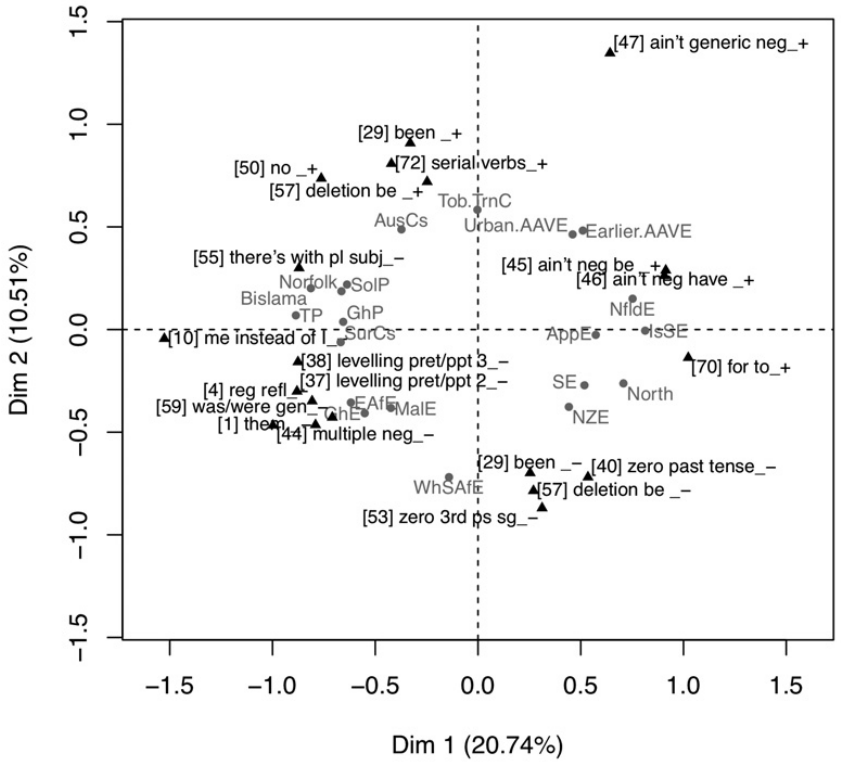

Figure 23.3 Multiple Correspondence Analysis (MCA) map, based on the morphosyntax survey from the *Handbook of Varieties of English* (from Szmrecsanyi 2017: figure 2) Proximity between features indicates co-occurrence patterns. Display is limited to the 20 features and varieties that have the highest contribution on the dimensions. "+" suffixed to a feature's label indicates presence of the feature, "–" indicates absence.
Note. Abbreviations: E/R/UAAVE: Earlier/Rural/Urban African American Vernacular English; AppE: Appalachian English; AusCs: Australian Creoles; EAfE: East African English; GhP/E: Ghanaian Pidgin/English; IsSE/SEAmE: Isolated South Eastern American English; MalE: Malaysian English; NfldE: Newfoundland English; North: English dialects in the north of England; NZE: New Zealand English; SolP: Solomon Islands Pidgin; SurC(s): Suriname Creoles; TobC: Tobagonian Creole; TP: Tok Pisin, New Guinea Pidgin, Neomelanesian; WhSAfE: White South African English.

include, e.g., unsplit *for to* in infinitival purpose clauses, while, e.g., deletion of *be* is typically absent. This distributional pattern is typical of British varieties of English (Kortmann and Szmrecsanyi 2004: 1162–1165). In the lower left-hand quadrant, we find primarily indigenized L2 varieties such as Malaysian English. MCA suggests that these varieties are characterized by the absence of features such as *them* instead of demonstrative *those*, and multiple negation. In all, therefore, the analysis would seem to suggest the following dialect typology: The most important dimension of variation (Dim 1) pits native varieties (right) against pidgins/creoles and L2 varieties (left). The vertical dimension (Dim 2) appears to be capturing a language–externally defined contrast between orientation toward North American English (top) versus orientation toward British English (bottom).

23.4 Parameters of Structural Diversity

Similarities between varieties of English are often discussed based on the trajectory of a variety's evolution (e.g. Schneider 2007), the status of English (e.g. the ENL/ESL/EFL distinction), the degree of contact (low vs. high), or shared linguistic (morphosyntactic) features (e.g. varioversals) as elucidated in Section 23.3. This section will introduce two sets of parameters to capture the structural diversity of World Englishes: analyticity vs. syntheticity (Section 23.4.1.), and complexity vs. simplicity (Section 23.4.2).

23.4.1 Analyticity vs. Syntheticity

The distinction between analytic and synthetic languages goes back to August Wilhelm von Schlegel (1818). Schlegel's original classification has been popular but has also received methodological criticism. Sapir (1921) proposes a classification that would allow languages to belong to more than one type and introduces a number of parameters along which languages should be categorized. Sapir's typology in turn influenced Greenberg (1960), who defined five indices to characterize languages, thus abolishing the need to assign languages categorically to one type (Greenberg 1960: 185). Greenberg proposed an essentially corpus-based method to classify languages and demonstrated that seemingly abstract typological notions are amenable to precise measurement through the calculation of text-based indices.

Drawing inspiration from Greenberg (1960), Szmrecsanyi (2009) and Szmrecsanyi and Kortmann (2009b) analyze the degree of grammatical analyticity and syntheticity, measured as an index of free versus bound grammatical markers per word, across a geographically widespread range of varieties of English. They define analyticity/syntheticity as follows:

- *Formal grammatical analyticity* includes all coding strategies where grammatical information is encoded with free grammatical markers defined as closed-class function words without any lexical meaning.
- *Formal grammatical syntheticity* includes all coding strategies where grammatical information is encoded with bound grammatical markers (see Szmrecsanyi 2009: 2).

A variety's degree of analyticity or syntheticity has been directly linked to mechanisms of simplification and complexification at work in dialect and language contact (see Section 23.2.3). According to Trudgill (2010), simplification results from widespread adult Second Language Acquisition (SLA) and manifests in two possible ways: regularization of irregularities and an increase in lexical and morphological transparency (Trudgill 2010: 307). Lexical and morphological transparency entails analytic structures where

"the relation between form and meaning is as transparent as possible" and "every single meaning is expressed in a separate form" (Kusters 2003: 21). Languages that fall under Kuster's TRANSPARENCY PRINCIPLE are arguably more analytic and thus easier to learn for adult speakers (Trudgill 2010: 312). The most extreme cases where language contact results in simplification (and hence increased transparency/analyticity) are pidgins and creoles (see also Leufkens 2013; McWhorter 2001). On the other hand, it has been postulated that low-contact situations lead to complexification and arguably increased syntheticity. These hypotheses (namely that simplification leads to increased analyticity and complexification leads to increased syntheticity) have been tested using language data from varieties of English.

For example, Siegel et al. (2014) test the claim that we should find heightened analyticity in creole languages. Their study explores the coding of grammatical information (free vs. bound morphemes) in two English-lexified creoles (Tok Pisin and Hawai'i Creole) and – for benchmarking purposes – in a number of rural dialects of British English, non-native indigenized L2 varieties, transplanted L1 varieties around the world, and language-shift varieties. Using the Greenberg-inspired typological profiling method discussed in this section, their analysis shows that, indeed, creoles are significantly less synthetic than other varieties of English. That is, they exhibit a greater ratio of analytic versus synthetic structure. At the same time, however, it turns out that creoles are not necessarily more analytic (in absolute terms) than indigenized or native varieties of English. Figure 23.4 illustrates this by locating varieties in a two-dimensional syntheticity–analyticity space. As can be seen, both Tok Pisin and Hawai'i Creole use synthetic markers less often than other varieties of English but they do exhibit a similar degree of analyticity. We also see that L1 varieties and traditional British dialects exhibit more syntheticity and analyticity than some L2 varieties where zero marking is relatively frequent (see Kortmann and Szmrecsanyi 2011: 275). Hence, while low-contact scenarios in general seem to lead to more grammatical marking (the sum of both analytic and synthetic indices), high-contact scenarios typically level grammatical marking, especially of the synthetic kind.

A similar pattern has been observed by Callies (2016) regarding processes of structural innovations in New Englishes. Tapping into learner as well as indigenized L2 varieties of English, Callies (2016: 244) concludes that the processes at play in structural innovations rely on and result in increased morphological transparency and maximal "explicitness of form-meaning relations" in both variety types. A similar preference for transparent forms in structural innovations has been found by Laporte (2012) who explores the use of *to*-infinitives in causative constructions (e.g. *to make someone to laugh*) in some ESL and EFL varieties. Similarly, Steger and Schneider (2012) find an

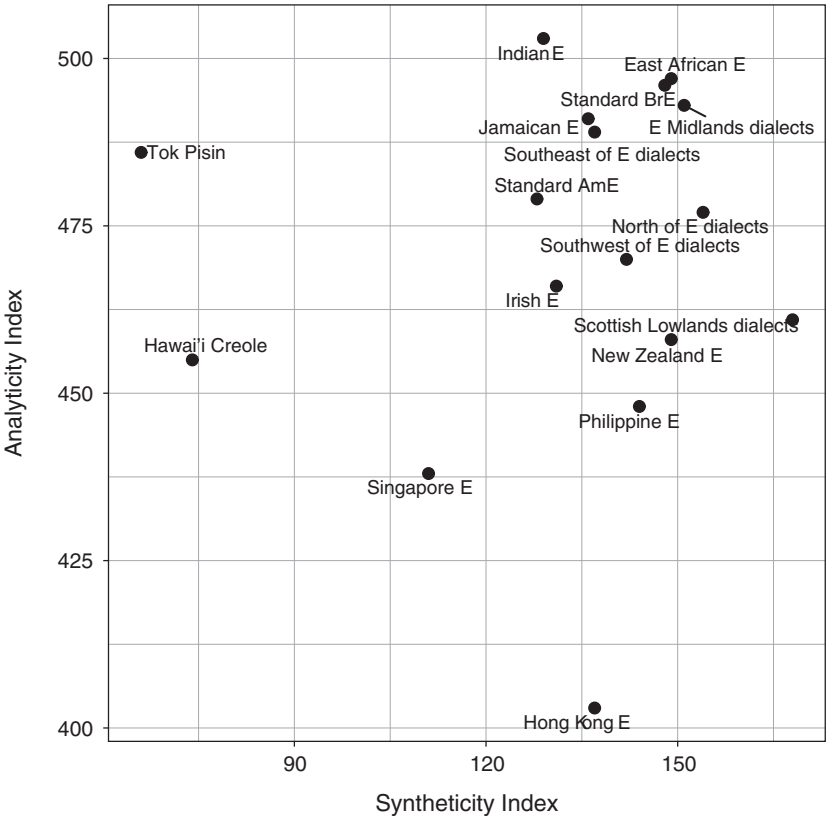

Figure 23.4 Tok Pisin and Hawai'i creole vis-à-vis varieties of English: total number of analytic types against total number of synthetic types (after Siegel et al. 2014: figure 2)

increase in the text frequency of overt complementizers in L2 varieties of English; Nesselhauf (2009) and Gilquin (2015) report usage of semantically redundant particles in (phrasal/prepositional) verbs in World Englishes (e.g. *enter into*) (see also Callies 2016: 246); and Mesthrie (2006) discusses "anti-deletion" in some L2 varieties, a tendency to use explicit markers where speakers of native varieties would omit them.

23.4.2 Complexity vs. Simplicity

Language complexity is a hot topic in both cross-linguistic typology and dialect typology. While twentieth-century structural linguists assumed all languages to be equally complex (see Sampson 2009 for an overview; also Newmeyer and Preston 2014), the issue of cross-linguistic and intra-lingual complexity differentials has increasingly received attention in the past few decades. The main idea behind the equi-complexity hypothesis had been the assumption of a trade-off

between the different subsystems of a language: simplicity in one linguistic subsystem would be compensated by more complexity in another subsystem. This trade-off hypothesis has been challenged at the beginning of the twenty-first century (see, e.g., Gil 2008; Nichols 2009; Shosted 2006 whose empirical analyses give no indication of a trade-off; see, however, Sinnemäki 2014).

How is this relevant to (dialect) typology and World Englishes? In a seminal paper challenging the equi-complexity hypothesis, McWhorter (2001) argued that creoles are less complex grammatically than their lexifier languages

> by virtue of the fact that they were born as pidgins, and thus stripped of almost all features unnecessary to communication, and since then have not existed as natural languages for a long enough time for diachronic drift to create the weight of "ornament" that encrusts older languages. (McWhorter 2001: 125)

There is an emerging consensus that language complexity is indeed variable (see the papers in, e.g., Kortmann and Szmrecsanyi 2012; Miestamo, Sinnemäki and Karlsson 2008; Sampson, Gil and Trudgill 2009). Among other things, scholarship has sought to link observable complexity levels to language variation and change for the sake of understanding simplification or complexification as processes, as well as their language-external triggers. Moreover, the impact of sociolinguistic factors on language complexity has received special attention, as well as the ways in which complexity can be actually measured.

Researchers have been especially interested in the extent to which language-external factors fuel complexity variation. Trudgill (2001: 372) links complexity to adult language learning when he states that "[a]dult language contact means adult language learning; and adult language learning means simplification, most obviously manifested in a loss of redundancy and irregularity and an increase in transparency." Childhood bilingualism on the other hand results in complexification (Trudgill 2011: 42; see Section 23.2.3). Another important factor is intensity of language contact: Low-contact varieties seem to exhibit more complexity than those communities that are, or have been, subject to intense contact with other languages or dialects (Trudgill 2011). Other language-external factors whose impact on a linguistic system's complexity has been explored include age, sex, class (Sampson 2001), population size (Sinnemäki 2011), and geography (Nichols 1992; see also Trudgill 2016).

Analysts have proposed various measures according to which the complexity of *langue* or *parole* can be gauged. Most generally speaking, complexity measures can be dichotomized as follows (Miestamo 2008):

– *Global complexity measure* versus *local complexity measures*: global complexity quantifies the complexity of an entire language/dialect. Local

complexity gauges the complexity of a domain-specific linguistic subsystem such as syntax or phonology.

– *Relative complexity measures* versus *absolute complexity measures*: measures of relative complexity assess subjective, user-oriented complexity (related to processing and learning). For instance, more complex linguistic phenomena are also more difficult to learn. Absolute complexity gauges objective, theory-oriented complexity by counting parts of the system such as the number of phonemes in a language.

More fine-grained categorizations would then go on to distinguish between, for example, absolute-quantitative complexity measures, where more material (bigger marker inventories etc.) equals "more complex," or redundancy-induced complexity, also called "ornamental complexity," where the amount of redundant linguistic material is counted toward a language's degree of complexity. SLA-based relative measures, too, are popular; they define language or dialect complexity as being proportional to the difficulty of learning the language or dialect in question.

In the World Englishes literature, the bulk of complexity-oriented research focuses on the difference in structural complexity between native and non-native varieties of English. For instance, Huber (2012) compares the results of structural nativization in the English relative clause system in Ghanaian English to present-day British English: The nativization of an already complex relativizer system can result in a similarly complex system where the structural factors that shape the choice of relativizer may have been assigned new importance but are still present overall. Also with regard to Ghanaian English, Schneider (2015) shows that future marker choice in Ghanaian English, an indigenized L2 variety, is less constrained by probabilistic factors (and thus, simpler) than it is in British English. Steger and Schneider (2012) explore the degree of iconicity and isomorphism in new varieties of English by looking at variable patterns in complement clause constructions. Adopting a cognitive-functionalist perspective, they define complexity as being a function of iconicity, that is, iconicity effects contribute to an increased transparency of grammatical encoding and hence an increase in simplicity. The authors find that iconic constructions are more popular in non-native than in native varieties and conclude that the cognitive principles at play during SLA lead to increased simplicity in second-language varieties of English (Steger and Schneider 2012: 187).

Kortmann and Szmrecsanyi (2009), finally, adopt a bird's-eye view and explore complexity patterns in, among other data sources, the morpho-syntax survey that accompanies the *Handbook of Varieties of English* (Kortmann et al. 2004). They specifically inspect the seventy-six morpho-syntactic features covered in the survey to identify (1) "simplifying features," that is, features or structures that simplify usage or the system, vis-à-vis standard English (an example would be leveling phenomena, such as

leveling of preterite and past participle verb forms), and (2) "L2-simple features," that is features that are known to recur in interlanguage varieties, such as resumptive relative pronouns of the type *This is the house which I painted* ***it*** *yesterday* (see, e.g., Hyltenstam 1984). Subsequently, Kortmann and Szmrecsanyi establish the number of simplifying and L2-simple features attested per variety; the distributional pattern is visually depicted in Figure 23.5 (the *x*-axis, labeled "Rule simplicity," plots the number of simplifying features; the *y*-axis plots the number of L2-simple features). Two observations should be highlighted. First, on the whole, rule simplicity predicts L2 simplicity and vice versa. Second, there is a split between

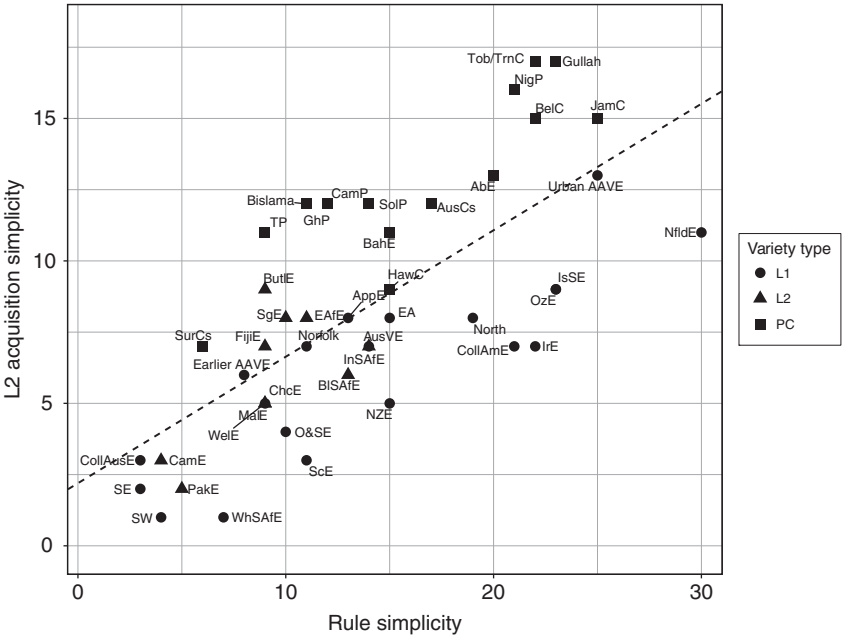

Figure 23.5 L2-simplicity by rule simplicity (after Kortmann and Szmrecsanyi 2009: diagram 1)
The dotted trend line represents linear estimate of the relationship.
Note. Abbreviations: E/R/UAAVE: Earlier/Rural/Urban African American Vernacular English; AbE/AborE: (Australian) Aboriginal English; AppE: Appalachian English; AusCs: Australian CreolesAusVE: Australian Vernacular English; BahE: Bahamian English; BelC: Belizean Creole; BlSAfE: Black South African English; ButlE: Butler English; CamP/E: Cameroon Pidgin/English; ChcE: Chicano EnglishCollAmE: Colloquial American English; CollAusE: Colloquial Australian English; EA: East Anglian English; EAfE: East African English; FijiE: Fiji English; GhP/E: Ghanaian Pidgin/English; HawC/E: Hawai'ian Creole/English; InSAfE: Indian South African English; IrE: Irish English; IsSE/SEAmE: Isolated South Eastern American English; JamC/E: Jamaican Creole/English; MalE: Malaysian English; NfldE: Newfoundland English; NigP/E: Nigerian Pidgin/English; North: English dialects in the north of England; NZE: New Zealand English; O&SE: Orkney & Shetland English; OzE: Ozarks English; PakE: Pakistani English; ScE: Scottish English, Scots; SE: English dialects in the southeast of England; SgE: Singapore English; SolP: Solomon Islands Pidgin; SurC(s): Suriname Creoles; SW: English dialects in the southwest of England; TobC: Tobagonian Creole; TrnC: Trinidadian Creole; TP: Tok Pisin, New Guinea Pidgin, Neomelanesian; WelE: Welsh English; WhSAfE: White South African English.

English-based pidgin and creole languages, which cluster in the upper right-hand quadrant (which means that they attest lots of simplifying and L2-simple features), and other varieties, which attest fewer simplifying and L2-simple features. What is surprising in this connection is that indigenized L2 varieties are not more clearly set apart from L1 varieties of English (see Kortmann and Szmrecsanyi 2009: 276 for more discussion).

23.5 Concluding Remarks

In this chapter, we have reviewed scholarship on World Englishes that is situated at the intersection between linguistic typology, a research field that is concerned with classifying human languages and with identifying structural similarities, and dialectology, which as a field is concerned with vernacular and regionally restricted and/or distinctive forms of language. Approaches that could not be discussed in this chapter but that are nonetheless relevant or at least neighboring to dialect typology include the extent to which variation patterns in particular varieties are rule-based or exemplar-based (see Baayen 2011 for some discussion); the potentially differential power of prescriptivism in different varieties (Hinrichs, Szmrecsanyi, and Bohmann 2015); and work that generates typologies of World Englishes by considering attitudes and transnational importance (Mair 2013). These are also topics whose interface with dialect typology would merit more attention in future research.

References

Aikhenvald, Alexandra Y. and Robert M. W. Dixon, eds. 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.

Anderwald, Lieselotte. 2003. *Negation in Non-Standard British English: Gaps, Regularizations and Asymmetries* (Routledge Studies in Germanic Linguistics). London, New York: Routledge.

Baayen, R. Harald. 2011. Corpus linguistics and naive discriminative learning. *Revista Brasileira de Linguística Aplicada 11*(2): 295–328. doi:10.1590/S1984-63982011000200003.

Callies, Marcus. 2016. Towards a process-oriented approach to comparing EFL and ESL varieties: A corpus-study of lexical innovations. *International Journal of Learner Corpus Research 2*(2): 229–250. doi:10.1075/ijlcr.2.2.05cal.

Chambers, Jack K. 2003. *Sociolinguistic Theory: Linguistic Variation and Its Social Implications*. Oxford: Blackwell.

Chambers, Jack K. 2004. Dynamic typology and vernacular universals. In Bernd Kortmann, ed. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter, 127–145.

Crystal, David. 2004. *The Cambridge Encyclopedia of the English Language* (2nd ed.). Cambridge: Cambridge University Press.

Dryer, Matthew S. and Martin Haspelmath, eds. 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/

Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink and Angela Terrill. 2008. Structural phylogeny in historical linguistics: Methodological explorations applied in Island Melanesia. *Language 84*(4): 710–759.

Filppula, Markku, Juhani Klemola and Heli Paulasto, eds. 2009. *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond* (Routledge Studies in Germanic Linguistics 14). New York: Routledge.

Gil, David. 2008. How complex are isolating languages? In Matti Miestamo, Kaius Sinnemäki and Fred Karlsson, eds. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 109–131.

Gilquin, Gaëtanelle. 2015. At the interface of contact linguistics and second language acquisition research: New Englishes and Learner Englishes compared. *English World-Wide 36*(1): 90–123.

Greenberg, Joseph H. 1960. A quantitative approach to the morphological typology of language. *International Journal of American Linguistics 26*(3): 178–194.

Greenberg, Joseph H. 1963. *The Languages of Africa*. Bloomington: Indiana University.

Greenberg, Joseph H. 1966. *Language Universals, With Special Reference to Feature Hierarchies*. The Hague: Mouton.

Harris, Alice C. 2010. Explaining typologically unusual structures: The role of probability. In Jan Wohlgemuth and Michael Cysouw, eds. *Rethinking Universals: How Rarities Affect Linguistic Theory*. Berlin: De Gruyter Mouton, 91–104.

Hinrichs, Lars, Benedikt Szmrecsanyi and Axel Bohmann. 2015. Which-hunting and the Standard English relative clause. *Language 91*(4): 806–836. doi:10.1353/lan.2015.0062.

Huber, Magnus. 2012. Syntactic and variational complexity in British and Ghanaian English: Relative clause formation in the written parts of the International Corpus of English. In Bernd Kortmann and Benedikt Szmrecsanyi, eds. *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin: Mouton De Gruyter, 218–242.

Hundt, Marianne. 1998. *New Zealand English Grammar – Fact or Fiction? A Corpus-Based Study in Morphosyntactic Variation* (Varieties of English Around the World G23). Amsterdam: John Benjamins.

Hyltenstam, Kenneth. 1984. The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Andersen, ed. *Second Languages*. Rowley, MA: Newbury, 39–58.

Kachru, Braj B., ed. 1992. *The Other Tongue: English across Cultures* (2nd ed.). (English in the Global Context). Urbana: University of Illinois Press.

Kortmann, Bernd, ed. 2004. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Kortmann, Bernd and Kerstin Lunkenheimer. 2012. *The Mouton World Atlas of Variation in English*. Berlin: de Gruyter Mouton.

Kortmann, Bernd and Kerstin Lunkenheimer, eds. 2013. *eWAVE*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://ewave-atlas.org/

Kortmann, Bernd and Benedikt Szmrecsanyi. 2004. Global synopsis: Morphological and syntactic variation in English. In Bernd Kortmann, Edgar Schneider, K. Burridge, R. Mesthrie and C. Upton, eds. *A Handbook of Varieties of English*, Vol. 2. Berlin: Mouton de Gruyter, 1142–1202.

Kortmann, Bernd and Benedikt Szmrecsanyi. 2009. World Englishes between simplification and complexification. In Lucia Siebers and Thomas Hoffmann, eds. *World Englishes – Problems, Properties and Prospects: Selected Papers from the 13th IAWE Conference*. Amsterdam: John Benjamins, 265–285.

Kortmann, Bernd and Benedikt Szmrecsanyi. 2011. Parameters of morpho-syntactic variation in World Englishes: Prospects and limitations of searching for universals. In Peter Siemund, ed. *Linguistic Universals and Language Variation*. Berlin: Mouton de Gruyter, 264–290.

Kortmann, Bernd and Benedikt Szmrecsanyi. 2012. *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin: Mouton de Gruyter.

Kortmann, Bernd, Edgar Schneider, Kate Burridge, Raj Mesthrie and Clive Upton, eds. 2004. *A Handbook of Varieties of English*. Berlin: Mouton de Gruyter.

Kruskal, Joseph B and Myron Wish. 1978. *Multidimensional Scaling* (Quantitative Applications in the Social Sciences). Newbury Park, CA: Sage Publications.

Kusters, Wouter. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Leiden: Leiden University.

Laporte, Samantha. 2012. Mind the gap! Bridge between World Englishes and Learner Englishes in the making. *English Text Construction* 5(2): 265–292. doi:10.1075/etc.5.2.05lap.

Lê, Sébastien, Julie Josse and François Husson. 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25(1): 1–18.

Leufkens, Sterre. 2013. The transparency of creoles. *Journal of Pidgin and Creole Languages* 28(2): 323–362.

Levshina, Natalia. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins.

Mair, Christian. 2003. Kreolismen und verbales Identitätsmanagement im geschriebenen jamaikanischen Englisch. In E. Vogel, A. Napp and W. Lutterer, eds. *Zwischen Ausgrenzung und Hybridisierung*. Würzburg: Ergon, 79–96.

Mair, Christian. 2013. The World System of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide 34*(3): 253–278. doi:10.1075/eww.34.3.01mai.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

McMahon, April, Paul Heggarty, Robert McMahon and Warren Maguire. 2007. The sound patterns of Englishes: Representing phonetic similarity. *English Language and Linguistics 11*(1): 113–142.

McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology 5*: 125–166.

Mesthrie, Rajend. 2006. Anti-deletions in an L2 grammar: A study of Black South African English mesolect. *English World-Wide 27*(2): 111–145.

Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki and Fred Karlsson, eds. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 23–41.

Miestamo, Matti, Kaius Sinnemäki and Fred Karlsson, eds. 2008. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins.

Murelli, Adriano and Bernd Kortmann. 2011. 28 Non-standard varieties in the areal typology of Europe. In Bernd Kortmann and Johan van der Auwera, eds. *The Languages and Linguistics of Europe*. Berlin: Mouton de Gruyter. www.degruyter.com/view/books/9783110220261/9783110220261.525/9783110220261.525.xml

Nerbonne, John and Peter Kleiweg. 2007. Toward a dialectological yardstick. *Journal of Quantitative Linguistics 14*(2): 148–166.

Nesselhauf, Nadja. 2009. Co-selection phenomena across New Englishes. *English World-Wide 30*(1): 1–26.

Newmeyer, Frederick J and Laurel B. Preston, eds. 2014. *Measuring Grammatical Complexity*. Oxford: Oxford University Press.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

Nichols, Johanna. 2009. Linguistic complexity: A comprehension definition and survey. In Geoffrey Sampson, David Gil and Peter Trudgill, eds. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 64–79.

Sampson, Geoffrey. 2001. *Empirical Linguistics*. London: Continuum Press.

Sampson, Geoffrey. 2009. A linguistic axiom challenged. In Geoffrey Sampson, David Gil and Peter Trudgill, eds. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 1–18.

Sampson, Geoffrey, David Gil and Peter Trudgill, eds. 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Company.

Schlegel, August Wilhelm von. 1818. *Observations sur la language et la littérature provençales*. Paris.

Schneider, Agnes. 2015. Aspect and modality in Ghanaian English: A corpus-based study of the progressive and the modal WILL. Unpublished doctoral dissertation, University of Freiburg.

Schneider, Edgar. 2004. Global synopsis: Phonetic and phonological variation in English world-wide. In Bernd Kortmann, Edgar Schneider, K. Burridge, R. Mesthrie and C. Upton, eds. *A Handbook of Varieties of English*, Vol. 1. Berlin: Mouton de Gruyter, 1111–1137.

Schneider, Edgar. 2007. *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge University Press.

Schreier, Daniel. 2016. A true split? Typological and sociolinguistic considerations on contact intensity effects. In Raffaela Baechler and Guido Seiler, eds. *Complexity, Isolation, and Variation* (Linguae and Litterae 57). Berlin, Boston: Mouton de Gruyter, 139–157.

Shackleton, Robert G. Jr. 2007. Phonetic variation in the traditional English dialects: A computational analysis. *Journal of English Linguistics 35*(1): 30–102.

Shosted, Ryan. 2006. Correlating complexity: A typological approach. *Linguistic Typology 10*. 1–40.

Siegel, Jeff, Benedikt Szmrecsanyi and Bernd Kortmann. 2014. Measuring analyticity and syntheticity in creoles. *Journal of Pidgin and Creole Languages 29*(1): 49–85. doi:10.1075/jpcl.29.1.02sie.

Siemund, Peter. 2008. Language contact: Constraints and common paths of contact-induced language change. In Peter Siemund and Noemi Kintana, eds. *Language Contact and Contact Languages*. Amsterdam: John Benjamins, 3–11.

Siemund, Peter. 2013. *Varieties of English: A Typological Approach*. Cambridge: Cambridge University Press.

Peter Siemund and Noemi Kintana, eds. 2008. *Language Contact and Contact Languages*. Amsterdam: John Benjamins.

Sinnemäki, Kaius. 2011. *Language Universals and Linguistic Complexity: Three Case Studies in Core Argument Marking*. Helsinki: University of Helsinki.

Sinnemäki, Kaius. 2014. Complexity trade-offs: A case study. In Frederick J. Newmeyer and Laurel B. Preston, eds. *Measuring Grammatical Complexity*. Oxford: Oxford University Press, 179–201.

Steger, Maria and Edgar W Schneider. 2012. Complexity as a function of iconicity: The case of complement clause constructions in New Englishes. In Bernd Kortmann and Benedikt Szmrecsanyi, eds. *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin: de Gruyter.

Szmrecsanyi, Benedikt. 2009. Typological parameters of intralingual variability: Grammatical analyticity versus syntheticity in varieties of English. *Language Variation and Change 21*(3): 319–353. doi:10.1017/S0954394509990123.

Szmrecsanyi, Benedikt. 2012. Typological profile: L1 varieties. In Bernd Kortmann and Kerstin Lunkenheimer, eds. *The Mouton World*

Atlas of Variation in English. Berlin: de Gruyter, 826–843. www.degruyter.com/view/books/9783110280128/9783110280128.826/9783110280128.826.xml

Szmrecsanyi, Benedikt. 2017. Featurometry. In Martijn Wieling, Gosse Bouma and Geertjan van Noord, eds. *From Semantics to Dialectometry: Festschrift in Honor of John Nerbonne*. Berlin: Language Science Press.

Szmrecsanyi, Benedikt and Bernd Kortmann. 2009a. Vernacular universals and angloversals in a typological perspective. In Markku Filppula, Juhani Klemola and Heli Paulasto, eds. *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*. London: Routledge, 33–53.

Szmrecsanyi, Benedikt and Bernd Kortmann. 2009b. Between simplification and complexification: Non-standard varieties of English around the world. In Geoffrey Sampson, David Gil and Peter Trudgill, eds. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 64–79.

Szmrecsanyi, Benedikt and Bernd Kortmann. 2009c. The morphosyntax of varieties of English worldwide: A quantitative perspective. *Lingua 119*(11): 1643–1663.

Szmrecsanyi, Benedikt and Bernhard Wälchli, eds. 2014. *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech* (Lingua and Litterae 28). Berlin: Walter de Gruyter.

Trudgill, Peter. 1996. Dialect typology: Isolation, social network and phonological structure. In Gregory R. Guy, Crawford Feagin, Deborah Schiffrin and John Baugh, eds. *Towards a Social Science of Language: Papers in Honor of William Labov, Vol. 1: Variation and Change in Language and Society*. Amsterdam: John Benjamins, 3–22.

Trudgill, Peter. 2001. Contact and simplification: Historical baggage and directionality in linguistic change. *Linguistic Typology 5*: 371–374.

Trudgill, Peter. 2004. Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology 8*: 305–320.

Trudgill, Peter. 2006. Dialect mixture versus monogenesis in colonial varieties: The inevitability of Canadian English? *Canadian Journal of Linguistics-Revue Canadienne De Linguistique 51*: 265–286. doi:10.1353/cjl.2008.0002.

Trudgill, Peter. 2008. The role of dialect contact in the formation of Englishes. In Miriam A. Locher and Jürg Strässler, eds. *Standards and Norms in the English Language*. Berlin: Mouton de Gruyter, 69–83.

Trudgill, Peter. 2009a. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil and Peter Trudgill, eds. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 98–109.

Trudgill, Peter. 2009b. Vernacular universals and the sociolinguistic typology of English dialects. In Marrku Filppula, Juhani Klemola and Heli Paulasto, eds. *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*. London: Routledge, 302–329.

Trudgill, Peter. 2010. Contact and sociolinguistic typology. In Raymond Hickey, ed. *The Handbook of Language Contact*. Malden, MA: Wiley-Blackwell, 299–319.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Trudgill, Peter. 2016. The sociolinguistics of non-equicomplexity. In Raffaela Baechler and Guido Seiler, eds. *Complexity, Isolation, and Variation*. Berlin: Mouton de Gruyter, 159–170.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis and Margaret MacLagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics 36*(2): 299–318.

Wohlgemuth, Jan and Michael Cysouw, eds. 2010. *Rara and Rarissima: Documenting the Fringes of Linguistic Diversity* (Empirical Approaches to Language Typology [EALT] 46). Berlin: de Gruyter Mouton.