Benedikt Szmrecsanyi and Bernd Kortmann

Introduction: Linguistic complexity

Second Language Acquisition, indigenization, contact

The present volume endeavors to summarize a workshop held at the Freiburg Institute for Advanced Studies (FRIAS) in May 2009 about »Linguistic complexity in interlanguage varieties, L2 varieties, and contact languages«. The presentations at the workshop, and the essays in this volume, reflect the intricacies of thinking about complexity in three major contact-related fields of (and schools in) linguistics: creolistics, indigenization and nativization studies (i.e. in the realm of English linguistics, the »World Englishes« community), and Second Language Acquisition (SLA) research. The authors in this volume are all leading linguists in different areas of specialization, and they were asked to elaborate on those facets of linguistic complexity which are most relevant in their area of specialization, and/or which strike them as being most intriguing. The result is a collection of papers that, when viewed collectively, may raise more questions than it does answer – but such is the way good science advances. And the fact of the matter is that linguistic complexity is complicated.

This general introduction is structured as follows. In Section 1, we review the literature on linguistic complexity to provide a backdrop against which the contributions in this volume are to be read. In Section 2, we sketch the objectives of this volume. Section 3 provides summaries of the contributions in this book, and contextualizes them in the context of the volume as a whole.

1. Linguistic complexity: a review of the research literature

In contrast to SLA researchers, who have always considered linguistic complexity a gauge for learners' proficiency in the target language, a descriptor for performance, and an index to benchmark development, and thus an intrinsically variable concept (see, for example, Ortega, this volume and Larsen-Freeman 1978), for most of the twentieth century most theoretical linguists have agreed that on the level of *langue*, all natural languages must be equally complex. Hence, linguistic complexity was supposed to be invariant,

kort_001.pod 7
07-01-05 16:25:03 -mlb- mlb

such that there were no ›simple‹ or ›complex‹ languages. This consensus – which was an article of faith rather than an insight backed up by empirical evidence – has been dubbed the *ALEC statement* (»*A*ll *L*anguages are *E*qually *C*omplex«) (Deutscher 2009: 243), or the *linguistic equi-complexity dogma* (Kusters 2003: 5). For a detailed and illuminating review of the history of thought on linguistic complexity and the equi-complexity dogma, we refer the reader to Sampson (2009). Suffice it to say here that the basic idea behind the equi-complexity dogma is that the total complexity of a language is fixed because sub-complexities in linguistic sub-systems trade off. Accordingly, simplicity in some domain A must be compensated by complexity in domain B, and vice versa. This view is nicely articulated in a 1958 textbook by descriptivist Charles Hockett (quoted in, e.g., Sampson 2009):

> Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology and syntax, is about the same as that of any other. This is not surprising, since all languages have about equally complex jobs to do, and what is not done morphologically has to be done syntactically. Fox, with a more complex morphology than English, thus ought to have a somewhat simpler syntax; and this is the case. (Hockett 1958: 180–181)

The popularity of the equi-complexity dogma during the twentieth century may be seen against the backdrop of the nineteenth century, when linguists such as Wilhelm von Humboldt put forward somewhat unfortunate claims to the effect that differences (in terms of complexity or otherwise) between languages can be traced back to the differential mental capacities of their speakers (Humboldt 1836: 37). One reason for the dominance of the equi-complexity dogma in the twentieth century, then, was that it meshed well with more modern and egalitarian perspectives, and specifically with the idea that all human speakers are endowed with the same mental, cultural, and biological capacities.

But be that as it may, the beginning of the twenty-first century has witnessed a number of critical reviews of the idea that all languages are, under all circumstances, equally complex (see Karlsson, Miestamo, and Sinnemäki 2008 for a detailed survey). This iconoclasm has no doubt been in the air for a while, but one of the primers that really got the debate going was a special issue of the journal *Linguistic Typology* (2001, volume 5, issue 2/3). In the issue's lead article, John McWhorter suggested, somewhat provocatively, that »the world's simplest grammars are creole grammars«:

> Creole languages are unique in having emerged under conditions which occasioned the especial circumstance of stripping away virtually all of a language's complexity (as defined in this paper), such that the complexity emerging in a

kort_001.pod 8
07-01-05 16:25:03 -mlb- mlb

> creole is arising essentially from ground zero, rather than alongside the results of tens of thousands of years of other accretions. As such, creoles tend strongly to encompass a lesser degree of complexity than any older grammar. (McWhorter 2001: 155)

In the same special issue, a range of luminaries responded to this claim – some concurring, some strongly objecting. In retrospect, it seems to us that many of the most vigorous dissenters were actually beholden to the Humboldtian notion that ›complex is beautiful, simple is retarded‹, an equation that few modern complexity analysts would subscribe to (on the contrary!). In this connection, we would like to emphasize that present-day challengers of the equi-complexity dogma do not, of course, resort to nineteenth century racism to posit or explain complexity variance. Instead, many modern complexity analysts seek to model variable complexity as a function of language-external variables (see Section 1.3), which all essentially boil down to ›cultural‹ factors (Deutscher 2010) in the broadest sense.

In subsequent years the discourse on linguistic complexity and the equi-complexity dogma has been nourished by several dedicated workshops and conferences (for example, the 2007 workshop on »Language Complexity as an Evolving Variable«, hosted by the Max Planck Institute for Evolutionary Anthropology in Leipzig,), a number of pertinent journal papers (e.g. Shosted 2006; Trudgill 2004), monographs (e.g. Dahl 2004; Kusters 2003), and edited volumes (e.g. Miestamo, Sinnemäki, and Karlsson 2008; Sampson, Gil, and Trudgill 2009). A succinct review of this literature will be presented in what follows. This review will be selective in that we shall focus on synchronic complexity (in)variance on some level of aggregation (language, dialects, or speaker communities) rather than, say, differences between individuals (e.g. Chipere 2009), or the evolutionary genesis of complexity (e.g. Givón 2009).

1.1. Local complexity versus global complexity

There is a sense in much of the literature that we need to distinguish carefully between, in the parlance of Miestamo (2008), (1) *global linguistic complexity* (that is, complexity of a language, dialect, etc. as such) and (2) *local linguistic complexity* (i.e. domain-specific linguistic complexity). While assessing a language's global complexity is a very ambitious and indeed probably hopeless endeavor (akin to a »wild goose chase« [Deutscher 2009, 2010], measuring local complexities in linguistic subdomains is seen as a more doable task. Previous scholarship has been concerned with measures of local complexity in the following linguistic subdomains:

kort_001.pod 9
07-01-05 16:25:03 -mlb- mlb

Phonological complexity (e.g. Nichols 2009; Shosted 2006). Aspects investigated include the size of the phoneme inventory, the incidence of marked phonemes, tonal distinctions, suprasegmental phonology, phonotactic restrictions, and the maximum complexity of consonant clusters.

Morphological complexity (e.g. Dammel and Kürschner 2008; Kusters 2003). What is the scope of inflectional morphology in a given language or language variety? Specifically, what is the extent to which we are dealing with allomorphies and morphophonemic processes as classical complexity-inducing »nuisance factors« (Braunmüller 1990: 627) not only in inflectional morphology but also in phonology?

Syntactic complexity (e.g. Givón 2009; Karlsson 2009). How many rules are mandated by the syntax of a language (motto: the more, the more complex), and what is the level of clausal embedding and recursion mandated and/or permitted by a language? Observe that the degree of embedding is incidentally also by far the most popular metric of complexity in SLA research (see Ortega 2003).

Semantic and lexical complexity (e.g. Fenk-Oczlon and Fenk 2008; Nichols 2009). How many homonymous and polysemous lexemes or phrases are we dealing with? Do personal pronouns have an inclusive/exclusive distinction? We note in this connection that establishing type-token-ratios as a customary SLA tool to assess lexical variation in corpora (for example, Kuiken and Vedder 2008) is essentially also a lexical complexity metric.

Pragmatic complexity, also known as »hidden complexity« (Bisang 2009). What is the degree of pragmatic inferencing on the part of hearers mandated by a given language?

Some studies have suggested that high levels of local complexity in some domain of a given language do not necessarily entail low local complexity in some other domain of that same language, as the equi-complexity dogma would predict. Shosted (2006), for example, investigates morphological and phonological complexity in a sample of 34 languages and finds that there is no significant correlation, and no trade-off, between morphological and phonological complexity scores; Nichols (2009) likewise explores local complexities in a wide range of languages and fails to obtain a trade-off. Fenk-Oczlon and Fenk (2008: 63) do diagnose »balancing effects« between local complexities, but not to the extent that the equi-complexity dogma could be assumed to hold true under all circumstances. Gil (2008) argues that isolating

kort_001.pod 10
07-01-05 16:25:03 -mlb- mlb

languages do not necessarily compensate for simpler morphology with more complexity in other domains.

1.2. Complexity notions and measures

Expectably, there is a daunting variety of ways in which analysts have sought to assess linguistic complexity. In most general terms, these measures fall into two categories (Dahl 2009; Lindström 2008; see, e.g., Miestamo 2008): *measures of absolute complexity* (theory-oriented and thus ›objective‹) versus *measures of relative complexity* (user-oriented and thus ›subjective‹, about ›cost‹ and ›difficulty‹). Miestamo defines the difference as follows:

> The absolute approach defines complexity in objective terms as the number of parts in a system, of connections between different parts, etc. […] The relative approach to complexity defines complexity in relation to language users: what is costly or difficult to language users (speakers, hearers, language learners) is seen as complex. Complexity is thus identified with cost and difficulty of processing and learning. (Miestamo 2009: 81)

Notice that the absolute approach is popular especially in the cross-linguistic typology camp, while most sociolinguistically and psycholinguistically inclined complexity analysts have a knack for the relative approach. Note also that some analysts (for example, P. Mühlhäusler 1974, 1992) have wondered whether there are viable language-neutral definitions of ›complexity‹ and ›simplicity‹ (or ›complexification‹ and ›simplification‹) at all.

We move on to a discussion of four more specific complexity notions and measures that have served as the workhorses in extant complexity scholarship: (1) *absolute-quantitative complexity*, (2) *redundancy-induced complexity*, (3) *irregularity-induced complexity*, and (4) *L2 acquisition complexity*.[1]

1.2.1. Absolute-quantitative complexity

Many studies marshal a genuinely absolute-quantitative measure of complexity which restricts attention to what McWhorter (for example, in this volume) refers to as ›structural elaboration‹. Arends has dubbed this notion the ›more is more complex‹ approach:

---

[1] Other complexity/simplicity measures that have been proposed but not explored extensively include *processing complexity* (J. A. Hawkins 2009; Seuren and Wekker 1986); *morphological naturalness* (Mayerthaler 1981), and *information-theoretic complexity* (Juola 2008; Sadeniemi et al. 2008).

kort_001.pod 11
07-01-05 16:25:03 -mlb- mlb

> A grammar is judged to be more complex if it has more (marked) phonemes, more tones, more syntactic rules, more grammatically expressed semantic and/or pragmatic distinctions, more morphophonemic rules, more cases of suppletion, allomorphy, agreement. Qualitative aspects of complexity, such as the internal complexity of the rules themselves, are not taken into account. (Arends 2001: 180)

Students of absolute-quantitative complexity have been interested in, e.g., the number of grammatical categories in a language (Shosted 2006), the number of phonemic contrasts (McWhorter 2001), or the length of the minimal description of a linguistic system (Dahl 2004).[2]

1.2.2. Redundancy-induced complexity

This notion is a hybrid between relative and absolute complexity. What takes center stage here is not the absolute number of linguistic elements in a system, but the extent of »overspecification« (McWhorter, this volume): the prevalence of redundant features, in the sense of communicatively dispensable elements that are »incidental to basic communication« (McWhorter 2001: 126). Such material has been called, somewhat polemically, »historical baggage« (Trudgill 1999: 149), »ornamental elaboration« (McWhorter 2001: 132), or »baroque accretion[s]« (McWhorter 2001: 126). In short, redundancy-induced complexity notions are concerned with linguistic elements that are in the language for no apparent reason other than historical accident. Such features may include

> ergativity, grammaticalized evidential marking, inalienable possessive marking, switch-reference marking, inverse marking, obviative marking, »dummy« verbs, syntactic asymmetries between matrix and subordinate clauses, grammaticalized subjunctive marking, verb-second, clitic movement, any pragmatically neutral word order but SVO, noun class or grammatical gender marking (analytic or affixal), or lexically contrastive or morphosyntactic tone […] (McWhorter 2001: 163)

Note that redundancy-induced complexity is an absolute notion because theory dictates what should count as redundant and what not. At the same time, however, redundancy-induced complexity is a relative concept because it is arguably language users, not the analyst, who bear the cost of this redundancy.

---

[2] In this connection, it is interesting to note that the design of Basic English, an international auxiliary language (B. Mühlhäusler and P. Mühlhäusler 2005; Ogden 1934), appears to be primarily indebted to the ›more is more complex‹ notion – its most prominent, supposedly simplifying design feature is its limited lexicon.

kort_001.pod 12
07-01-05 16:25:03 -mlb- mlb

1.2.3. Irregularity-induced complexity

Irregularity-induced complexity (a.k.a. ›non-transparency‹; see Nichols forthcoming) is a hybrid concept as well that to some extent combines absolute and relative complexity notions. The idea is to classify as complex the outcome of irregular inflectional and derivational processes, which are opaque and non-transparent (see also McWhorter, this volume; P. Mühlhäusler 1974; Trudgill 2004):

> as often as not inflection leads to the development of morphophonological processes, which constitute an added component of a grammar to be learned […] Meanwhile, suppletion also complexifies an area of grammar according to our metric. The various suppletive forms of *be* in English (*am, are, is, was, were, been*, and similarly in most Indo-European languages) render these languages more complex in this area than languages where the copula is invariable across person and number in the present (McWhorter 2001: 137)

Irregularity-based complexity is an absolute notion in that the definition of what should count as ›irregular‹ is theory-driven, but once again the cost associated with such irregularity is incurred by language users.

1.2.4. L2 acquisition complexity

As a genuinely relative complexity measure, L2 acquisition complexity – a metric which is particularly popular in the sociolinguistic community – is concerned with the degree to which a language or language variety (or some aspect of a language or language variety) is difficult to acquire for adult language learners. Thus, Kusters (2003) (see also Kusters 2008) defines

> complexity as the amount of effort an outsider has to make to become acquainted with the language in question […] an outsider is someone who learns the language in question at a later age, and is not a native speaker. Therefore, phenomena that are relatively difficult for a second language learner in comparison with a first language learner are more complex […] Phenomena that are easy to acquire for a second language learner but difficult for a first language learner are less complex. (Kusters 2003: 6)

In exactly the same spirit, Trudgill (2001: 371) defines linguistic complexity as being commensurate with »difficulty of learning for adults«.

1.3. Language-external conditioning factors

Much of the recent discourse on linguistic complexity (in)variance is oriented towards the question of whether and to what extent linguistic complexity is a function of certain language-external factors. On the most general level, Deutscher (2010) has argued that *culture* is partly to blame for com-

kort_001.pod 13
07-01-05 16:25:03 -mlb- mlb

plexity variance – for example, complex, literate cultures need complex, sizable vocabularies (Deutscher 2010: 110). More narrowly conceptualized proposals include McWhorter's claim that the *age of a language* can affect its complexity (see also Parkvall 2008). In this perspective, the world's simplest languages are creole languages because

> 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)

Other researchers have suggested that while older languages may indeed be more complex than younger languages (all other things being equal), a history of *language contact and concomitant adult SLA* may render a *per se* old language simpler than it otherwise would have been (see also McWhorter 2008; Thomason 2001). The reason is that, for one thing, speakers often »simplify their native language for specific social purposes in contact situations« (Thomason and Kaufman 1988: 177). Secondly, adult language learners regularly adopt simplifying strategies (Selinker 1972) to avoid irregularities and to increase transparency in the sense of Seuren and Wekker (1986) (cf. also Klein and Perdue 1997; Muysken and Smith 1995). Trudgill succinctly sums up this argument as follows:

> Adult 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. (Trudgill 2001: 372)

A third line of research has endeavored to show that linguistic complexity is sensitive to certain sociological parameters. Sinnemäki (2008, 2009), for example, demonstrates that in a typological sample of fifty languages, population size correlates with complexity in core argument marking. In much the same vein, Trudgill shows that »small, isolated, low-contact communities with tight social network structures« (2004: 306) tend to have more complex languages than high-contact communities. Lastly, Nichols (forthcoming) suggests that linguistic complexity is to some extent predicted by geographic altitude, a first-rate proxy for the degree of isolation of a speaker community.

1.4. Data sources

It is fair to say that much of the recent literature on linguistic complexity is based on rather unsystematic and intuition-based, if not anecdotal, evidence. To the extent that claims are based on real data, we find that most researchers interested in cross-linguistic variation have tapped reference grammars in conjunction with typological sampling techniques (for example,

kort_001.pod 14
07-01-05 16:25:03 -mlb- mlb

Miestamo 2009; Sinnemäki 2009). Parkvall (2008) accesses a major typological database, the *World Atlas of Language Structures* (Haspelmath et al. 2005). Analysts interested in language-internal variation typically rely on fieldwork and standard dialectological data sources such as dialect atlases (for example, Trudgill 2009). Few researches have based claims on major, balanced naturalistic text corpora (but see, e.g., Karlsson 2009). Using an approach that is

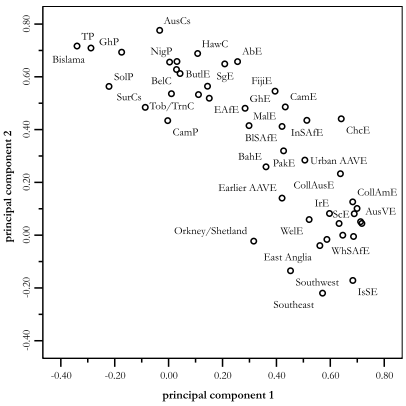

Figure 1. Locating varieties of English in a two-dimensional plane (adapted from Szmrecsanyi and Kortmann 2009a: Fig. 4)

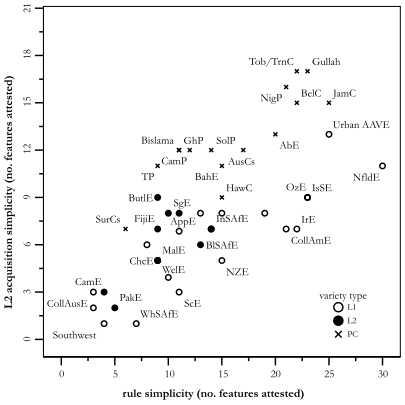

Figure 2. Rule simplicity by L2 acquisition simplicity in varieties of English (adapted from Kortmann and Szmrecsanyi 2009: Figure 1)

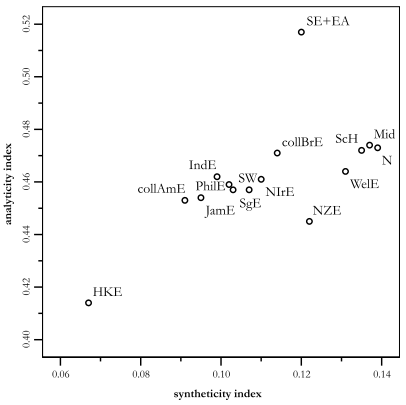

Figure 3. Analyticity by syntheticity in varieties of English (adapted from Szmrecsanyi and Kortmann 2009c: Fig. 5.1)

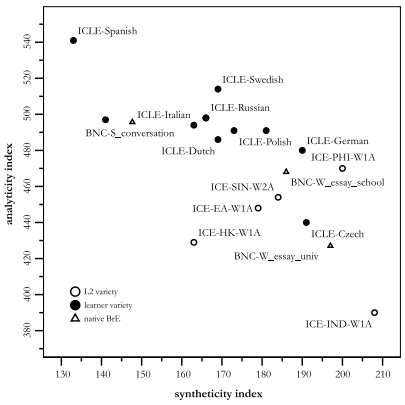

Figure 4. Analyticity by syntheticity: indigenized L2 varieties versus learner varieties (adapted from Szmrecsanyi and Kortmann 2011: Figure 1)

kort_001.pod 15
07-01-05 16:25:03 -mlb- mlb

corpus-linguistic in spirit, Juvonen (2008) studies Chinook Jargon, a pidgin language, on the basis of a fictional text. Some analysts have also explored Bible translations (e.g. Dahl 2008). Gil (2008) is a rare example of a complexity study that uses an experimental design.

1.5. The Freiburg perspective

The Freiburg program endeavors to marry methodologies and interpretational approaches familiar from the study of large-scale cross-linguistic variation to the analysis of language-internal variation (see Kortmann 2004 for a collection of papers in this spirit). It is precisely in this tradition that recent Freiburg-based work has sought to explore linguistic complexity variance in varieties of English, for the sake of understanding the comparatively simple (i.e. language-internal complexity variation) before approaching the comparatively complicated (i.e. cross-linguistic complexity variation). Our research agenda has had two specific objectives: first, to elucidate large-scale patterns of (complexity) variance in varieties of English and in English-based pidgin and creole languages; and second, to develop the necessary metrics for passing judgment on degrees of complexity and claims concerning processes of simplification and complexification of grammars.

A first line of research has tapped into the morphosyntactic survey of the multimedia reference tool that accompanies the *Handbook of Varieties of English* (Kortmann et al. 2004). Kortmann and Szmrecsanyi (2004) describe the survey design in ample detail; suffice it to say here that the survey describes 46 varieties of English (specifically, 20 L1 varieties, 11 indigenized L2 varieties of English, and 15 English-based pidgin and creole languages) in terms of the presence or absence of 76 non-standard morphosyntactic features, which fall into eleven broad areas of morphosyntax: pronouns, the noun phrase, tense and aspect, modal verbs, verb morphology, adverbs, negation, agreement, relativization, complementation, and discourse organization and word order. What can this database tell us about complexity variance in varieties of English? To start with, when applying satellite-view exploratory statistical analysis techniques to the dataset, it turns out that varieties of English can be arranged along two major dimensions. In Figure 1, for example, we find a plot that uses Principal Component Analysis to assign each variety of English a coordinate in two-dimensional principal component space. The picture that emerges is that varieties of English cluster very neatly according to variety type: English-based pidgin and creole languages cluster in the top-left quadrant while L1 varieties cluster in the bottom-right quadrant; indigenized L2 varieties of English cover the middle ground. The meaning of the

kort_001.pod 16
07-01-05 16:25:03 -mlb- mlb

axes can be interpreted as follows. Component 2 (the vertical axis in Figure 1) plots a given variety's degree of *analyticity*, a notion that is defined as bringing about a greater number of features that are autonomous, invariable and periphrastic in nature. Component 1 (the horizontal axis in Figure 1) is taken to indicate increased *morphosyntactic complexity*, which is broadly defined in the ›more is more complex‹ spirit following McWhorter (2001). According to this interpretation, then, English-based pidgin and creole languages are least complex, L1 varieties are most complex, and indigenized L2 varieties of English exhibit intermediate complexity. Digging deeper into complexity variance in the morphosyntax survey, Kortmann and Szmrecsanyi define three more compartmentalized complexity measures:

Ornamental rule complexity (Kortmann and Szmrecsanyi 2009; Szmrecsanyi and Kortmann 2009c): the relative incidence in a variety's inventory of features that add contrasts, distinctions, or asymmetries without providing a communicative or functional bonus (a definition that is, of course, reminiscent of McWhorter 2001's approach). Prime examples of such features include *she/her* used for inanimate referents, or *be* as a perfect auxiliary.

Rule simplicity (Kortmann and Szmrecsanyi 2009): as the mirror image of ornamental rule complexity, this measure is concerned with the relative incidence in a variety's inventory of features that simplify the system, vis-à-vis the standard system. An example of a simplifying feature would be regularized reflexives-paradigms (where oppositions in the Standard English system are leveled) or the loosening of the sequence of tenses rule (where a rule in the Standard English system is dispensed with).

L2 acquisition simplicity (Kortmann and Szmrecsanyi 2009; Szmrecsanyi and Kortmann 2009c): as a relative complexity metric, this measure establishes the relative incidence in a variety's inventory of features that SLA research has shown to recur in interlanguage varieties. Examples include invariant *don't* in the present tense (learners tend to avoid inflection; see, e.g., Klein and Perdue 1997), and regularization of irregular verb paradigms (learners tend to overgeneralize [e.g. Towell and R. Hawkins 1994]).

The results in terms of ornamental rule complexity are fairly clear-cut. While traditional low-contact L1 vernaculars (e.g. traditional dialects in the British Isles) attest between two and three ornamentally complex features on aver-

kort_001.pod 17
07-01-05 16:25:03 -mlb- mlb

age, high-contact L1 varieties, indigenized L2 varieties, and English-based pidgin and creole languages typically have only about one ornamentally complex feature in their inventory. Thus, ornamental rule complexity appears to be a function of the degree of contact (and of a history of L2 acquisition among adults), which is in line with, e.g., Trudgill (2001: 372). As for rule simplicity and L2 acquisition simplicity, consider Figure 2, which plots the two simplicity indices against each other. As with Figure 1, we find variety clusters such that English-based pidgin and creole languages attest many simple features, while L1 varieties and, curiously, indigenized L2 varieties of English seem to attest few simplifying features as well as few L2-simple features.

A second line of research pursued in Freiburg has based claims about language-internal complexity variance in English not on survey data but – drawing on a method in quantitative morphological typology originally suggested by typologist Joseph Greenberg (1960) – on frequency vectors extracted from naturalistic corpus data. In this research, four frequency-based complexity indices are distinguished (Kortmann and Szmrecsanyi 2011; Szmrecsanyi 2009; Szmrecsanyi and Kortmann 2009c):

An *overt grammatical analyticity index*, which measures the text frequency (normalized to 1,000 words of running text) of free grammatical markers (also known as function words).

An *overt grammatical syntheticity index*, which measures the normalized text frequency of bound grammatical markers, i.e. inflections (recall that synthetic marking is supposedly particularly complex, thanks to inevitable allomorphies and the fact that learners tend to avoid inflectional marking).

An *irregularity index*, which establishes the proportion of bound grammatical markers that are irregular and lexically conditioned (e.g. *catch* > *caught*, as opposed to *catch* > *catched*).

An *overt grammaticity* index, which is concerned with the normalized text frequency of any grammatical markers, analytic or synthetic.

The findings can be summarized as follows. Consider Figure 3, which plots syntheticity levels of a range of varieties of English for which naturalistic corpus data are available against their analyticity levels. We find that traditional, low-contact British English dialects (e.g. East Anglian English) tend to be both fairly analytic and synthetic, while some ›deleting‹ (Mesthrie 2006) indigenized L2 varieties of English (e.g. Hong Kong English) tend to be

kort_001.pod 18
07-01-05 16:25:03 -mlb- mlb

neither particularly synthetic nor particularly analytic. Overall, the generalization that emerges from Figure 3 is that there is a theoretically interesting *positive* correlation between analyticity and syntheticity levels, instead of a trade-off. This is another way of saying that varieties of English differ primarily in terms of their *grammaticity* (that is, the text frequency of grammatical marking, analytic or synthetic): thus, L1 varieties exhibit above-average grammaticity, while indigenized L2 varieties exhibit below-average grammaticity (and do not selectively avoid synthetic marking only). As for the irregularity index, it turns out that in indigenized L2 corpus material, 82 per cent of all bound grammatical allomorphs are regular, a percentage that goes down to 71 per cent in the case of high-contact L1 varieties and to 65 per cent in traditional L1 vernaculars. In other words, indigenized L2 varieties are least complex and most transparent morphologically whereas traditional, low-contact L1 vernaculars are most complex and least transparent – again, in line with what one would expect from perusing the literature. Szmrecsanyi and Kortmann (2011) advance this approach[3] to cover differences between written registers in indigenized L2 varieties as sampled in the *International Corpus of English* (ICE; Greenbaum 1996) and the written material sampled in the Louvain *International Corpus of Learner English* (ICLE; Granger 2003). In short, the name of the game is comparing essays written by, say, Hong Kong English writers to essays written by e.g. advanced Spanish learners of English. Figure 4 thus arranges the data points in a customary two-dimensional analyticity-syntheticity plane. The fact of the matter is that in this perspective, learner varieties and indigenized L2 varieties are clearly different beasts in that learner Englishes are significantly more analytic than indigenized L2 varieties of English. There is also a notable tendency for indigenized L2 varieties to be more synthetic than learner Englishes. In addition, Szmrecsanyi and Kortmann (2011) show that the complexity profiles of learner Englishes (for example, German learner English) cannot necessarily be predicted by the complexity profile of learners' native languages (e.g. German). In other words, there do not appear to be reliable substrate effects.

2. The present volume

There is a new angle from which this volume tries to shed light on the multifacetted and highly controversially discussed concept of linguistic complexity. This angle is determined by three independent, but clearly related, lan-

[3] For further extensions, see Ehret (2008) and Szmrecsanyi (2009) on text type variability, and Szmrecsanyi (2012) on long-term historical variability.

kort_001.pod 19
07-01-05 16:25:03 -mlb- mlb

guage-external factors conditioning linguistic complexity, namely language contact (including dialect contact), adult language acquisition, and indigenization. The present volume thus meets the increasingly voiced need for approaches integrating what we know about native Englishes, non-native Englishes (including English-based pidgin and creole languages), and learner Englishes with what we know about the processes involved in SLA, indigenization and creolization, in general (for studies in the same spirit compare, for example, Davydova to appear a,b). In doing so, it brings together for the first time leading representatives of three fields of research in each of which, as we have shown in the literature review, the notion of linguistic complexity has figured in recent debates. More specifically, the authors were asked to address one or more of the following three questions in their papers:

1. How can we adequately assess complexity in interlanguages (e.g. Finnish learner English), indigenized L2 varieties (e.g. Hong Kong English), and pidgin and creole languages (e.g. Jamaican Creole)?
2. Should we be interested in absolute complexity or rather relative complexity? (One may spontaneously assume that most authors would go for a relative complexity measure, given the nature of the languages under consideration in this volume.)
3. What is the extent to which language contact and/or (adult) language learning might lead to morphosyntactic simplification? In which contact situations does it lead to morphosyntactic simplification, at all? (As presented in Section 1.5 above, simplification in grammar clearly correlates with a high degree of language and dialect contact for varieties of English. However, from these results it must not be concluded that intensive language contact necessarily results in simplification processes. Sarah Thomason has repeatedly stated [p.c.] that complexification processes are equally observable among contact languages and contact varieties.)

All authors were asked to provide explicit statements on what exactly they understand by ›complexity‹, how to measure complexity, and which role complexity plays in their line of research or their research framework for the varieties they are interested in.

It so happens that all papers in this volume primarily address varieties of English and English-based pidgin and creole language. We believe, however, that this does not impinge on the validity of the general points concerning many of the crucial issues sketched above. Ultimately, this volume offers a unique opportunity to explore both the different views on, but also the degree of common ground among three research communities each of which is concerned with linguistic complexity in, broadly speaking, (current or past) conditions of overcoming the language barrier.

kort_001.pod 20
07-01-05 16:25:03 -mlb- mlb

3. Summaries of the contributions in this volume

A perfect representative of the present volume's inter-disciplinary spirit, **Jeff Siegel**'s »Accounting for Analyticity in Creoles« draws on evidence and observations deriving from research on SLA, interlanguage varieties, indigenized L2 varieties, restricted pidgins, and expanded pidgin and creole languages. The contribution is a meticulously argued piece of detective work illuminating why creole languages are, as a rule, pervasively analytic and therefore, by Siegel's reckoning, non-complex. To fix terminology, Siegel distinguishes between two types of complexity, *componential complexity* (an absolute-quantitative notion: more components will render a language more complex) and *structural complexity* (the degree to which a grammar is hard to understand or analyze – hence, a relative notion); the contribution is concerned with morpho-grammatical (and thus, local) structural complexity. Siegel proceeds from the widely accepted assumption that all other things being equal, analytic morphemes are simpler than synthetic morphemes thanks to the perceptual salience and concomitant acquisitional ease that comes with analytic marking (see McWhorter's contribution for a critical stance towards this view). But what is the exact role that SLA plays in the genesis of creole analyticity? In principle, analytic simplicity in creole languages may be the consequence of (i) a process of reduction (in Siegel's parlance, *reductive simplification* with lexifier speaker agentivity), or (ii) a lack of development (*developmental simplicity,* with substrate speaker agentivity). Siegel points out that (ii) is the more likely scenario. However, if normal adult SLA were implicated in the process, it would follow that any analytic creole morphology is simply the product of the expansion of an earlier interlanguage, in which case creole languages should exhibit certain traits they share with interlanguages and indigenized L2 varieties – a corollary that, according to Siegel, does not always mesh with the facts. The interim conclusion is, therefore, that adult SLA cannot entirely account for the analytic nature of creole languages. Instead, Siegel offers that what may better explain pervasive creole analyticity is the notion of *functional transfer*, defined as a process where agents apply »the properties of a grammatical morpheme of one language to a syntactically congruent lexical item of another language«. Unlike in the case of indigenized L2 varieties, in the genesis of which agents may have drawn on material (possibly synthetic) in the target language for expanding an interlanguage, speakers of restricted pidgins may not have that sort of access to the lexifier. Instead, speakers may tap into knowledge of their first language(s) via functional transfer, thus mapping grammatical information on lexical items drawn from the lexifier. The outcome is pervasive creole analyticity.

kort_001.pod 21
07-01-05 16:25:03 -mlb- mlb

Like Jeff Siegel, **Terence Odlin** in »Nothing will come of nothing« is also interested in transfer effects, though Odlin's particular focus is on theoretical value of the interface between transfer effects and linguistic complexity for SLA research. The specific complexity notion that Odlin explores is *descriptive complexity* (see also Mühlhäusler's contribution), an absolute-quantitative notion that defines complexity as being proportional to the length of the minimal adequate description of the system being described (cf. Dahl 2004; Rescher 1998). Observe here that Odlin's idea of complexity also partially overlaps with Kusters' (2008) idea of *outsider complexity*, the crucial difference being that while Kusters abstracts away from transfer effects, Odlin is explicitly interested in them. The contribution investigates preposition and article usage in written learner English by native speakers of Finnish and Swedish in Finland, two learner groups with native languages that are typologically quite different (Swedish being related to English, and Finnish not at all), but with minimal cultural differences. The paper relies on a cross-sectional dataset collected by Scott Jarvis that documents task-oriented written learner English (movie retells, to be precise) by different learner groups (Finnish-speaking Finns versus Swedish-speaking Finns) with varying numbers of years of formal instruction. It turns out that in this dataset, zero prepositions (as in *Chaplin go ∅ a restaurant*) and zero articles (as in *∅ girl run away but ∅ police[man] take ∅ girl*) make all the difference: Finnish-speaking Finns tend to omit prepositions and (even more) articles that are normally obligatory in the specific syntactic environments, while Swedish-speaking Finns are better at providing prepositions and articles where appropriate. How can we account for this difference? Both Finnish-speaking and Swedish-speaking Finns have native languages whose spatial reference system with overt formal elements aids eventual acquisition of correct English prepositional usage, so both learner groups seem to eventually manage to come to terms with English prepositional usage. By contrast, Finnish offers little help, via positive transfer, with the English article system (note that Finnish does not have real articles), whereas Swedish-speaking Finns have an advantage because Swedish does have an article system that is broadly similar to the English system. Odlin observes that unlike English preposition usage English article usage is descriptively fairly complex: while the article system does not have many formal elements, the determinants of the choice between zero and explicit articles are descriptively complex and thus present a tall order to learners whose first language does not have a comparable article system. In any case, the combination of the absence of positive transfer effects and lots of complexity in the target language go a long way towards explaining why Finnish-speaking learners of English have particular difficulties with the English article system.

kort_001.pod 22
07-01-05 16:25:03 -mlb- mlb

Zero and non-zero take center stage in Odlin's contribution, and so they do in »Deletions, antideletions and complexity theory, with special reference to Black South African and Singaporean Englishes« by **Rajend Mesthrie**. Focusing on the strikingly different morphosyntactic behavior – economy versus explicitness – of two indigenized L2 varieties of English, Mesthrie draws on previous work (Mesthrie 2006) to define a ›deletion‹ parameter that yields a fairly reliable typology of indigenized L2 varieties of English. In the explicitness camp, we find varieties, such as Black South African English, that are ›anti-deleters‹. These obey the following principles:

1. If a grammatical feature can be deleted in Standard English, it can be undeleted in the anti-deleting variety.
2. If a grammatical feature can't be deleted in Standard English, it almost always can't be deleted in the anti-deleting variety.
3. If *X* is a grammatical feature of the anti-deleting variety that is not covered by Principles 1 and 2, then *X* almost always involves the presence of a morpheme not found in Standard English.

Across the aisle in the economy cabin, we are dealing with indigenized L2 varieties of English (for example, Singaporean English) which are ›deleters‹. Their morphosyntax profile is governed by the following principles:

4. If a grammatical feature can be deleted in Standard English, it can be deleted in the deleting variety.
5. If a grammatical feature can't be deleted in Standard English, it almost always can be (variably) deleted in the deleting variety.

Mesthrie points out that the deleting/anti-deleting polarity between varieties such as Black South African English and Singaporean English can be interpreted in terms of complexity variance. Hence, the most straightforward interpretation of the behavior of Black South African English is one along the lines of an inverted absolute-quantitative complexity notion: due to its supposedly user-friendly explicitness, Mesthrie considers Black South African English comparatively simple (motto: ›more is [cognitively] less complex‹). However, by virtue of logic this would mean that a deleting variety such as Singaporean English should count as complexifying, a conclusion that is hardly palatable to many in the World Englishes community (which is beholden to the view that deletion must be simplifying). In an attempt to resolve this issue, Mesthrie emphasizes that substrate languages (for example, isolating languages in the case of Singaporean English) need to be taken into account, as these can explain away a good deal of what might look at first blush as complexity variance (we note that Huber's contribution is more critical about this line of explanation). Mesthrie's study thus highlights the fact that comparing complexity levels across contact lan-

kort_001.pod 23
07-01-05 16:25:03 -mlb- mlb

guages with different substrate profiles, as is customary in the literature, can be problematic.

In **Peter Mühlhäusler**'s »The complexity of the personal and possessive pronoun system of Norf'k«, we are likewise reminded of a number of shortcomings in conventional thinking about linguistic simplicity and complexity. Mühlhäusler illustrates these drawing on Pitkern Norf'k as a case study. Pitkern Norf'k is a creole language spoken on Pitcairn and Norfolk Island in the Pacific Ocean by the descendants of the Bounty mutineers, and the language operates in a very specific social ecology. For example, from the beginning there has been a pressing need to distinguish between insiders and outsiders, and this is expressed not only lexically but the distinction is also grammaticalized. Mühlhäusler's overall point – besides drawing attention to the deplorably sketchy nature of published accounts on Pitkern Norf'k – is that this setting must be taken into account when evaluating Pitkern Norf'k's linguistic simplicity (or, rather, complexity). Mühlhäusler focuses on local complexity in the pronoun system and uses, or rather critiques, an absolute-quantitative complexity measure according to which more contrasts and distinctions yield more »descriptive« complexity. It turns out that, overall, the Pitkern Norf'k pronominal system is staggeringly complex according to customary definitions: the inventory is large, and the forms are irregular. For instance, there is plenty of synthesis and fusion, which violates the one-form-one-meaning principle; there are pronouns that must be used when referring to insiders; there is a gender contrast in the third person singular and dual; and so on. All in all, these are complexities that exceed the pronominal complexity of Pitkern Norf'k's source languages and probably all English-based pidgin and creole languages. However, Mühlhäusler highlights the fact that most of Pitkern Norf'k's pronominal complexity resides in the system of deictic reference (whose function it is to carve up »people space«), and not in the system of anaphoric reference. Why is this? Mühlhäusler presents an extended argument that Pitkern Norf'k's complex social setting has promoted a complex system of deictic reference. Thus, what looks like linguistic complexity at first glance is actually just about adequate given the communicative and social-indexical job Pitkern Norf'k has to do: to carve up people space given the islanders' needs. Mühlhäusler thus demonstrates that failing to take into account communicative function results in linguistic complexity ratings that are maybe descriptively accurate, but that do not tell the whole story – akin to asking ›What is the simplest tool‹ rather than asking the better question ›What is the simplest tool *for a certain job*?‹.

Mühlhäusler's idea of toolbox complexity is echoed, to some extent, by **Lourdes Ortega**'s plea for complexity metrics in SLA research that take

kort_001.pod 24
07-01-05 16:25:03 -mlb- mlb

language function seriously. Thus, »Interlanguage complexity: A construct in search of theoretical renewal« reviews previous SLA approaches to linguistic (specifically, interlanguage) complexity, and sketches a research agenda to guide future SLA research in this spirit. Ortega's point of departure is that while extant SLA research has sought to measure interlanguage complexity for a variety of purposes, the literature is dominated by three notions of linguistic complexity, all of which come into the remit of ›more is more complex‹ absolute-local complexity: (1) length of selected linguistic units (e.g. increased mean T-unit length is supposed to index increased complexity), (2) density of subordination (e.g. more subordination is more complex), and (3) frequency of complex forms (higher frequencies of complex forms equate with increased complexity). Ortega argues that all of these notions are somewhat simplistic, selective, and do not consistently do justice to the intricacies of interlanguage complexity, which is why they suffer from what Ortega calls »construct reductionism«. To address these shortcomings, Ortega advocates drawing on ideas and concepts originally developed in *Systemic Functional Linguistics* (SFL) (e.g. Halliday and Matthiessen 1999), a framework that usefully distinguishes between *dynamic styles* (low formality, typically oral) and *synoptic styles* (high formality, typically written). It turns out that while subordination – which, to reiterate, is one of the workhorse complexity diagnostics in extant SLA research – is crucial for the development of dynamic styles, it is less relevant to the development of synoptic styles, where processes such as nominalization and grammatical metaphor (e.g. *to feel* [verb] *feeling* [noun]) take center stage. Ortega cites preliminary evidence that measures of subordination may not be appropriate to measure interlanguage complexity under all circumstances, and that they may be actually inappropriate when dealing with advanced levels of proficiency. This is another way of saying that what is needed is a set of (at least) two complexity metrics: one that measures complexities in dynamic styles, typically at early levels of proficiency (where semantic content is mapped onto grammatical categories via, e.g., subordination), and one that gauges complexity in synoptic styles, more often than not at advanced levels of proficiency (where we find less prototypical form-meaning mappings, more nominalizations, and more grammatical metaphors). Ortega's proposal highlights the empirical potential of nominalization and grammatical metaphor as complexity indicators, phenomena that are virtually unexplored in the existing complexity literature. Also, the added bonus of rejecting one-size-fits-all complexity measures is that the analyst is encouraged to take into account register and modality differences. Such an endeavor dovetails nicely with recent usage-based approaches to language (see,

kort_001.pod 25
07-01-05 16:25:03 -mlb- mlb

for example, Ellis and Larsen-Freeman 2006) and in so doing opens up research interfaces with corpus-linguistic research programs concerned with the dynamics of contact languages, such as indigenized L2 languages, at different stages of development.

A corpus-based perspective on indigenized L2 varieties, then, is precisely what underpins **Maria Steger and Edgar W. Schneider**'s approach in »Complexity as a function of iconicity: The case of complement clause constructions in New Englishes«. Drawing on a bedrock principle in the functionalist literature – *iconicity* – the contribution probes complexity variance in a number of indigenized L2 varieties of English (specifically: Singapore English, Indian English, Hong Kong English, and East African English) vis-à-vis Standard British English for benchmarking purposes (note that the study is entirely based on the publicly available *International Corpus of English*). Steger and Schneider utilize a complexity notion that has hitherto received less attention than it deserves: they define linguistic complexity as being inversely proportional to linguistic *iconicity* of various types, which all boil down to the degree to which form-meaning pairings are motivated and transparent (in this sense, the contribution ties in with Mesthrie's paper, which considers anti-deletion simple because explicitness is maintained). The view that iconicity may translate into linguistic simplicity, and un-iconicity into complexity, is essentially a hybrid absolute-relative notion: cognitive-functionalist theory dictates which structures are considered iconic and which are not, but the key assumption is relative in that iconic constructions are posited to be prioritized in settings of multilingualism/multidialectism and concomitant adult SLA (thanks to speaker-oriented processing optimization). It is precisely because of this that Steger and Schneider expect that iconicity should be more pervasive in New Englishes than in, e.g., Standard British English. The study explores local complexity in the verbal complexity domain, operationalizing iconicity-based complexity such that, for example, non-finite complementation patterns (as in *John expects Mary to come*) count as less iconic and hence more complex than finite complementation patterns (for example, *John expects that Mary comes*). In exactly this spirit, Steger and Schneider spell out a number of structural hypotheses that should be borne out in the data if iconicity were indeed prioritized in the genesis of indigenized L2 varieties of English, as Steger and Schneider conjecture. Thus, indigenized L2 varieties of English should exhibit (1) a preference for finite (rather than non-finite) complementation patterns; (2) a preference for patterns with an overt (as opposed to zero) complementizer; (3) a preference for explicit expressions of modality; (4) a preference for explicit structural marking of double conceptual functions; and (5) a dispreference for raising

kort_001.pod 26
07-01-05 16:25:03 -mlb- mlb

constructions. Steger and Schneider subsequently present a meticulous qualitative and quantitative analysis to put these hypotheses to the empirical test, and they demonstrate that, indeed, New Englishes – and especially so Hong Kong English and East African English – tend towards the iconic pole compared to Standard British English. In the context of the present volume, the contribution highlights the fact that the corpus-based complexity literature would profit from more studies along the lines of Steger and Schneider's that go beyond, e.g., mere counting of elements or some such and that instead rely on key notions in functionalist theory to conceptualize the notion of linguistic complexity.

In »Acquisitional complexity: What defies complete acquisition in Second Language Acquisition«, **ZhaoHong Han and Wai Man Lew** center – much as Steger and Schneider as well as, to some extent, Ortega do – on the interplay between form and function. The paper specifically discusses a novel take on conceptualizing complexity in SLA research and beyond. In most general terms, Han and Lew view complexity as a linguistic as well as psycholinguistic notion – more precisely, as an »intricate relationship, relative rather than absolute, between linguistic elements, with both a clinal and a temporal dimension«. Han and Lew differentiate between two pertinent types of complexity, *developmental* and *acquisitional*; the focus of the contribution is on the latter type. Traditional complexity metrics in the SLA literature, such as, e.g., number of T-units per sentence, come within the remit of developmental complexity, which is all about what is acquirable at a given time. Against this backdrop, the authors define acquisitional complexity as what is ultimately non-acquirable by learners. Acquisitional complexity is mostly idiosyncratic and static; it focuses on form-meaning pairings; it is a function of stabilized interactions between exogenous and endogenous contingencies; and, despite its static nature, empirically it can be captured on the basis of longitudinal data only. We note, along these lines, that acquisitional complexity in the spirit of Han and Lew is one of the most radically relative complexity notions discussed in the present volume. How exactly do we find out about acquisitional complexity, then? Han and Lew offer that the notion of *fossilization* (i.e. cessation of learning) and, specifically, Han's *Selective Fossilization Hypothesis* (Han 2009) are à propos. The Selective Fossilization Hypothesis identifies L1 markedness and L2 input robustness as crucial determinants of likely fossilization candidates and, hence, of acquisitional complexity; L1 markedness and L2 input robustness, in turn, are a function of formal frequency and formal-functional variability. Han and Lew's argument, which they illustrate on the basis of two case studies, boils down to an argument that linguistic features that are marked (i.e. infrequent and/or vari-

kort_001.pod 27
07-01-05 16:25:03 -mlb- mlb

able) in a learner's L1 and robust (i.e. frequent and invariable) in L2 input are unlikely to fossilize and thus should count as acquisitionally simple. By contrast, features that are unmarked in the L1 and non-robust in L2 input are acquisitionally complex. A crucial element in this nexus is Slobin's (1987) *Thinking for Speaking Hypothesis*. In all, Han and Lew's contribution is an inspiring piece that combines ideas about first language influence, fossilization, markedness, formal frequency, and formal-functional (in)variability with cross-disciplinarily relevant reasoning about linguistic complexity. The proposal's focus on L2-endstate grammars in particular should provide inspiration to scholars interested in complexity variance among indigenized L2 varieties.

By comparison to Han and Lew (and also Odlin), **Magnus Huber**, in »Syntactic and variational complexity in British and Ghanaian English: Relative clause formation in the written parts of the International Corpus of English«, is relatively skeptical about the explanatory potency of transfer effects when it comes to explaining the grammatical blueprint of indigenized L2 varieties. Huber's contribution is a careful variationist investigation, based on the *International Corpus of English*, into what may happen to a particular grammatical domain (such as restrictive relative clause formation) when users of an indigenized L2 variety (such as Ghanaian English) at a particular evolutionary stage (in the case of Ghanaian English, nativization and endonormative stabilization in the parlance of Schneider 2007) are faced with certain complexities in the norm-providing input variety (in the Ghanaian scenario, Standard British English). Interpretationally, Huber draws on a hybrid complexity notion that is partly absolute and partly relative. The absolute component consists of defining complexity in some linguistic domain as being proportional to the number of elements and, particularly, communicative redundancies in that domain. The relative component highlights those communicative redundancies that are uncommon either cross-linguistically and/or in the relative adstrate languages, for these are supposedly hard to acquire for adult language users. Consequently, ›uncommon‹ redundancies should be disadvantaged in the genesis of indigenized L2 varieties. In addition, Huber considers inherent variability as complexifying (as ›variationally‹ complex, that is). Against the backdrop of this particular complexity definition, the (Standard) British English relativization domain seems fairly complex, since (i) it exhibits several relativization strategies (pronouns, particles, and zero) with functional overlap; (ii) there is inherent variability between these relativization strategies, which is governed by somewhat exotic factors such as antecedent animacy; and (iii) relative pronouns in particular are crosslinguistically relatively uncommon. On the empirical plane, Huber ana-

kort_001.pod 28
07-01-05 16:25:03 -mlb- mlb

lyzes a dataset in which relative clause occurrences in the corpus database, which covers Ghanaian English as well as British English, were annotated for several contextual variables (syntactic function of the relativizer, voice of the relative clause, animacy of the antecedent). Huber goes on to present an extended empirical argument that relative clause formation in Ghanaian English can be seen as being less complex than in (Standard) British English: for example, Ghanaian English has a stronger preference – one that is in line with the cross-linguistic picture – for the invariant relative particle *that*. Crucially, Huber shows the Ghanaian »reinterpretation« of the (Standard) British English relativization system cannot be entirely explained away by properties of Ghanaian adstrate languages. In sum, the study merits particular attention thanks to its focus on gradience and variation – as Huber points out, Ghanaian relative clauses are typically not ungrammatical in (Standard) British English, but in the big picture the underlying variationist constraints appear to be subtly different.

John McWhorter's »Complexity hotspot: The copula in Saramaccan and its implications« also takes a critical stance toward the explanatory power of adstrate or substrate effects. The study sets its sights on local grammatical complexities in Saramaccan creole, an English-lexified contact language spoken in Suriname that is related to Sranan creole but also deeply influenced by Fongbe, an African language spoken by its creators. McWhorter marshals a complexity metric that, in the tradition of McWhorter (2001), is absolute-quantitative in nature and consists of the following three components:

– *overspecification*: complexity is a function of the number of overt and obligatory marked distinctions
– *structural elaboration*: complexity is a function of the number of rules required to generate well-formed grammatical output
– *irregularity*: complexity is a function of the amount of irregularity and suppletion

The grammatical domain in Saramaccan that McWhorter inspects with regard to the above metric is that of copula constructions, which turn out to be unusually complex – against the backdrop of overall grammatical complexity in Saramaccan, but also in the wider context of creole languages in general. The Saramaccan copula exhibits overspecification in that constructions vary as a function of the semantics of the predicate. We also find plenty of structural elaboration in the form of, e.g., allomorphies, and the Saramaccan copula features a substantial amount of irregularity and suppletion. McWhorter discusses diachronic and synchronic evidence suggesting that this sort of local complexity must have developed gradually over time in Saramaccan, and that there is no good reason to believe that these complexities

kort_001.pod 29
07-01-05 16:25:03 -mlb- mlb

were transferred from one of Saramaccan's substrate languages (especially since, as McWhorter points out, copula constructions are typically omitted, rather than preserved, in creole genesis and SLA). In short, the complexity of the Saramaccan copula constructions is somewhat puzzling. Where does it come from? McWhorter offers an essentially usage-based explanation. What probably happened in Saramaccan was that initially, a deictic element was subject to reanalysis as a copula construction, as is customary cross-linguistically. Next, phonetic erosion created allomorphies which – thanks to high discourse frequency – were fairly resistant to simplification. In addition, we find in Saramaccan recurrent reanalysis of topic-comment constructions into subject-predicate constructions, a process that additionally creates complexities. According to McWhorter, then, the complexity of the Saramaccan copula, which comes into the remit of a typologically analytic domain, highlights the fact that complexity need not necessarily reside in synthetic grammatical marking or clausal embedding, which are the usual suspects in the complexity literature. What is more, the Saramaccan copula scenario demonstrates that linguistic complexity does not necessarily result from language-external forces (such as language contact), for in the Saramaccan case we find that complex copulas are the result of language-internal processes along the lines of, e.g. grammaticalization theory. The argument presented in »Complexity hotspot« thus overlaps substantially with Edward Sapir's notion of ›drift‹, and in so doing comes back full circle to structuralist lines of analysis that, ironically, triggered the emergence of the equi-complexity axiom in the early twentieth century.

References

Arends, Jacques 2001 Simple grammars, complex languages. *Linguistic Typology* 5 (2/3): 180–182.

Bisang, Walter 2009 On the evolution of complexity: sometimes less is more in East and mainland Southeast Asia. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 34–49. Oxford: Oxford University Press.

Braunmüller, Kurt 1990 Komplexe Flexionssysteme – (k)ein Problem für die Natürlichkeitstheorie? *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 43: 625–635.

Chipere, Ngoni 2009 Individual differences in processing complex grammatical structures. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 178–191. Oxford: Oxford University Press.

Dahl, Östen 2004 *The growth and maintenance of linguistic complexity.* Amsterdam, Philadelphia: Benjamins.

Dahl, Östen 2008 Grammatical resources and linguistic complexity: Sirionó as a

kort_001.pod 30
07-01-05 16:25:03 -mlb- mlb

language without NP coordination. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 153–164. Amsterdam, Philadelphia: Benjamins.

Dahl, Östen 2009 Testing the assumption of complexity invariance: the case of Elfdalian and Swedish. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 50–63. Oxford: Oxford University Press.

Dammel, Antje, and Sebastian Kürschner 2008 Complexity in nominal plural allomorphy. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 243–262. Amsterdam, Philadelphia: Benjamins.

Davydova, Julia to appear a English as a second language vs. English as a foreign language: Bridging the gap. *World Englishes.*

Davydova, Julia to appear b *The Present Perfect in non-native Englishes.* Berlin, New York: de Gruyter Mouton.

Deutscher, Guy 2009 »Overall complexity«: a wild goose chase? In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 243–251. Oxford: Oxford University Press.

Deutscher, Guy 2010 *Through the language glass: why the world looks different in other languages.* New York: Metropolitan Books.

Ehret, Katharina Luisa 2008 *Analyticity and syntheticity in East African English and British English: a register comparison*. URN: urn:nbn:de:bsz:25-opus-58047, URL: http://www.freidok.uni-freiburg.de/volltexte/5804/. Freiburg: University of Freiburg.

Ellis, Nick C., and Diane Larsen-Freeman 2006 Language emergence: Implications for applied linguistics – Introduction to the special issue. *Applied Linguistics* 27 (4): 558–589.

Fenk-Oczlon, Gertraud, and August Fenk 2008 Complexity trade-offs between the subsystems of language. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 43–65. Amsterdam, Philadelphia: Benjamins.

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

Givón, T. 2009 *The genesis of syntactic complexity: diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam, Philadelphia: Benjamins.

Granger, Sylviane 2003 The International Corpus of Learner English: A New Resource for Foreign Language Learning and Teaching and Second Language Acquisition Research. *TESOL Quarterly* 37 (3): 538–546.

Greenbaum, Sidney 1996 *Comparing English worldwide: the International Corpus of English*. Oxford, New York: Clarendon Press/Oxford University Press.

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

Halliday, M. A. K., and Christian M. I. M. Matthiessen 1999 *Construing experience through meaning: A language-based approach to cognition*. London: Cassell.

Han, Zhaohong 2009 Interlanguage and fossilization: towards an analytic model. In: Vivian Cook and Li Wei (eds.), *Contemporary Applied Linguistics*, 137–162. London: Continuum.

Haspelmath, Martin, Matthew S Dryer, David Gil, and Bernard Comrie (eds.) 2005 *The World Atlas of Language Structures.* Oxford: Oxford University Press.

kort_001.pod 31
07-01-05 16:25:03 -mlb- mlb

Hawkins, John A. 2009 An efficiency theory of complexity and related phenomena. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 252–268. Oxford: Oxford University Press.

Hockett, Charles F 1958 *A Course in Modern Linguistics*. New York: Macmillan.

Humboldt, Wilhelm von 1836 *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Dümmler.

Juola, Patrick 2008 Assessing linguistic complexity. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 89–108. Amsterdam, Philadelphia: Benjamins.

Juvonen, Päivi 2008 Complexity and simplicity in minimal lexica: the lexicon of Chinook Jargon. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 321–340. Amsterdam, Philadelphia: Benjamins.

Karlsson, Fred 2009 Origin and maintenance of clausal embedding complexity. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 192–202. Oxford: Oxford University Press.

Karlsson, Fred, Matti Miestamo, and Kaius Sinnemäki 2008 Introduction: The problem of language complexity. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, vii-xiv. Amsterdam, Philadelphia: Benjamins.

Klein, Wolfgang, and Clive Perdue 1997 The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–347.

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

Kortmann, Bernd, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie, and Clive Upton (eds.) 2004 *A Handbook of Varieties of English*. Berlin, New York: Mouton de Gruyter.

Kortmann, Bernd, and Benedikt Szmrecsanyi 2004 Global synopsis: morphological and syntactic variation in English. In: Bernd Kortmann, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie, and Clive Upton (eds.), *A Handbook of Varieties of English*, Vol. 2, 1142–1202. Berlin, New York: Mouton de Gruyter.

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*, 265–285. Amsterdam, Philadelphia: Benjamins.

Kortmann, Bernd, and Benedikt Szmrecsanyi 2011 Parameters of morphosyntactic variation in World Englishes: prospects and limitations of searching for universals. In: Peter Siemund (ed.), *Linguistic Universals and Language Variation*, 264–290. Berlin, Boston: De Gruyter Mouton.

Kuiken, Folkert, and Ineke Vedder 2008 Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing* 17 (1): 48–60.

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

Kusters, Wouter 2008 Complexity in linguistic theory, language learning and language change. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 3–22. Amsterdam, Philadelphia: Benjamins.

kort_001.pod 32
07-01-05 16:25:03 -mlb- mlb

Larsen-Freeman, Diane 1978 An ESL Index of Development. *TESOL Quarterly* 12 (4): 439–448.

Lindström, Eva 2008 Language complexity and interlinguistic difficulty. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 217–242. Amsterdam, Philadelphia: Benjamins.

Mayerthaler, Willi 1981 *Morphologische Natürlichkeit*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.

McWhorter, John 2001 The world's simplest grammars are creole grammars. *Linguistic Typology* 6: 125–166.

McWhorter, John 2008 Why does a language undress? Strange cases in Indonesia. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 167–190. Amsterdam, Philadelphia: Benjamins.

Mesthrie, Rajend 2006 World Englishes and the multilingual history of English. *World Englishes* 25 (3/4): 381–390.

Miestamo, Matti 2008 Grammatical complexity in a cross-linguistic perspective. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 23–42. Amsterdam, Philadelphia: Benjamins.

Miestamo, Matti 2009 Implicational hierarchies and grammatical complexity. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 80–97. Oxford: Oxford University Press.

Miestamo, Matti, Kaius Sinnemäki, and Fred Karlsson 2008 *Language complexity: typology, contact, change*. Amsterdam, Philadelphia: Benjamins.

Mühlhäusler, Beverly S., and Peter Mühlhäusler 2005 Simple English in the South Seas Evangelical Mission: Social context and linguistic attributes. *Language Problems & Language Planning* 29 (1): 1–30.

Mühlhäusler, Peter 1974 *Pidginization and simplification of language*. Pacific linguistics. Series B. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Mühlhäusler, Peter 1992 Twenty years after: a review of Peter Mühlhäusler's pidginization and simplification of language. In: Martin Pütz (ed.), *Thirty Years of Linguistic Evolution*, 109–117. Amsterdam, Philadelphia: Benjamins.

Muysken, Pieter, and Norval Smith 1995 The study of pidgin and creole languages. In: Jacques Arends, Pieter Muysken, and Norval Smith (eds.), *Pidgins and Creoles: An Introduction*, 1–14. Amsterdam, Philadelphia: Benjamins.

Nichols, Johanna forthcoming The vertical archipelago: Adding the third dimension to linguistic geography. In: Peter Auer, Martin Hilpert, Anja Stukenbrock, and Benedikt Szmrecsanyi (eds.), *Space in language and linguistics: geographical, interactional, and cognitive perspectives*. Berlin, New York: Walter de Gruyter.

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

Ogden, C. K 1934 *The system of Basic English*. New York: Harcourt.

Ortega, Lourdes 2003 Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24: 492–518.

Parkvall, Mikael 2008 The simplicity of creoles in a cross-linguistic perspective. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 265–285. Amsterdam, Philadelphia: Benjamins.

kort_001.pod 33
07-01-05 16:25:03 -mlb- mlb

Rescher, Nicholas 1998 *Complexity: a philosophical overview*. New Brunswick: Transaction Publishers.

Sadeniemi, Markus, Kimmo Kettunen, Tiina Lindh-Knuutila, and Timo Honkela 2008 Complexity of European Union Languages: A Comparative Approach. *Journal of Quantitative Linguistics* 15 (2): 185–211.

Sampson, Geoffrey 2009 A linguistic axiom challenged. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language complexity as an evolving variable*, 1–18. Oxford: Oxford University Press.

Sampson, Geoffrey, David Gil, and Peter Trudgill 2009 *Language complexity as an evolving variable*. Oxford linguistics. Oxford, New York: Oxford University Press.

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

Selinker, Larry 1972 Interlanguage. *International Review of Applied Linguistic* 10: 209–230.

Seuren, Pieter, and Herman Wekker 1986 Semantic transparency as a factor in creole genesis. In: Pieter Muysken and Norval Smith (eds.), *Substrata versus Universals in Creole Genesis*, 57–70. Amsterdam, Philadelphia: Benjamins.

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

Sinnemäki, Kaius 2008 Complexity trade-offs in core argument marking. In: Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 68–88. Amsterdam, Philadelphia: Benjamins.

Sinnemäki, Kaius 2009 Complexity in core argument marking and population size. In: Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*, 126–140. Oxford: Oxford University Press.

Slobin, Dan I. 1987 Thinking for speaking. *Proceedings of the Berkeley Linguistics Society* 13: 435–444.

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

Szmrecsanyi, Benedikt 2012 Analyticity and syntheticity in the history of English. In: Terttu Nevalainen and Elisabeth Closs Traugott (eds.), *Rethinking the History of English*. Oxford: Oxford University Press.

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

Szmrecsanyi, Benedikt, and Bernd Kortmann 2009b 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*, 33–53. London, New York: Routledge.

Szmrecsanyi, Benedikt, and Bernd Kortmann 2009c 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*, 64–79. Oxford: Oxford University Press.

Szmrecsanyi, Benedikt, and Bernd Kortmann 2011 Typological profiling: learner Englishes versus indigenized L2 varieties of English. In: Joybrato Mukherjee and Marianne Hundt (eds.), *Exploring Second-Language Varieties of English and Learner Englishes: Bridging a Paradigm Gap*, 167–187. Amsterdam, Philadelphia: Benjamins.

kort_001.pod 34
07-01-05 16:25:03 -mlb- mlb

Thomason, Sarah Grey 2001 *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah Grey, and Terrence Kaufman 1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Towell, Richard, and Roger Hawkins 1994 *Approaches to second language acquisition*. Clevedon, Philadelphia: Multilingual Matters.

Trudgill, Peter 1999 Language contact and the function of linguistic gender. *Poznan Studies in Contemporary Linguistics* 35: 133–152.

Trudgill, Peter 2001 Contact and simplification: historical baggage and directionality in linguistic change. *Linguistic Typology* 5 (2/3): 371–374.

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

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