8

How British is Gibraltar English?[1]

Manfred Krug, Ole Schützler and Valentin Werner
University of Bamberg

1. Introduction

Gibraltar is a British Overseas Territory at the southern tip of Spain which was ceded to Great Britain by the Treaty of Utrecht in 1713. Today, Gibraltar has a population of approximately 30,000 and the only official language is English (Levey 2015: 66; Central Intelligence Agency 2016). Also widely spoken are Spanish and the local Andalusian-based Spanish variety called Llanito (or Yanito), characterized by extensive code-switching and lexical borrowing from English (as well as from other languages and dialects; see Kellermann 2001; Levey 2008). In this linguistic context, the present study focuses on concepts for which a (traditionally more) British variant and a (traditionally more) American variant coexist. More precisely, we investigate two referentially synonymous expressions that are known, or consistently reported in standard reference works and textbooks, to have differed in usage between British English (BrE) and American English (AmE) in the late twentieth century (cf., e.g., Algeo 2006; Krug and Sönning 2017 for web data). On the basis of $n = 312$ questionnaires (i.e. a sample of roughly one per cent of the population), we investigate what linguistic choices Gibraltarians make in cases such as *lorry* vs. *truck* or *parcel* vs. *package.* In order to place Gibraltar with regard to the two major norm-providing varieties of English and in order to identify patterns of variation and change, we compare questionnaire data from Great Britain and the US.

The status of Gibraltar English (GibE) is special in a number of ways. Gibraltar has been labelled an 'unusual British colony' (Weston 2015: 647), and attempts at linguistic description have been characterized as suffering from a 'mismatch between its colonial status and sociolinguistic theory' (Weston 2011: 339), for example in terms of categorizations of GibE within models of World Englishes (such as Kachru's Concentric Circles Model or Schneider's Dynamic Model of Postcolonial Englishes). While Spanish-English language contact is omnipresent (Kramer 1998; Suárez-Gómez 2012: 1746–8; see also Table 8.2 below), it is clear that GibE does not epitomize what can be considered a typical L2 variety for a number of reasons. First, the vast majority of the current population in Gibraltar speaks English as a first language or can at least be considered highly fluent (see www.ethnologue.com/language/eng),[2] and our own data

– with a preponderance of younger speakers – seem to reflect the age-grading identified in earlier studies (e.g., Levey 2008: 95, 2015: 53; Weston 2012: 18) in terms of an increasing bias towards English as an L1. Second, *prima facie*, it appears that the political status as British Overseas Territory should lead (as in earlier stages of other colonial contexts) to Gibraltarians perceiving 'themselves as outposts of Britain, deriving their social identity primarily from their common territory of origin and a feeling of culturally belonging there' (Schneider 2007: 37). Interestingly, however, in the sociolinguistic section of our questionnaire, 55 per cent (172/312) of the raters submitted Gibraltar as the country or region with which they identify most, while a mere 12 per cent provided either UK, Britain or England as an answer. For the majority of our raters, therefore, a strong sense of a separate local identity seems to have emerged (*pace* Kellermann 2001: 111; cf., Levey 2015: 51; Weston 2015: 676; Loureiro-Porto and Suárez-Gómez 2017: 96).

Further usage- and domain-related aspects offer important insights. As in all British Overseas Territories, English is the sole official language (and thus also used as medium of instruction; Weston 2011: 359; Levey 2015: 66). English in Gibraltar, then, is not a co-official language, as in other L2 contexts like India or Jamaica. Also, the socio-linguistic information gleaned from our questionnaire supports the view that English is not restricted to official and formal contexts: 56 per cent (175/312) of the raters belong to the category defined as 'English-dominant' based on language use at home (cf. Table 8.2); 71 per cent report to use English at home at least to the same extent as Spanish and a mere 5 per cent report to use no English at all at home (cf. also Loureiro-Porto and Suárez-Gómez 2017: 96–7).[3]

The remainder of this chapter is organized as follows: Section 2 discusses the data and questionnaire structure. Section 3 offers exploratory analyses in the form of distance matrices and identifies broad regional patterns and individual outlier informants, while Section 4 focuses on individual items and rater means. Socio-linguistic factors will be investigated in detail, and apparent-time studies reveal trends of change for individual items as well as recurrent patterns of change in progress in Gibraltar English. Our data suggest that, while Gibraltarians broadly conform in their usage to speakers of the main contact variety, BrE, younger speakers of English in Gibraltar display tendencies of Americanization and globalization in opting less consistently for (conservative) British forms than older Gibraltarians. And while older speakers of both sexes behave conservatively British to the same extent, a gender gap emerges for middle-aged and younger raters: in our Gibraltar data, it is the men who are leading in this trend towards a less rigidly British-oriented usage. Interestingly, however, we also find occasional instances of increasing Britishization for individual items, that is, more consistent usage of traditional British variants among the younger cohorts.

2. Data and questionnaire structure

The questionnaire developed at the University of Bamberg investigates morphosyntactic and lexical variation in varieties of English world-wide (see Krug and Sell 2013 for

more detail). The full version takes about sixty minutes to complete and requires high levels of concentration as well as quiet environments, because 138 audio files are played (twice) and rated. For these reasons, it is conducted in educational institutions and larger groups only. The complete questionnaire elicits, *inter alia*, differences between one both medially and conceptually spoken register (a casual conversation among friends) and one semi-formal written register (an email to a former school teacher).

In this contribution, we concentrate on lexical items. In addition to material collected in larger groups and educational contexts, we use material elicited from individuals, mostly in public places (e.g., streets and parks), but also cafés, supermarkets, shops, etc. For this, the lexical part of the Bamberg Questionnaire was used (see Appendix A), which elicits the informants' personal preference in cases where there is a choice between (at least) two (near-)synonyms that are known to differ in British and American usage. For ambiguous terms, the semantic domains are specified. In the pair *lorry vs. truck*, for instance, reference to 'large motor vehicles for carrying goods by road' is made, while for the pair *parcel* vs. *package* a paraphrase 'something you send by mail' is given in the questionnaires.[4]

For each of the 68 lexical binaries[5] (see Appendix A), informants can choose one of five options, or not choose any, that is, opt out. If they choose 'I only use this expression', this translates into a score of +2 if the (more) British variant is chosen; −2 if the (more) American variant is chosen; 'I prefer this expression' translates into +1 if the (more) British variant is chosen; −1 if the (more) American variant is chosen; 'I have no preference' translates into a score of zero. If the informant ticks that he or she uses neither of the two lexical expressions, this item is left out of the calculation. In addition, informants can add brief comments or mention what they would use instead of the two options given.

Informant metadata for this questionnaire study (see Appendix A for detail) include age, gender, nationality, ethnic self-identification, parents' native languages, languages used at home while growing up as well as the informants' educational profile, their current occupation and their parents' highest qualification and (potentially last) occupation. In addition, we ask informants to provide the places of residence for their entire lifespan. In order to test what impact attitudes and identity considerations have on linguistic choices, we also ask informants to provide the region or country they identify with most.

The present study uses questionnaire data from three varieties of English: BrE, AmE and GibE. The vast majority of the data are from Gibraltar (n = 312), while BrE and AmE are represented by much smaller samples (n_{BrE} = 14; n_{AmE} = 25). The main focus of the analysis is therefore on English in Gibraltar, while the latter two varieties are treated as reference points. Data from Gibraltar were collected in 2010; the AmE and BrE data were collected in 2008 and 2010, respectively.

To control for a possible gender bias, the sample was split into male and female raters for the global apparent-time analyses. The GibE data are divided into seven age groups as shown in Table 8.1. Apart from a slight overrepresentation of male raters aged 30–39 (who account for 59 per cent of their cohort), age groups are fairly balanced by gender for raters who are twenty years old and older. The youngest age group (< 20), however, is dominated by female raters, who constitute about two thirds of their cohort,

Table 8.1 Numbers of raters by age group and gender (n_{total} = 312)

	<20	20–29	30–39	40–49	50–59	60–69	>69
			Age group				
Male	51	15	20	13	14	10	11
Female	103	15	14	16	12	8	9
NA	1						
Σ	155	30	34	29	26	18	20

which in turn is due to the fact that large numbers of questionnaires were collected in a girls' school.

For some analyses, we subdivided the GibE sample by raters' language background. The criterion applied was the raters' response to the questionnaire item 'language(s) used at home when growing up'. While alternative language-related parameters could be used (e.g., the native language of raters' mothers or fathers), we consider actual language use during the formative years of childhood and youth to be the best indicator of linguistic background. Responses to this questionnaire item can be placed on the scale shown in the left-hand column of Table 8.2. The extreme poles of this scale are exclusive use of either English (top row) or Spanish (bottom row); in addition, there are three intermediate values.

Two subdivisions can be applied based on this scale, as shown in columns two and three of Table 8.2. One method (Subdivision A) is to decide which of the two languages – English or Spanish – is the dominant one, collapsing 'English (only)' and 'mostly English, some Spanish' into a single category, for example; in this case, only the use of English and Spanish to roughly similar extents remains as an intermediate category. The alternative approach (Subdivision B) is to collapse raters into a single category if they were exposed to both languages while growing up, irrespective of whether English or Spanish was the dominant one. This results in a very large intermediate category and

Table 8.2 Categorization of raters by language background

'Language(s) used at home while growing up'	Subdivision A	Subdivision B
English (only) n = 65	English-dominated n = 175	English n = 65
mostly English, some Spanish n = 110		
English and Spanish n = 48	English and Spanish n = 48	English and Spanish n = 230
mostly Spanish, some English n = 72	Spanish-dominated n = 89	
Spanish (only) n = 17		Spanish n = 17

smaller groups of raters who were influenced by English or Spanish exclusively. In Section 4 we investigate which of the subdivisions charted in Table 8.2 yield significant results or prompt the most insightful interpretations.

3. Explorative data analysis

As a first step towards describing the structure of large quantitative datasets such as the present one, we rely on explorative aggregative data analysis. This is a hypothesis-generating approach that, in addition to other methods such as multidimensional scaling (MDS; see, e.g., Ruette, Ehret and Szmrecsanyi 2016), has recently proved useful in a number of linguistic studies (see, e.g., Szmrecsanyi and Wolk 2011; McMahon and Maguire 2013; Werner 2014, 2016; Wälchli and Szmrecsanyi 2014; Schützler 2015; Fuchs and Gut 2016). Its aim in the present study is to determine whether data points cluster according to the sociolinguistic variables elicited (such as variety, rater age, gender, language background, etc.), thus both establishing general patterns in the data and highlighting areas that potentially deserve more detailed investigation (e.g., through descriptive and inferential statistics commonly used for hypothesis-testing; see Section 4). At the same time, an analysis of this type may facilitate the establishment of generalized characteristic patterns based on results for individual raters (e.g., association with a reference variety, 'old' vs. 'young' raters, etc.) and may further be used as a tool to identify outliers in the data at a glance.

3.1. Methodological notes

Aggregative analysis has been described as data-driven, that is, it represents a bottom-up approach in not relying on prior assumptions on the structure and distribution of the data. It is aggregative in the sense that it represents a cumulative analysis of multidimensional data (in our case, acceptability ratings of sixty-eight lexical items by different raters from Gibraltar, the UK and the US). Aggregative analysis is a largely visual approach, where graphical representations are eventually used as a means to identify associations and structures in the data (Krug, Schützler and Werner 2016: 37–8). This approach, therefore, exploits that 'humans are very good at seeing things' (Wainer 1992: 15). In other words, the underlying rationale is that visual representations in general (Braithwaite and Goldstone 2013: 1928), and spatial representations (Vessey 1991) like the ones used for the present study in particular, facilitate the comprehension of complex relationships between multiple data points, which would be difficult to process in extended tables, for instance. In concrete terms, in our data we have to compare the values for sixty-eight lexical items across 312 raters of GibE (plus thirty-nine raters from the reference varieties BrE and AmE). It goes without saying that, with such an extended number of dimensions to compare, complexity increases beyond what the human mind can possibly handle.

The type of aggregative analysis we use relies on the calculation of distance matrices[6] based on the ratings for sixty-eight lexical items with non-hierarchical phenograms as visual output. We subsequently describe how to arrive at both the former and the latter

(see further Krug, Schützler and Werner 2016: 38–41 for a detailed description of procedures and arguments for using non-hierarchical graphical representations).

3.2. Calculating a distance matrix

The distance matrix is calculated with *R* (R Core Team 2016) and quantifies the Euclidean distance (see Szmrecsanyi 2011: 54; Krug, Schützler and Werner 2016: 42) between any two of the 312 raters in the dataset. A spreadsheet containing the transformed raw rating values (on which see below) provides the input for the distance matrix. Note that the sixty-eight lexical items are not internally weighted and thus contribute equally to the calculation.

The computation of distance matrices relies on complete datasets, but in questionnaires there is bound to be missing data (in other words, individual raters may submit fewer than sixty-eight ratings; see Section 2). As a compensating strategy for missing values we decided to apply a method that takes into account average values (i) for the variety as a whole, (ii) for the individual item and (iii) for the individual rater. Thus, missing values are calculated as follows for GibE raters:

$$missing\ value = \frac{average\ across\ all\ raters\ (missing\ item) \times overall\ average\ of\ individual\ rater}{overall\ average\ score\ GibE}$$

These calculations rely on a transformed dataset (with the original scale from −2 to +2 changed to a scale ranging from 1 to 5), in order to avoid multiplication with zero, the potential 'neutral' rating option. Calculated values may have to be manually adjusted so that they observe the upper and lower limits of the transformed scale. This is necessary to avoid nonsensical imputed average values like 0.8 or 5.3, for instance, which would be manually adjusted to 1.0 and 5.0, respectively. Once the missing values are inserted into the spreadsheet, we are able to calculate the Spearman rank correlation between all raters in R, and then use the *dist* method (default settings; Euclidean distance) from the *stats* package to arrive at a distance matrix.

3.3. Visual output

With the help of a purpose-built *R* script called *MakeNex*,[7] we then formally adapt the emerging distance matrix (from CSV spreadsheet format) to a format readable by *SplitsTree* (Huson and Bryant 2006; splitstree.org). This program creates the visual output in the form of phenograms, using the NeighborNet algorithm, and offers different formatting and exporting options for the graphs.[8] For our purposes, we use non-hierarchical 'equal angle' representations that allow a fine-grained view of differences and potential associations between raters (see Figure 8.1). At the same time, this type of display reduces multidimensional complexity in the data, that is, the differences between all possible pairings of categories in an *n*-dimensional space, to a two-dimensional display, which is much more accessible to the human mind (see above).

Note the following (intuitive) properties of the phenograms: (i) the distance between two categories ('nodes', i.e., in our case, raters) consistently is the shortest
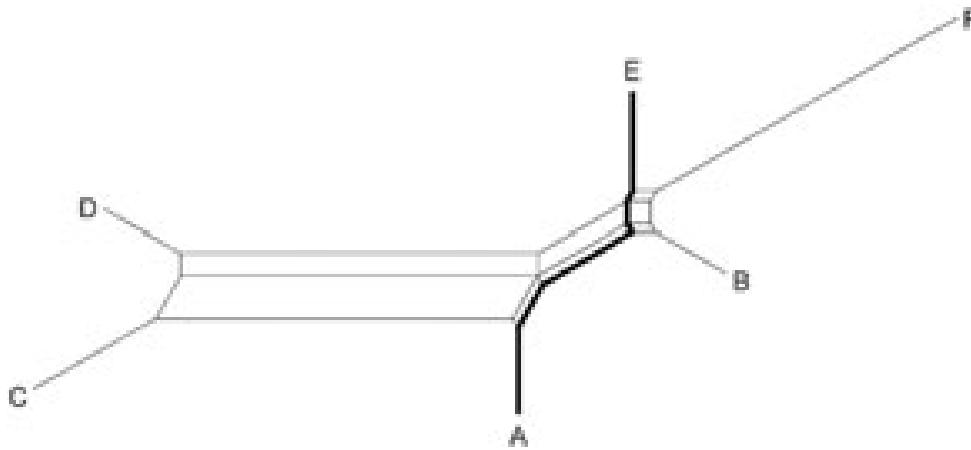
Figure 8.1 Phenogram (NeighborNet) for a simple assumed dataset with six categories (adapted from Krug, Schützler and Werner 2016: 42).

possible way along the edges of the graph and corresponds to the (numerical) Euclidean distance value as represented in the tabular distance matrix (e.g., between categories A and E, highlighted in bold in Figure 8.1); (ii) associations (in our case, in terms of similar rating patterns) exist between neighbouring nodes, for instance between categories C and D in Figure 8.1. Thus, both distances and spatial location play a part in interpreting the graphical output.

3.4. Overall perspective

The overall picture that emerges from a comparison of the ratings of the sixty-eight lexical binaries across 312 raters for GibE, our target variety, is set against the ratings provided by two sets of reference variety raters ($n = 14$ for BrE, $n = 25$ for AmE) in Figure 8.2. The strengths of visual representation are exploited in that nodes for raters receive different shapes (GibE = circles, BrE = grey squares, AmE = black squares). This will facilitate the identification of general structures and potential irregularities in the data.

Two outliers apart (which will be treated shortly), Figure 8.2 provides a clear picture in terms of variety associations. First, it emerges that raters from each variety clearly align with other raters of their variety, which implies that the notion of *variety* is not a mere theoretical sociolinguistic construct, but an empirically valid one when lexical ratings are analysed. More specifically, we find an American cluster towards the left and a British(-influenced) cluster towards the right, with a less densely populated area in the middle, which we interpret as an indication of the relative distance between the two clusters. In other words, Figure 8.2 yields a generalizable clear split (i.e. a large distance) between AmE raters, on the one hand (left), and GibE and BrE raters on the other (right). This finding will receive further support in the other quantitative analyses presented in the remainder of this study.

The robust association between GibE and BrE is plausible from a sociolinguistic and historical point of view, as BrE can be seen as the 'parent' variety of GibE, so that close historical and cultural ties are likely to be reflected in language use.[9] The graphical
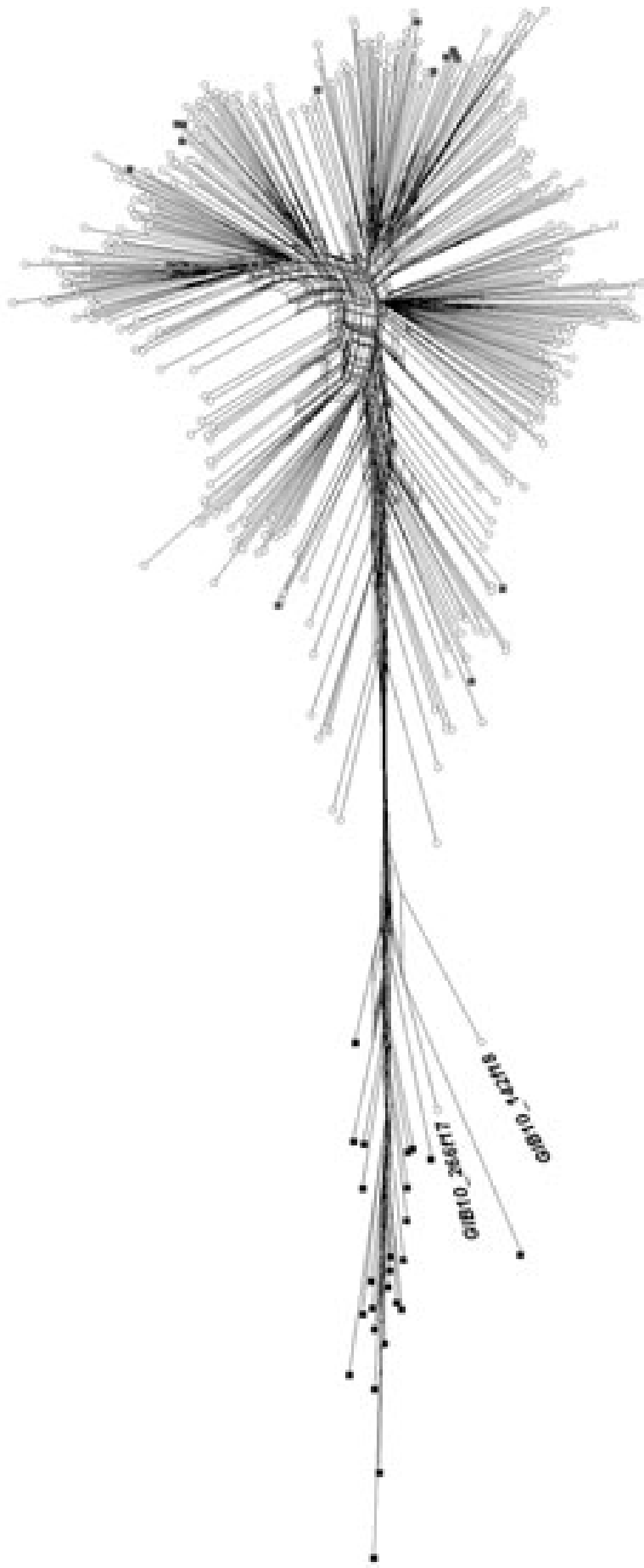
Figure 8.2 Phenogram (NeighborNet) for lexical questionnaire ratings for GibE (all data; BrE (grey squares) and AmE (black squares) as reference varieties; node labels masked).

display suggests that these ties indeed persist in terms of lexical choices when BrE and GibE are compared. An additional inference in terms of Britishness vs. Americanness that we can draw concerns colonial lag, as some of the GibE raters appear even 'more British' (or more traditionally so) than the British themselves. That is, the overall distance between some of the GibE raters to the AmE cluster is bigger than the one observable for some of the BrE raters (on which see further below and Section 4).

As described elsewhere (Krug, Schützler and Werner 2016: 43), one application of our exploratory aggregative analysis is the identification of outliers. With the help of the phenogram shown in Figure 8.2, we are in a position to identify two obvious outliers in terms of associations between GibE and AmE raters at a glance. GibE raters 266 and 142 (labelled nodes in Figure 8.2) cluster with the AmE nodes towards the left of the phenogram. Possible attempts at explaining this surprising association may profit from relating back to the sociolinguistic information elicited by the questionnaire, which, unfortunately, is incomplete for rater 266: the only possible rationale for his AmE-like rating pattern can be found in the reported Irish (hence probably Irish English) language background of the rater's father. Since we have of course no information on the vita of this rater's father, this interpretation remains speculative to a certain extent. Nor can a potentially relevant explanation be provided for rater 142, due to the fact that most sociolinguistic information is actually missing and what is provided (e.g., Spanish-English language background, three years abroad in England as a child) does not enhance our understanding.

We have to accept, then, that for these particular raters, our approach reaches its limits. We nevertheless decided to include the two outliers in the analysis as we have no indication of extended language contact with AmE. Also, due to their minimal relative weight (2/312, i.e. 0.6 per cent of the data), their diverging behaviour will have only a very minor impact on our overall results, as all remaining GibE raters cluster around a common centre (see below). The noise which these two outliers potentially introduce into further analyses (see Section 4) is therefore negligible. Other researchers, depending on their aims, might proceed differently, especially if an extended number of outliers occurs in their data (see Krug, Schützler and Werner 2016: 43–4 for discussion).

3.5. Refining the display

In Figure 8.3, following suggestions developed in Krug, Schützler and Werner (2016: 43), we add three modifications. As the representation in Figure 8.2 is close to being visually overcharged, we decided to present additionally a random sample of fifty GibE raters, which leads to greater clarity of exposition. Furthermore, we introduce nodes for a hypothetical 'average' rater for each variety (labelled AVER_* in Figure 8.3), which represents a meta-category based on average values across all relevant raters for all of the sixty-eight lexical items. Lastly, we adjust the node design for the GibE raters according to gender (female = filled circles, male = unfilled circles) as a test case for identifying potential effects of sociolinguistic factors on the clustering.

First of all, even though the hypothetical average raters for Gibraltar (AVER_GIB) and Britain (AVER_GB) display a similar distance from the American (AVER_US) average rater in Figure 8.3, it is the American and British average raters that emerge as
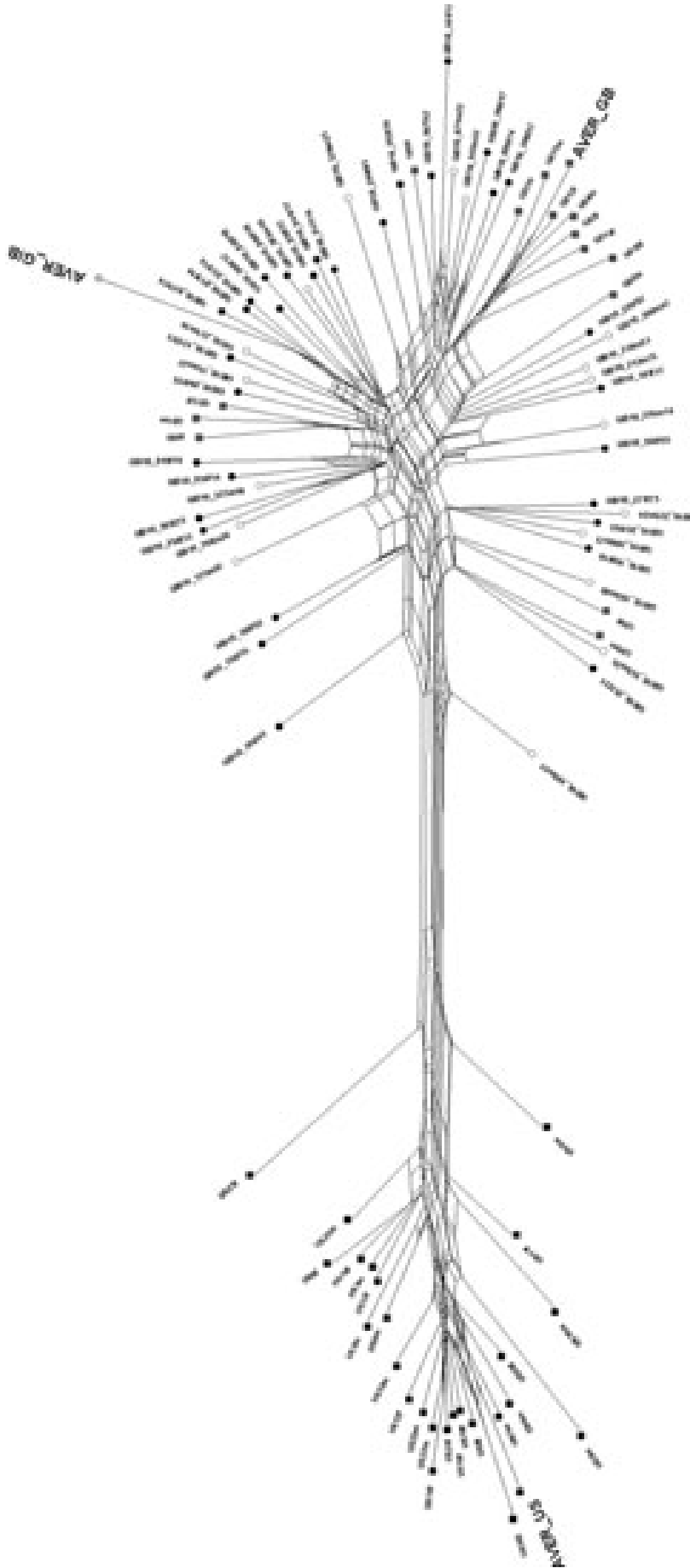
Figure 8.3 Phenogram (NeighborNet) for lexical questionnaire ratings for GibE (sampled data; female = filled circles, male = unfilled circles; BrE (grey squares) and AmE (black squares) as reference varieties). [Node labels appear in the format Variety(Year)_RaterGender(Age), e.g., GIB10_253f31]

poles in the dataset, as they exhibit the maximal distance among the regional averages. AVER_GIB and the majority of GibE raters lie somewhere between these poles, although a few GibE raters are even more distant from AVER_US than AVER_GB is. This result is in line with our earlier hypothesis that 'the average L2 user will figure between the major poles of the average BrE and AmE language users [so that] in graphical representations of large and varied datasets, L2 varieties will typically be found in the space *between* BrE and AmE' (Krug, Schützler and Werner 2016: 45; emphasis added). The hypothesis that (raters of) L2 varieties are influenced by linguistic contact with both BrE and AmE (i.e. one form of linguistic globalization) can therefore be extended to the postcolonial language-contact scenario investigated in the present dataset.

In what follows, we will continue to integrate our present findings into the fourfold typology of linguistic globalization established in Krug, Schützler and Werner (2016: 55), which we repeat here for convenience:

1. Convergence (understood as 'a process of increased or increasing similarity') on (former) British norms.[10]
2. Convergence on (former) American norms.
3. Convergence on free variation between (former) British and American norms.
4. Convergence on a norm that is neither a former (standard) British nor a former (standard) American norm, but potentially originates in a low-prestige dialect or L2 variety.

It has been argued that for socio-cultural reasons English in Gibraltar is used as a 'powerful means of expressing affiliation to Britain and proving an ethnic identity separate from Spain' (Kellermann 2001: 412). On a related note, Kellermann (2001: 414) diagnoses a continuing stigmatization of deviations from the exonormative BrE model. Such input- and usage-related aspects, ultimately traceable to issues of linguistic identity, at least partly serve to explain the very close association between GibE and BrE in our dataset. The same applies to the fact that some GibE raters exceed the 'Britishness' values of the average BrE rater, while most show a globalizing tendency that places them somewhere between the British and American poles in our dataset. On a more general note, our findings suggest that individual (post-)colonial scenarios and their potential linguistic consequences (cf. Schneider 2007) have to be considered when attempting an interpretation of empirical data in terms of nativizing or globalizing tendencies. We will explore this further in Section 4 when we treat individual items. In the present case, the late, comparatively slow and as yet unfinished decolonization of Gibraltar (Kellermann 2001: 140) seems to have led to a persisting linguistic orientation towards BrE, which in all likelihood is motivated essentially by political circumstances (Weston 2011: 365, 2015; Levey 2015: 51, 57).

A secondary finding that emerges from the display in Figure 8.3 relates to the issue of gender as an often-described sociolinguistic factor. The representation in Figure 8.3, with male nodes in unfilled circles and female nodes infilled circles, yields no clear pattern or clustering. This analysis suggests that gender is not an immediately obvious factor in the present study, which may appear surprising at first sight. Note that we arrive at similarly indistinct displays if we (re-)shape the nodes of the GibE raters for

language background or age. (We refrain from displaying such non-results here.) In our more fine-grained regression analysis in the following sections, however, we will see that speakers under seventy diverge subtly and that our male raters show stronger tendencies than our female raters of the second globalization type sketched above, in moving slightly away from rigidly British linguistic behaviour into the direction of American behaviour.

4. Rater means, individual items and sociolinguistic factors

This part of the study first re-inspects some issues discussed in the context of the explorative analysis from a slightly different angle, using mean values instead of distance metrics. On this basis, a more fine-grained analysis of individual lexical binaries can be presented. We believe that in complex datasets like the one we are looking at, both approaches – the one presented above (hypothesis-generating) and the one shown in this section (hypothesis-testing) – are not only legitimate but necessary: using explorative aggregative analyses in combination with approaches that are more transparently grounded in the underlying data may reveal different aspects of variation and change. Therefore, the following analyses will refer back to what has been stated above, where appropriate.

4.1. Methodological notes

In what follows, two methods of quantification need to be understood and kept distinct: (i) rater means, which are average values calculated for individual raters from all sixty-eight ratings and (ii) item means, which are average values calculated for individual lexical items across all raters in the respective sociolinguistic subgroup. Whenever item means are plotted, their number will invariably be sixty-eight, while the number of raters varies depending on the subsample inspected. Results based on item means will normally not be in conflict with results based on rater means – rather, both approaches highlight somewhat different aspects of the same dataset. Therefore, in some plots they will be shown in combination. As discussed above, in the explorative aggregative analyses ratings were transformed into a scale with $1 \leq rating \leq 5$ for methodological reasons. In the following paragraphs, ratings are centred round zero, that is, they take values of $-2 \leq rating \leq +2$. This way, values below zero indicate a preference of the lexical variant traditionally associated with AmE (e.g., *truck*), while values above zero indicate a preference of the (traditionally) BrE variant (e.g., *lorry*), which is a more intuitively plausible representation of results. In this case, missing values are not a problem, and we therefore work with the original data and do not make use of imputation for missing values.

4.2. Global comparison of varieties

As a starting point for this part of the analysis, we compare GibE to the two reference varieties (BrE and AmE) at a global level. In Figure 8.4, the horizontal bar in each plot
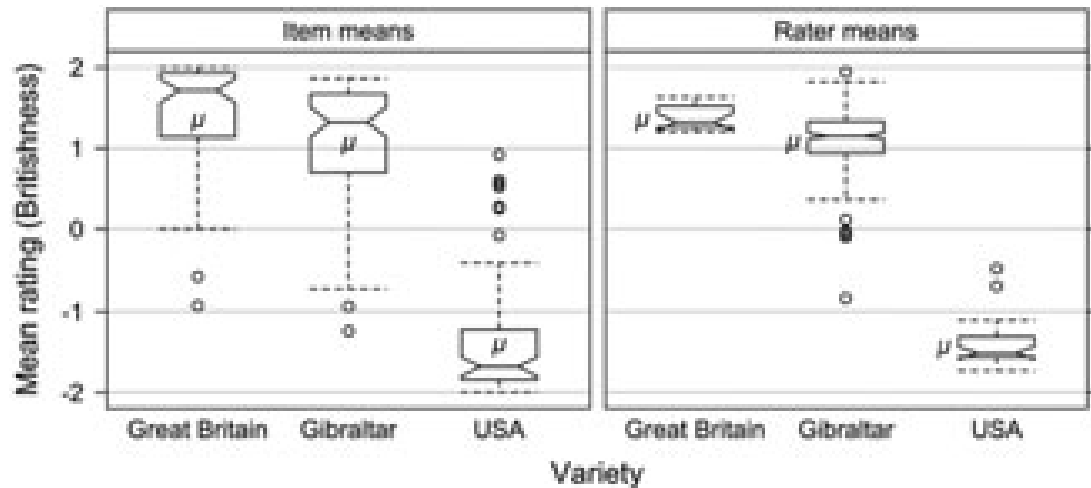
Figure 8.4 Global comparison of GibE, BrE and AmE.

represents the value of the median and the symbol μ additionally indicates the position of the arithmetic mean. Each box contains the central 50 per cent of cases and the whiskers embrace the remaining two quartiles, excluding outliers. The notches in each plot can be treated as confidence intervals, indicating a significant difference between groups if they do not overlap.

The left-hand panel of Figure 8.4 shows the sixty-eight item means for all three varieties, with GibE positioned between the major standard dialects. The right-hand panel makes the same comparison based on rater means. Both panels show even more clearly than the distance matrix in Figure 8.3 above the intermediate position of GibE when compared to BrE and AmE. Individual items and individual speakers are not considered here, although important outliers will be discussed briefly below. Two aspects of variation are readily apparent. On the one hand, there is a clear ordering of varieties in terms of their general degree of Britishness, irrespective of which of the two views on the data we adopt: British raters tend to award higher (i.e. more British) values to the lexical items under investigation and American raters on average award the lowest values (close to −2), which indicates that rather clear norms of usage exist with regard to the phenomena under investigation. Secondly, Figure 8.4 confirms what was also evident in the NeighborNets shown in Figures 8.2 and 8.3 above, namely that GibE is very similar to BrE as far as our lexical items are concerned. When we base the analysis on rater means instead (right-hand panel of Figure 8.4), the dispersion of values becomes considerably more compact, but does not change fundamentally otherwise. Importantly, in either panel of Figure 8.4 the notches of the boxes do not overlap, which is interpreted as indicating statistically robust differences between all three varieties. In addition, for both items and raters we can observe that the interquartile ranges (the heights of the boxes) are greatest for GibE. This ties in with a pattern which we have observed before (Krug, Schützler and Werner 2016: 61), *viz.* that our questionnaire data from L2 varieties (and, by concomitance, L2 speakers) are less uniform than those from standard BrE or AmE (and their raters). It was argued above that GibE is not a typical L2 variety of English. Based on the current data, we therefore expand our hypothesis from L2 varieties to postcolonial varieties in the present context of globalization.

A few notable outlier values need to be discussed. In the plot of individual items the two outliers in BrE are *to licence*, whose putative American variant *to license* is in fact preferred (mean rating = −.93), and *compare X to Y*, whose variant *compare X with Y* is preferred (mean rating = −.57). In GibE, the two outliers are *sport*, with a mean rating of −1.24 (i.e. a clear preference of *sports*), and *to let*, with a mean rating of −.94 (reflecting a preference of *for rent*). In AmE, the four most notable outliers are *typical of* (preferred to its supposedly AmE variant *typical for*, with a mean rating of +.92), *subway* (preferred to the allegedly more AmE variant *underpass*; mean rating = +.60), *backwards* (preferred to the variant *backward*; mean rating = +.56) and *a book about chemistry* (preferred to the variant *a book on chemistry*; mean rating = +.52). Those (and other) items are plotted in detail in Figure 8.5 below, which is followed by further discussion. What we learn from such results is that dictionaries, textbooks and style books often simplify or represent an earlier historical state when labelling an item as 'BrE' or 'AmE'.

The panel based on rater means in Figure 8.4 features one outlier rater in GibE with exceptionally low mean ratings; this is the same rater discussed in the context of the explorative analysis in Section 3. Another rater is exceptional in reporting to use only the BrE variant for sixty-two out of sixty-eight items, thus obtaining sixty-two times the rating +2 (while skipping five items and reporting to use only *baked potato*, the more AmE variant), which results in an overall mean of +1.94 for those items that were evaluated. In this case, a closer inspection of his personal information sheet is instructive. This rater is an extremely highly qualified forty-nine-year-old male lawyer holding a PhD. Interestingly, both parents' native language is given as Spanish, while at the same time English was the only language used at home while growing up. Not surprisingly, this rater lived in the UK from age 19 to 29, where in all probability he received his university education and professional training, as there was no tertiary education in Gibraltar at the time. The consistency with which this rater opted for the BrE variants is nevertheless astounding; apparently, he has access to an extremely clear intuition regarding the difference between the two variants that were offered. What probably underpins his linguistic Britishness is that he lived in the UK during the pre-internet 1980s, that is, at a time when the distinction between our 'British' and 'American' variants was still more categorical than in the more globalized, early twenty-first century. Nevertheless, language contact with twentieth-century BrE alone would probably not have effected such categorical usage. More likely, his linguistic sensitivity was further strengthened through subsequent extended periods in international contexts outside Britain: from age 30 to 35 he lived in Hong Kong and the following seven years in Finland. An additional likely factor is that both parents of our informant were teachers. These would belong to a less globalized generation and, on a more speculative note, were probably aware of transatlantic lexical differences and followed (and maybe passed on to their son) a rather traditional British model.

4.3. General behaviour of individual lexical items

In Figure 8.5, the item means of all sixty-eight lexical binaries are shown for the three varieties included in this study. Items are not ordered alphabetically or according to
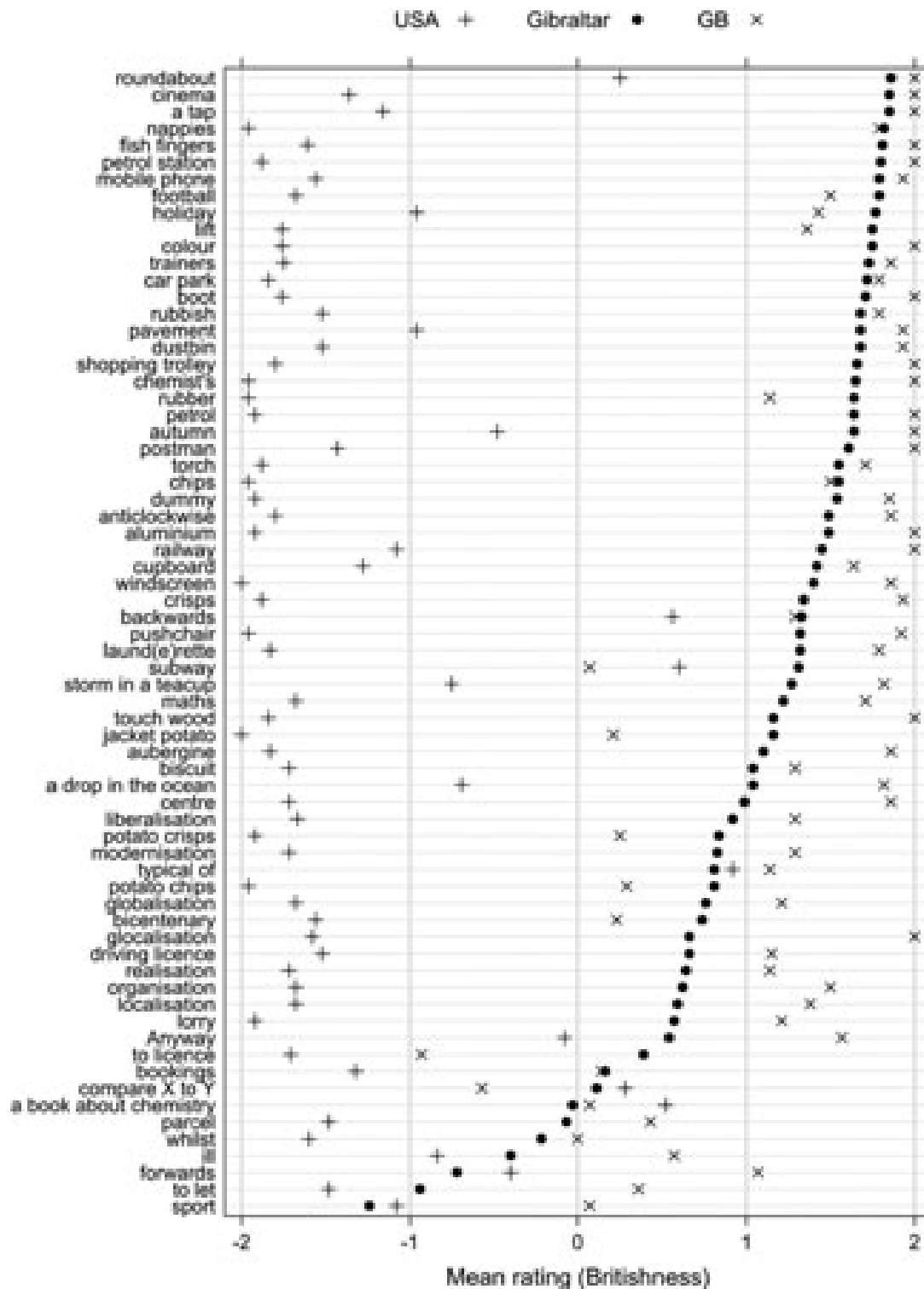
Figure 8.5 Ratings for individual items in AmE, GibE and BrE.

their arrangement in the questionnaire, but by their mean rating in GibE – that is, items appearing at the top of the figure received the highest average rating by GibE raters. The plot is useful for three purposes. First, it provides a vast amount of descriptive detail concerning specific lexical preferences across varieties and items. Second, individual items not in accordance with the generally expected pattern (that is, with

ratings in Gibraltar not intermediate between the other two varieties; cf., Figure 8.4) can be quickly identified. And, finally, the figure provides yet another perspective on the relatively close alignment of GibE and BrE, on the one hand, and the rather different preferences in AmE, on the other.

The outlier items that were discussed above can be clearly seen in this display, the additional advantage being that their positions in all three varieties are apparent at a glance. For example, it is evident that the two BrE outliers *to licence* and *compare X to Y* – with ratings of −.93 and −.57, respectively – are positioned relatively far towards the bottom of the plot. This indicates that in GibE those two items are also not very strongly drawn towards the supposedly BrE pole of the scale and that their relatively loose orientation towards the putative BrE norm may be more general. *Compare X to Y* takes intermediate values close to zero in all three varieties and appears to be in fairly free variation, that is without particularly clear preferences of either form. Items that are given ratings below zero in GibE are *sport* (−1.24), *to let* (−.94), *forwards* (−.72), *ill* (−.40), *whilst* (−.22), *parcel* (−.07) and *a book about chemistry* (−.03), that is, the seven items at the bottom of Figure 8.5. If zero is taken as an arbitrary threshold value, those items can be taken as a set that deserves further investigation.

There are fifteen items whose behaviour is exceptional in that they receive clearly higher (i.e. more British) ratings in GibE than in BrE: *football*, *holiday*, *lift*, *rubber*, *subway*, *jacket potato*, *potato crisps*, *potato chips*, *bicentenary*, *to licence*, *compare X to Y*, *nappies*, *chips*, *backwards* and *bookings*.[11] *Jacket potato* also featured in the context of the outlier rater discussed above, who only made an exception to his otherwise perfectly uniform ratings (+2) for this item. The patterns found for *typical of* and *compare X to Y* are striking: the former is rated lowest in GibE, that is, even our AmE raters tend to prefer the supposedly BrE variant more strongly. It has to be noted, however, that ratings in all three varieties are relatively close together (between +.81 and +1.14). Therefore, the important aspect of the pattern is perhaps not so much the cross-varietal ranking but the fact that *typical of* seems to be generally preferred irrespective of variety.

In contrast, the pattern found for the item *compare X to Y* is truly puzzling, since it is an inversion of what would be expected: ratings are lowest (i.e. least British) among BrE raters and highest among AmE raters, with GibE closer to the latter. There may be other factors at work here, which complicate (or even confound) the analysis of this particular item. For example, the two options that were offered (*compare to* vs. *compare with*) may not be perfect synonyms in all varieties investigated but may involve semantic nuances that made a straightforward decision difficult for raters. There is also *compare X and Y* as a third option, which, although not stated in the questionnaire, may have affected ratings in a way not transparent to the researcher. In any case, which of the three options is truly more British or more American in present-day usage is far from clear and deserves further investigation.

Concerning the global picture, the data shown in Figure 8.5 contribute yet another facet to the explorative analysis presented above. There is a very clear positive correlation between individual items' mean ratings by GibE raters and BrE raters (Pearson's $r = .71$; $p = .000$). In Figure 8.5, this is reflected in the fact that towards the bottom of the plot, ratings given by both GibE and BrE raters clearly tend towards the

left-hand (i.e. more AmE) pole of the scale. There is a negative correlation between BrE and AmE ratings, which is also statistically significant but considerably weaker ($r = -.25$; $p = .036$). Finally, the negative correlation between GibE and AmE ratings is of a similar nature as the one between BrE and AmE ratings, but it is weaker and fails to reach statistical significance by a very small margin ($r = -.23$; $p = .056$). Those findings once more confirm the general tendency of GibE to pattern like BrE, but also to be positioned somewhat closer to AmE.

The results presented in Figure 8.5 suggest that the global differences between BrE and GibE on the one hand vs. AmE on the other are a function of clearly patterned preferences associated with specific lexical items. In other words, certain items appear to serve as variety markers by diverging strongly between the two linguistic (and cultural) spheres influenced by Great Britain and the US, respectively, while others do not show this kind of divergence. The former tend to be found nearer the top of Figure 8.5, the latter nearer the bottom. The results for individual items and the correlation measures that were applied not only confirm the global analysis summarized in Figure 8.4 but also corroborate what was revealed in the explorative analyses: there is a much closer association between BrE and GibE, in this case based on specific items.

At a more general level, Figure 8.5 casts serious doubts on standard classifications of certain lexical variants as either 'British' or 'American'. Binary classifications obviously do not work; statistical preferences are the rule rather than the exception. Furthermore, based on our data, some items' label as (more) 'British' or 'American' will even have to be reversed, at least for an adequate description of English(es) in the early twenty-first century.[12]

4.4. Apparent-time trends

The synchronic lexical differences between GibE and the two reference varieties are an important step forward in the description and contextualization of GibE among World Englishes. However, we are also interested in the emergence and (ongoing) development of those differences. The present section focuses on this aspect by inspecting lexical variation in the GibE ratings with the help of apparent-time studies.

Figure 8.6 shows average ratings across seven age groups, again based on item means and rater means (for an identification of individual items see below). The picture that emerges is not suggestive of any strong ongoing process of change. There appears to be an Americanizing trend in apparent time, but this pattern is not very pronounced, and the oldest group of raters does not conform to it.

For a more fine-grained picture, male and female raters were separated. This was also important to control for the gender imbalance among the youngest raters (see Table 8.1 above). Moreover, precise ages, not age groups, were used. In Figure 8.7, rater means of male and female raters are plotted using different symbols; a smoothed local regression line indicates general tendencies within both groups.

Among male raters (see dashed regression line in Figure 8.7), there is a fairly monotonic, if not particularly steep, downward trend towards less British-oriented preferences. Female raters appear to follow this trend only hesitatingly: only young raters are affected, while the generally high preference of BrE variants is rather robust
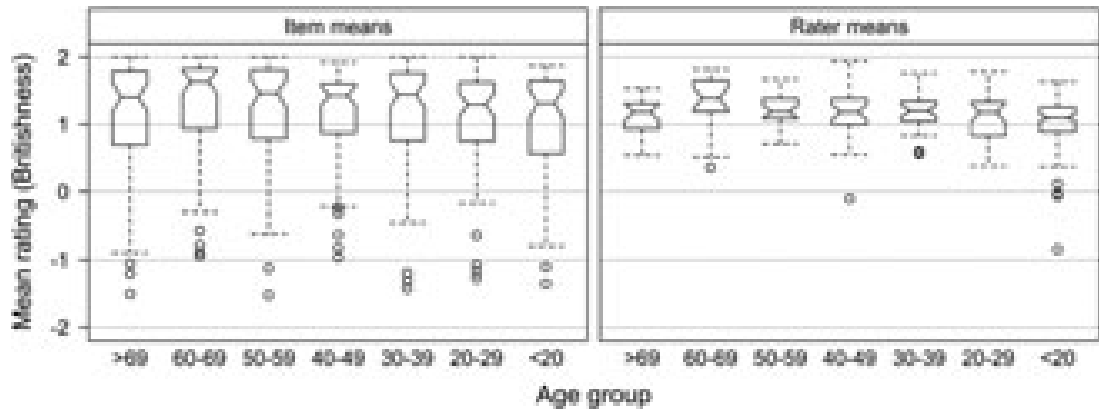
Figure 8.6 Differences according to age groups.
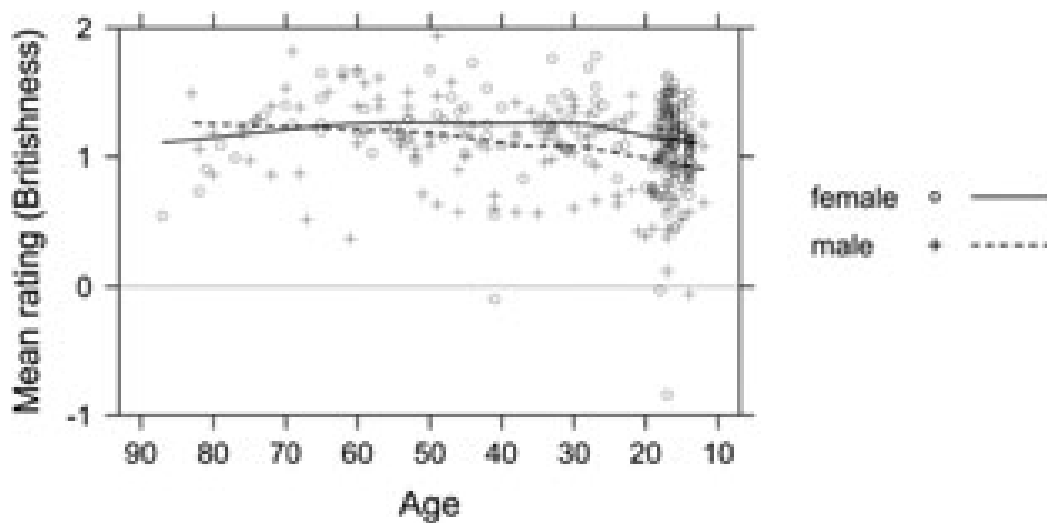
Figure 8.7 Rater means by age and gender; general trends indicated by smoothed regression lines.

for older groups. Female raters over seventy have strikingly low average values, which results in the slope of the smoothed regression line for this portion of the data. However, not too much weight should be placed on this finding, since there are only few data points for the relevant age range. The pattern may be explicable in terms of the amount of formal education that the oldest raters have received: for eight out of nine female raters aged seventy and above – including all those older than eighty – the respective socio-linguistic information ('qualification') is not available (i.e. it was not volunteered). This can be tentatively interpreted as indicating that the respective raters may indeed have undergone less formal education and qualification, resulting in lower linguistic awareness and looser norms of usage.

When seen as an instance of linguistic change, the pattern displayed by the more innovative Gibraltarian males may, *prima facie* at least, appear surprising, considering that it is often females who are in the vanguard of linguistic changes (see, e.g., the synopsis in Labov 1990; or Nevalainen 1996; Raumolin-Brunberg and Nurmi 1997).

What we observe is a particular (and less often described) kind of change: on the increase among males is a competing variant which diverges from the traditional BrE norm and whose status in terms of overt vs. covert prestige is in all likelihood different for different social groups. In this case, therefore, the Gender Paradox as discussed by Labov (1990: 213–15) does not help to explain the pattern. In essence, then, there appears to be a slow-moving general trend towards more globalized (i.e. less rigidly British-oriented) usage, but the older BrE norm is largely intact, and it is the female raters who adhere to it somewhat more strongly. An alternative – or rather complementary – sociolinguistic explanation considers the interaction between social circumstances and age. Speakers between the age of thirty and fifty-five have been shown to peak in the usage of (in our case, BrE) standard variants. This 'sociolectal retrenchment' (Tagliamonte 2012: 49) has been ascribed to increased social pressure during that period, caused, for instance, by starting a family, raising children or career advancement (see also Chambers 2009: 195). Female speakers in our sample (see Figure 8.7) seem to conform to this pattern, while it is less pronounced for the males.

Rater age and rater mean are positively correlated for male raters (Pearson's $r = .31$; $p = .000$). In other words, the older the male raters, the higher (i.e. more British) their average ratings become. A linear regression model (in which the effect of age is naturally also significant) predicts a decrease in average rating of 0.55 over a time span of 100 years for males. For female raters, the correlation of rater age and rater mean fails to reach statistical significance (Pearson's $r = .129$; $p = .087$). Results from a linear regression model suggest that there is a decrease of 0.23 for women over a period of 100 years. In sum, there is a very clear and consistent (if not particularly dramatic) Americanizing tendency among male Gibraltarians at this global level, looking at rater means only; more recently, a similar process seems to affect females as well.

In Figure 8.7, only five female raters over seventy-five with their relatively low average ratings (including a nearly ninety-year-old with particularly low overall means) are responsible for the curvature of the regression line at the highest age values. In addition, among themselves, these five female raters exhibit a trend which is the inverse of the general tendency, that is, the older ones prefer the putatively AmE variants. With only five informants each for raters over seventy-five, sparse data are in fact a problem for both genders in our dataset. If we concentrate in our analysis on speakers aged seventy-five and younger, as displayed in Figure 8.8, developments in apparent time can be interpreted with more confidence.

For raters up to seventy-five, gender-based linear regression models yield statistically significant findings that are also more readily interpretable: if we fit separate models to the data for males and females, the estimated expected rating at age seventy-five is remarkably similar for both groups, namely 1.307 for males and 1.308 for females. This is also clearly reflected in the starting points of the smoothed regression lines for male and female raters in Figure 8.8. The plot also indicates that ratings expected from men and women follow monotonic negative trajectories that gradually diverge from each other as younger speakers are inspected; as in the previous analysis discussed above, the downward (i.e. Americanizing) trend is also somewhat more pronounced among male raters (coefficient = .0061) than among female raters (coefficient = .0034).[13] The
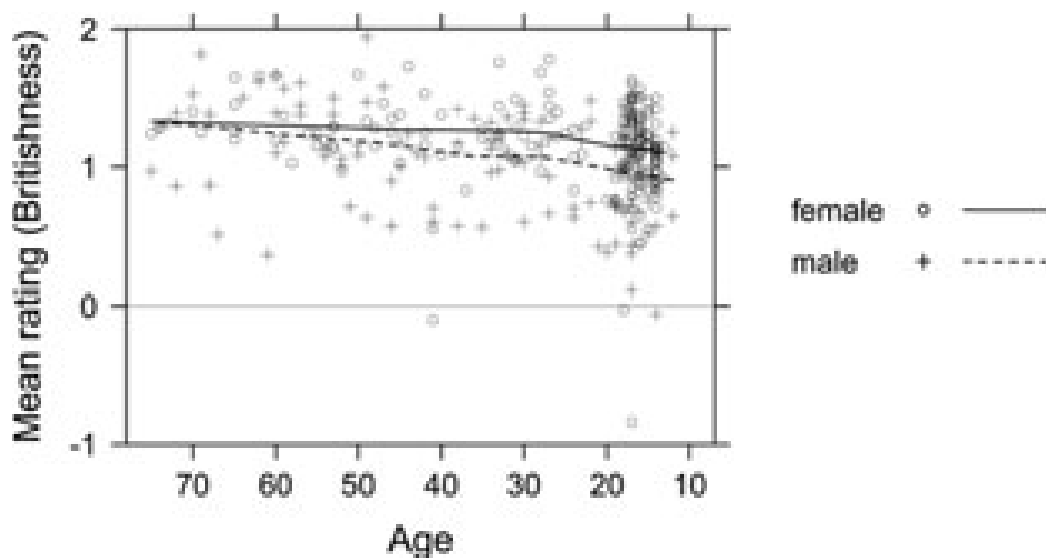
Figure 8.8 Rater means by age and gender; general trends indicated by smoothed regression lines for speakers aged up to seventy-five.

coefficients of the linear regression indicate to what extent average ratings of our lexical items change per year; the yearly decrease for men (from older to younger informants) is about twice that for women. Thus, over an eighty-year span, men's average ratings decrease by 0.49, while women's average ratings decrease by 0.27. As in the previous regression analyses based on the entire dataset, therefore, female raters diverge from the older British norms somewhat more slowly than their male counterparts. And while both analyses converge in suggesting Americanizing (or globalizing) tendencies for all Gibraltarians, the analysis based on the entire dataset and displayed in Figure 8.7 (and the contingent linear regressions) suggests a later onset of Americanization for women by about ten years. Due to the reasons advanced above (first and foremost, data scarcity for over-seventy-five-year-olds), however, we diverge from the usual convention and give preference here to the results obtained through the analysis of the incomplete dataset.

4.5. Language backgrounds

Figure 8.9 displays the distributions of item means (average of all sixty-eight binaries) for each of the six language background groups identified in subdivisions A and B of Table 8.2 above. No significant pattern emerges, irrespective of whether we use Subdivision A or Subdivision B for the categorization of language backgrounds detailed in Table 8.2.

Figure 8.10 displays rater means for each of the six language background groups identified in Table 8.2. Again, no significant pattern emerges if we apply Subdivision A. However, if Subdivision B is applied, the group of raters with mixed language backgrounds (i.e. those whose household usage is not purely English or Spanish) do exhibit some tendency to Americanize when compared to the other two, the picture being somewhat obscured by the presence of outlier values, however. If we trim the sample by omitting all outlier raters seen in the right-hand panel of Figure 8.10, the
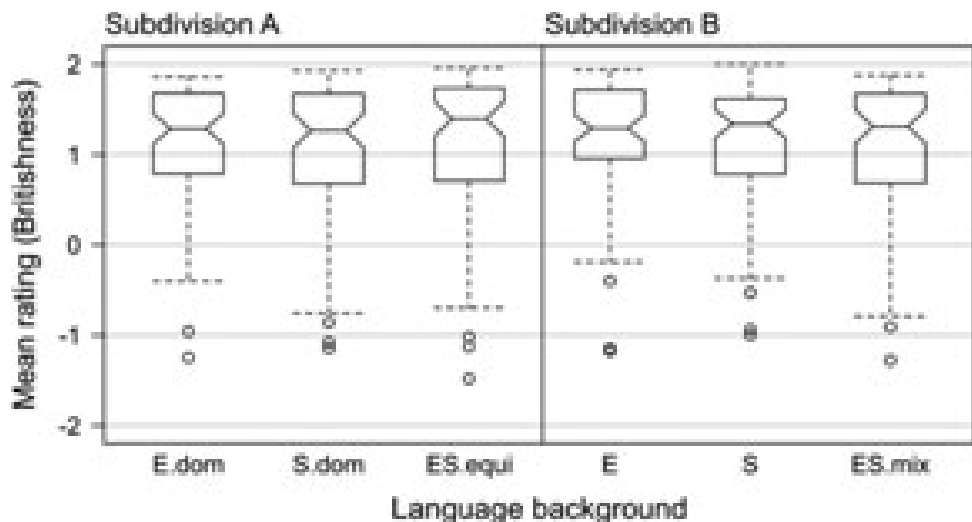
Figure 8.9 Item means (average of all sixty-eight items) by language backgrounds (six groups).
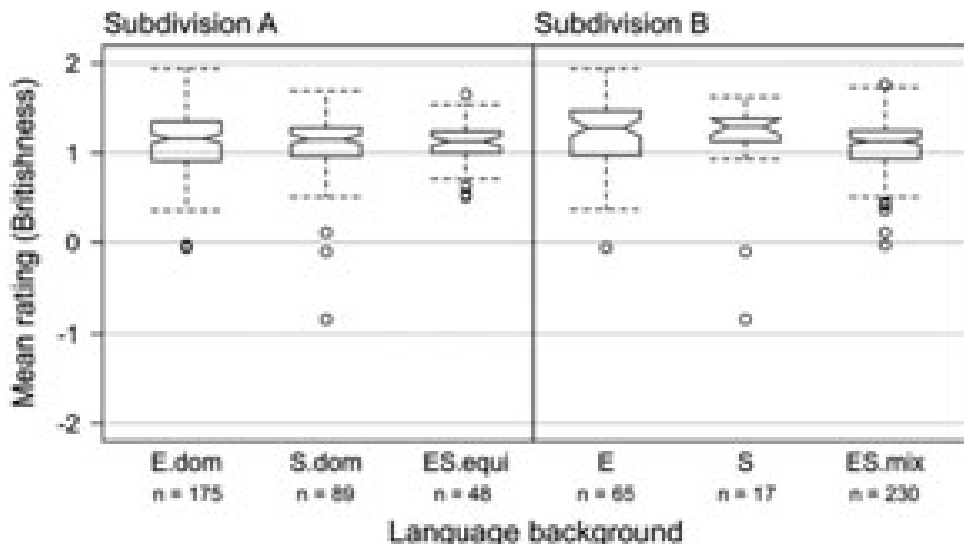
Figure 8.10 Rater means by language backgrounds (six groups).

respective arithmetic mean values for groups are $\mu_E = 1.21$, $\mu_S = 1.30$ and $\mu_{ES.mix} = 1.10$. Two-tailed *t*-tests applied to the three possible pairings of groups yield the following results: there is no significant difference between groups 'E' (English-only) and 'S' (Spanish-only), with $t(47.6) = 1.55$ and $p = .127$; the difference between groups 'E' and 'ES.mix' is statistically significant, with $t(82.3) = 2.09$ and $p = .040$; and the difference between groups 'S' and 'ES.mix' is also significant, with $t(18.7) = 4.23$ and $p = .000$. What is interesting about these results is that raters of mixed language background do not take a position intermediate between the purely monolingual raters. Whenever there is a mix of English and Spanish spoken at home, speakers exhibit a more globalized usage of our lexical items than in either English-only or Spanish-only familial language contexts. This is somewhat difficult to interpret but seems to suggest

that when English or Spanish are the only languages spoken in the home, language contact (or the model adopted) is more exclusively British than when both Spanish and English are spoken at home. In the latter cases, on the basis of our findings, there should be more extensive contact with American and global Englishes.

4.6. Individual lexical items in apparent time

In the final part of our analysis we inspect the likely diachronic behaviour of individual items based on an alternative apparent-time approach: first, the item mean for all sixty-eight items was calculated for each of the seven age groups defined in Table 8.1 above; that is, a total of 476 mean values. For each item, the seven values were correlated with the seven age groups (i.e. the integers 1–7), using Pearson's product-moment coefficient.[14] Items were then ranked and ordered based on the correlation coefficient, as documented in Appendix B. A positive correlation indicates a development towards more British usage, while a negative coefficient indicates Americanization. Most correlations (43/68 = 63 per cent) are negative (i.e. Americanizing), although one needs to bear in mind that many patterns hardly suggest a trend towards more British or more American usage, even if the associated correlation coefficient naturally will be positive or negative. Also note that a high correlation coefficient merely reflects linearity and consistency of a pattern and does not necessarily indicate a strong diachronic effect.

Diachronic developments of all sixty-eight items are plotted in Appendix C. Here we discuss only those twenty-one items whose mean ratings correlated significantly (or nearly significantly) with rater age ($p < .1$). The details for age-related differences of these items are presented in Figure 8.11.

- Items with positive trends (becoming more British) include *jacket potato*, *maths*, *subway*, *whilst*, *book about chemistry* and *cinema*.
- Items with negative (Americanizing) trends include *to licence*, *biscuit*, *Anyway, . . .*, *storm in a teacup*, *lorry*, *parcel*, *drop in the ocean*, *nappies*, *driving licence*, *dustbin*, *petrol station*, *pushchair*, *postman*, *shopping trolley* and *railway*.
- NB: The four items *cinema*, *railway*, *shopping trolley* and *postman* are merely 'trending' in that the correlation with rater age is marginally non-significant ($.5 \leq p < .1$).

The six items moving towards the BrE pole of the scale can be subdivided into three sets:

- *Jacket potato*, *maths* and *subway* start from an intermediate value that reflects relatively free variation between the AmE and the BrE variant, and undergo very clear processes of change (*jacket potato*: +.20 → +1.39, with an even slightly higher peak at +1.41; *maths*: +.40 → +1.40, with an even higher peak at +1.50; and *subway*: −.22 → +1.59, peaking at +1.79).
- *Whilst* and *a book about chemistry* move towards the 'middle ground', starting from the American side of the continuum (*whilst*: −.89 → −.04; *book about chemistry*: −.78 → +.12, if we look at the extreme points).
- *Cinema* is already very British-oriented but is further consolidated, with scores between +1.56 and +2.00.
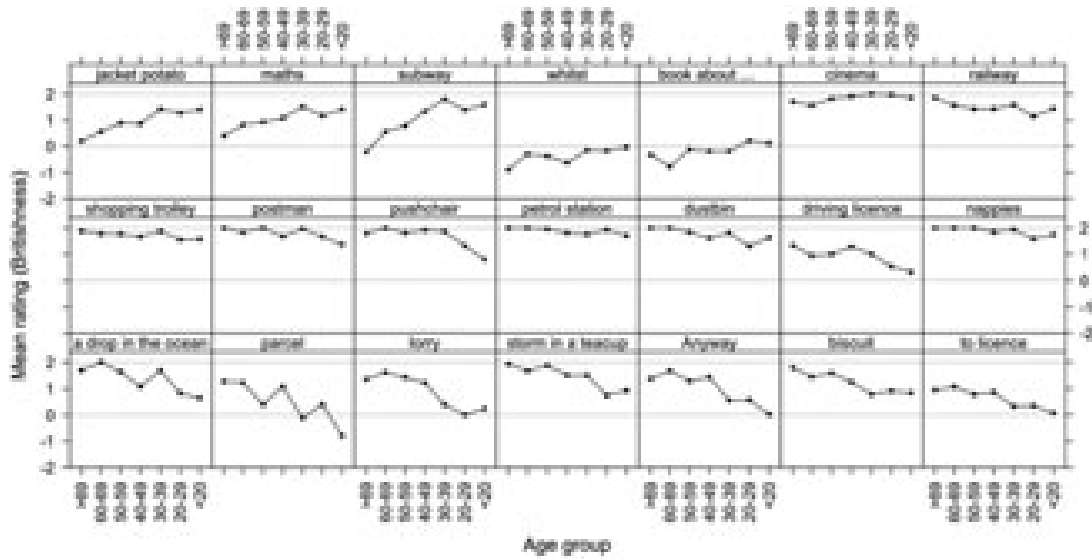
Figure 8.11 Apparent-time trends for selected items in GibE.

On the other hand, those items that become more Americanized over (apparent) time invariably start at rather high levels of Britishness. For some there is only a minor loosening of their orientation towards BrE norms (*railway*, *shopping trolley*, *postman*, *petrol station*, *nappies*), while others undergo more dramatic changes (e.g., *pushchair*, *driving licence*, *parcel*, *lorry*, *Anyway, . . .*). The three items *shopping trolley*, *petrol station* and *nappies* remain clearly oriented towards the BrE norm, but the significant small-scale trends that we detect in them nevertheless indicate that this norm is beginning to weaken. The item *pushchair* has rather stable mean values for older raters, followed by a very rapid change in the two youngest age groups (i.e. raters under thirty), the preference of the BrE variant dropping from a nearly categorical value of +1.87 to +.80. Finally, *parcel* and *lorry* are noteworthy for the particularly strong overall change they undergo, with values dropping from +1.26 to −.79 for *parcel* and from +1.61 to ±0 for *lorry* (if we look at the peak values of the latter).

5. Discussion and aspects of globalization

Questionnaire results reflect individuals' intuitions about language rather than actual performance, which for some linguists working in a usage-based framework casts serious doubts on their validity (cf., Dollinger 2015: ch. 3 for related aspects and ways to counterbalance potential shortcomings of questionnaires). The current project has addressed such issues by collecting large numbers of questionnaires for our contact-varieties (thus de-individualizing the intuitions collected and making the data amenable to sociolinguistic analysis) and by comparing, in an earlier publication, questionnaire-based results with web-based data (see Krug and Sönning 2017). The two approaches yielded highly significant correlations between rank-based hierarchies, which makes global spurious findings for our lexical items rather unlikely. In addition,

we have included several control items in the questionnaire, which address issues relating to the reliability of rater responses (e.g., a number of *-isation* spellings spread across the questionnaire).[15]

Furthermore, the consistency with which we find correlations between parent and (post-)colonial varieties (e.g., between British and Maltese English, on the one hand, and American and Puerto Rican English, on the other; see Krug, Schützler and Werner 2016 for details) makes us optimistic as regards the usefulness of questionnaire-based research into lexical (and potentially other types of) variation. The present investigation provides further support for optimism: based on the – unsurprising but hitherto also unexplored – striking similarity between BrE and GibE for lexical binaries, it is tempting to label present-day GibE on the lexical level as Gibraltar British English, or GiBrE, for short (rather than GibE or GibrE, for instance).

What we can conclude from the previous section is that the lexical binaries investigated in this study clearly have to be treated as individual categories, both in terms of their likely diachronic behaviour and their synchronic position on the continuum between BrE and AmE. Against the background of apparent-time analyses of individual items, it is not surprising that the global approach (based on overall rater means) yields less striking results: the apparent stability at a global surface level only smoothes over multiple individual patterns, some stable, others changing in different directions, sometimes towards a more British, more often towards a more American orientation. We found five significant apparent-time trends towards a more BrE usage pattern in our Gibraltar data, most notably so for *jacket potato, maths* and *subway.* These compare with twelve Americanizing items that exceed the 5 per cent significance level, the most obvious ones being *to licence*, *biscuit*, *Anyway, . . .*, *storm in a teacup*, *lorry* and *parcel*, which are increasingly being replaced by their more American counterparts *to license*, *cookie*, *Anyways, . . .*, *tempest in a teapot*, *truck* and *package*.

Some items display similar patterns of development in many regions of the world, and can thus be interpreted more confidently as instances of globalization. On the basis of this and previous studies into other traditionally British-oriented varieties of English (in Malta, the Channel Islands and indeed the UK itself; cf., Krug and Rosen 2012; Krug, Schützler and Werner 2016), we can assume globalization in the guise of Americanization (our globalization type 2) for *lorry*, *parcel* and *to licence*, onto whose territory encroach their American counterparts. Another example is *sport*, which is being superseded by *sports* in many regions of the world, too. In our Gibraltar data, there is only a statistically non-significant apparent-time trend towards *sports* (see Appendices B and C). Significantly from a linguistic perspective, however, all age-groups display solidly negative (i.e. more American) usage ratings for this item. Seen from this perspective, *sports* is in fact the most Americanized item in Gibraltar, together with *for rent* (which is preferred over *to let* in GibE, and which shows strong signs of Americanization in the Channel Islands as well as in Malta). A similar case is *sick*, which seems to be spreading at the expense of the once more British variant *ill* in many English-speaking regions of the world. As with *sport(s)*, the Americanizing apparent-time trend is negligible in GibE but, like *sport*, *ill* belongs to the small group of only six items that have solidly negative ratings overall.[16]

Evidently, therefore, different types of globalization coexist (see our typology developed in Krug, Schützler and Werner 2016: 55, and sketched in Section 3 above), as well as different methods of identifying individual items participating in globalization. Apparent-time approaches are one such method; comparing overall item means represents an alternative (i.e. identifying which traditional British forms have the lowest means in British-oriented varieties). Where the results point in the same direction, and not only in British-oriented but also in American-oriented varieties, we can be most confident. But it would be naïve to assume that evidence will typically converge in all varieties of English around the world. What we may hope to find are regional patterns or parallel patterns of development in disparate regions of the world that can be motivated. What we must expect, however, are also distributions which cannot be motivated, and ever-changing patterns of preference. Charting such complexity, we believe, is nevertheless worth its while. It is, we would in fact argue, indispensable if we want to improve our understanding of the nature of language contact and its relationship with language change.

Appendix A. Gibraltar version of the Bamberg Questionnaire
(lexical part only)

Gibraltar English Questionnaire
Informant Information Sheet

Informant ID # _____

Date _____

Personal Information

Age
_____

Gender
○ male ○ female

Nationality
_____

Ethnic Self-identification
_____

Country or region you identify with most
_____

Language(s) used at home while growing up
○ English
○ mostly Spanish, some English
○ mostly English, some Spanish
○ Spanish
○ other: _____

Mother's native language(s)
_____

Father's native language(s)
_____

Education Profile

Primary School
○ State
○ Private
○ Church
○ other (please specify below)
_____

Secondary School
○ State
○ Private
○ Church
○ other (please specify below)
_____

Name and place of secondary school
_____

Qualifications (completed or ongoing)
○ Vocational classes (please specify below)
○ Apprenticeship (please specify below)
○ Bachelor
○ Master's
○ PhD
○ other (please specify below)
_____

Your current occupation
_____

Mother's highest qualification
_____

Mother's (last) occupation
_____

Father's highest qualification
_____

Father's (last) occupation
_____

Partner's highest qualification
_____

Partner's (last) occupation
_____

Location Timeline

Years lived outside Gibraltar
_____

Age	Location lived at from age 0–100 (please indicate city/town or country, if abroad)
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
35	
40	
45	
50	
55	
60	
65	
70	
75	
80	
85	
90	
95	
100	

Rev. 2017-10-1 JGJ

Gibraltar English Questionnaire
Lexical Items

Informant ID #: _____

Date: _____

	I always use this expression	I use this expression more often	I have no preference	I use this expression more often	I always use this expression		I never use either expression	Explanation / Comment
a drop in the ocean	○	○	○	○	○	a drop in the bucket	○	
a faucet	○	○	○	○	○	a tap	○	
aluminum	○	○	○	○	○	aluminium	○	
anticlockwise	○	○	○	○	○	counter clockwise	○	
eggplant	○	○	○	○	○	aubergine	○	(fruit/vegetable)
fall	○	○	○	○	○	autumn	○	(season of the year)
backward	○	○	○	○	○	backwards	○	
bicentenary	○	○	○	○	○	bicentennial	○	
cookie	○	○	○	○	○	biscuit	○	(something sweet to eat)
bookings	○	○	○	○	○	reservations	○	
trunk	○	○	○	○	○	boot	○	(of a car)
car park	○	○	○	○	○	parking lot	○	
center	○	○	○	○	○	centre	○	(spelling)
chemist's	○	○	○	○	○	drugstore / drug store	○	
ill	○	○	○	○	○	sick	○	
French fries	○	○	○	○	○	potato chips	○	(warm, sometimes greasy)
fries	○	○	○	○	○	chips	○	(warm, sometimes greasy)
cinema	○	○	○	○	○	movie theater	○	
color	○	○	○	○	○	colour	○	(spelling)
cupboard	○	○	○	○	○	closet	○	(for clothes)
driver's license	○	○	○	○	○	driving licence	○	
dummy	○	○	○	○	○	pacifier	○	(for babies)
trash can	○	○	○	○	○	dustbin	○	
fish fingers	○	○	○	○	○	fish sticks	○	
soccer	○	○	○	○	○	football	○	(kicking-game, only goalkeeper uses hands)
forwards	○	○	○	○	○	forward	○	
globalization	○	○	○	○	○	globalisation	○	
glocalisation	○	○	○	○	○	glocalization	○	
vacation	○	○	○	○	○	holiday	○	
liberalization	○	○	○	○	○	liberalisation	○	
baked potato	○	○	○	○	○	jacket potato	○	
laund(e)rette	○	○	○	○	○	laundromat	○	
potato crisps	○	○	○	○	○	potato chips	○	(crunchy, cold)
crisps	○	○	○	○	○	chips	○	(crunchy, cold)

Rev. 2017-04-3.202

Gibraltar English Questionnaire
Lexical Items

Informant ID # _____

Date _____

	I always use this expression	I use this expression more often	I have no preference	I use this expression more often	I always use this expression		I never use either expression	Explanation / Comment
to licence	○	○	○	○	○	to license	○	
elevator	○	○	○	○	○	lift	○	
localisation	○	○	○	○	○	localization	○	
truck	○	○	○	○	○	lorry	○	(large motor vehicle for carrying goods by road)
maths	○	○	○	○	○	math	○	
cell phone	○	○	○	○	○	mobile phone	○	
modernisation	○	○	○	○	○	modernization	○	
diapers	○	○	○	○	○	nappies	○	(for babies)
organisation	○	○	○	○	○	organization	○	
package	○	○	○	○	○	parcel	○	(something you send by mail)
pavement	○	○	○	○	○	sidewalk	○	(for pedestrians, next to street)
gasoline	○	○	○	○	○	petrol	○	
petrol station	○	○	○	○	○	gas station	○	
mailman	○	○	○	○	○	postman	○	
pushchair	○	○	○	○	○	stroller	○	(for toddlers)
railroad	○	○	○	○	○	railway	○	
realisation	○	○	○	○	○	realization	○	
traffic circle	○	○	○	○	○	roundabout	○	(for cars)
rubber	○	○	○	○	○	eraser	○	
trash	○	○	○	○	○	rubbish	○	
shopping trolley	○	○	○	○	○	shopping cart	○	
sports	○	○	○	○	○	sport	○	
storm in a teacup	○	○	○	○	○	tempest in a teapot	○	
underpass	○	○	○	○	○	subway	○	(path for pedestrians under a road)
to let	○	○	○	○	○	for rent	○	
flashlight	○	○	○	○	○	torch	○	(electric lamp)
touch wood	○	○	○	○	○	knock on wood	○	
sneakers	○	○	○	○	○	trainers	○	
whilst	○	○	○	○	○	while	○	
windshield	○	○	○	○	○	windscreen	○	
a book about chemistry	○	○	○	○	○	a book on chemistry	○	
compare X to Y	○	○	○	○	○	compare X with Y	○	
typical of	○	○	○	○	○	typical for	○	
Anyways, ...	○	○	○	○	○	Anyway, ...	○	

Rev. 2017-10-13G4

Appendix B. Correlations of mean ratings of individual items and seven age groups in GibE: Americanization (negative correlations) vs. Britishization (positive correlations)

lexical item	*r*	*p*		lexical item	*r*	*p*	
jacket potato	.94	.002	***	petrol	−.32	.487	
maths	.89	.008	**	trainers	−.34	.453	
subway	.89	.007	**	typical of	−.37	.417	
whilst	.78	.040	*	to let	−.39	.389	
book about chemistry	.77	.042	*	car park	−.47	.282	
cinema	.71	.072	.	ill	−.49	.269	
football	.63	.131		laund(e)rette	−.49	.266	
forwards	.60	.151		holiday	−.53	.220	
chips	.55	.200		bicentenary	−.55	.203	
compare X to Y	.47	.286		cupboard	−.56	.190	
mobile phone	.44	.325		windscreen	−.56	.191	
crisps	.38	.399		fish fingers	−.57	.180	
roundabout	.32	.480		aubergine	−.58	.169	
a tap	.30	.516		dummy	−.59	.165	
boot	.30	.510		potato chips	−.60	.157	
bookings	.29	.535		sport	−.60	.154	
autumn	.20	.672		modernisation	−.64	.124	
centre	.18	.696		pavement	−.64	.119	
globalisation	.18	.695		rubbish	−.66	.104	
organisation	.18	.705		railway	−.69	.084	.
colour	.16	.739		shopping trolley	−.73	.060	.
anticlockwise	.14	.759		postman	−.74	.056	.
liberalisation	.14	.771		pushchair	−.76	.046	*
backwards	.05	.917		petrol station	−.78	.038	*
localisation	.01	.980		dustbin	−.79	.035	*
touch wood	−.04	.939		driving licence	−.80	.030	*
glocalisation	−.11	.809		nappies	−.80	.031	*
rubber	−.11	.820		a drop in the ocean	−.82	.023	*
torch	−.11	.807		parcel	−.84	.018	*
lift	−.12	.803		lorry	−.89	.007	**
realisation	−.15	.742		storm in a teacup	−.89	.007	**
potato crisps	−.22	.628		Anyway,…	−.89	.007	**
aluminium	−.29	.529		biscuit	−.92	.003	***
chemist's	−.31	.495		to licence	−.93	.003	***

Appendix C. Americanization vs. Britishization in Gibraltar English: apparent-time trends for all items (ordered by strength and polarity of correlation; seven age groups)
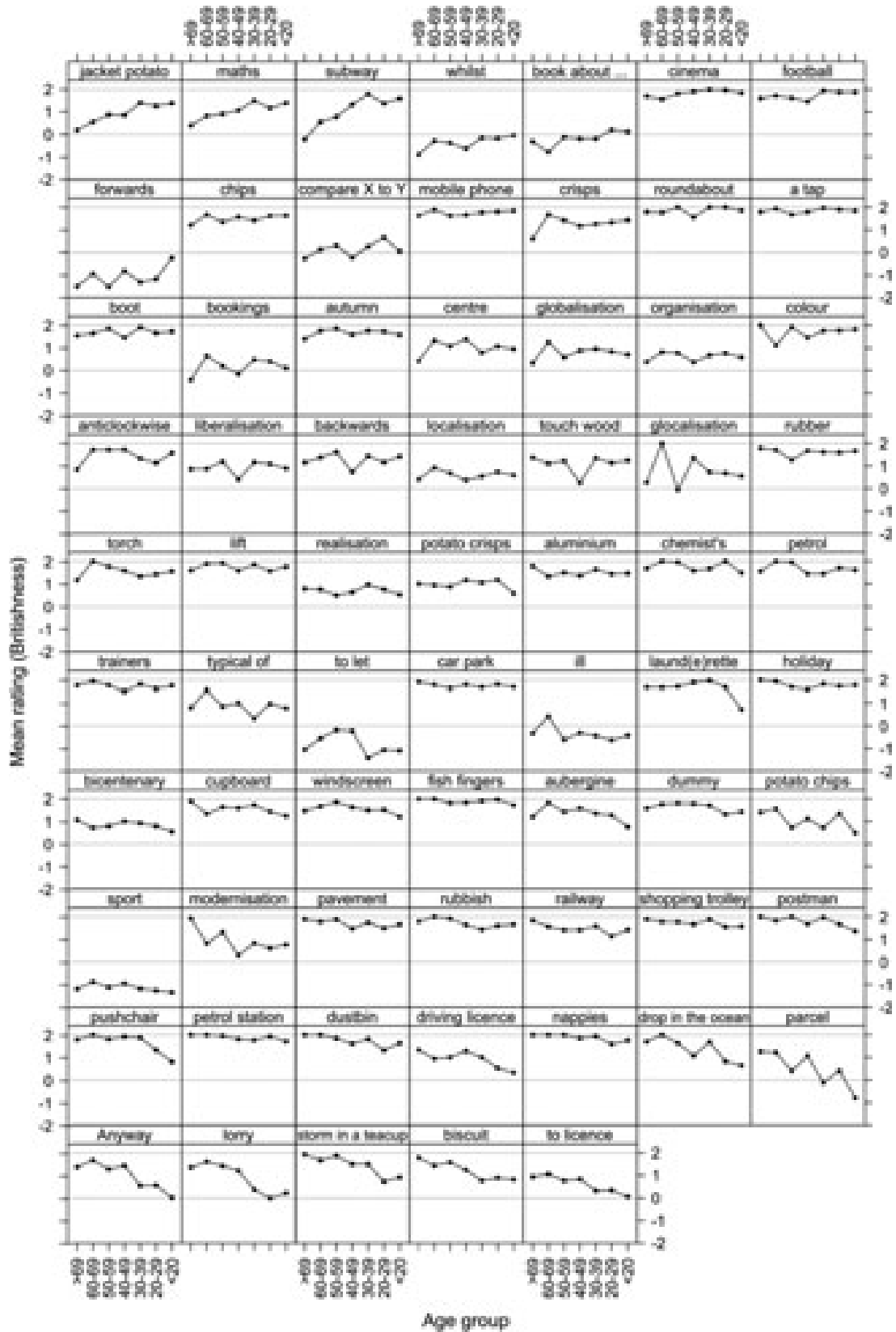
Notes

1 We dedicate this publication to Teresa Fanego, a linguist admirable for her ideas, perseverance, constructive communications, as well as the way in which she established and maintained academic networks and personal friendships. Teresa created an enduring link with the University of Bamberg through a project on variation, linguistic change and grammaticalization, which provided ample opportunities and invaluable support especially for junior researchers. This collaboration, formally as well as informally, became a constant source of inspiration that sparked reciprocal invitations and exchanges of ideas at all levels. We are grateful to two anonymous reviewers for their valuable comments. The usual disclaimers apply.

2 The most recent census data for the entire Gibraltarian population provide a value of *c.* 93 per cent English fluency based on self-reports (HM Government of Gibraltar 2012: xxxviii).

3 This contrasts strongly with figures for this item found in other L2 varieties like Maltese or Puerto Rican English, where English (or a combination of English and Maltese/Spanish) is submitted by only 5 per cent (Malta) and 13.5 per cent (Puerto Rico) of the raters. Note that, while not directly comparable, the findings of our questionnaire study are in line with Levey (2008: 59), but stand in stark contrast to Neidig (2008: 49–51), who claims that Spanish is the conversational language for two thirds of the Gibraltarian population and that near-exclusive use of Spanish characterizes parent-child interaction. However, even Neidig (2008: 55) acknowledges a growing influence of English in the latter domain (see also Weston 2012: 21).

4 For *parcel* vs. *package*, we considered as an alternative paraphrase 'something you send by *surface* mail' in order to avoid confusion with emails and computer programs. However, association tests with IT experts and students in 2008 and 2017 showed that both associate surface mail contexts for this test item rather than IT contexts. The longer paraphrase was therefore dismissed as minimal disambiguation gains were outweighed by extra efforts for explaining the term *surface mail* to older informants.

5 Of course, the expressions *binaries* and *British* vs. *American English usage* are simplifications which are used here for expository clarity. Some items have more than two alternatives, e.g., *dummy – pacifier – soother*, or *compare X with/to/and Y*. Similarly, we simplistically use *BrE* (or *AmE*, as the case may be) when we refer to more British (e.g., *backwards* vs. *backward*), exclusively (e.g., *-isation* spellings) or traditionally British terms (e.g., *lorry* vs. *truck*).

6 Alternative labels found in the literature are 'similarity matrix' or 'dissimilarity matrix'.

7 Thanks are due to Fabian Vetter for scripting. *MakeNex* is available upon request.

8 Originally, representations of this type were used in evolutionary biology and related areas to chart relationships between relevant categories, such as species or bacteria.

9 If the term were not laden with creolist associations, it would be tempting, in this investigation of lexical items, to speak of a 'lexifier' variety. See Krug, Schützler and Werner (2016: 45) for a similar interpretation based on a study of Maltese and Puerto Rican Englishes and their alignment with their BrE and AmE colonial 'parent' varieties. Note, however, that the alignment for those groupings is less close than the one observed between GibE and BrE here. We assume that this similarity is due to sociolinguistic factors applying to GibE, but also to language-external, historical factors (see further below and Weston 2011, 2015).

10 As we will see below, we are sometimes only dealing with alleged former norms. Often statistical preferences actually involve complex social patterns and sometimes they are even the inverse of what textbooks suggest.

11 The items *nappies*, *chips*, *backwards* and *bookings* receive only marginally higher ratings in GibE.

12 We plan to collect more British and American data, which in turn may lead to a reconsideration of individual items as well as to methodological adaptations.

13 The precise results of the linear regression model for male raters are INTERCEPT = .845 and AGE = .006 (p = .000); for female raters they are INTERCEPT = 1.055 and AGE = .003 (p = .019). That is, for example, a decrease of one year in the age of a male rater is expected to result in a rating that is lower by the value of .006.

14 For an alternative approach using Kendall's tau, see Hilpert and Gries (2009).

15 This does not mean that we take every rater's response at face value. In the case of our notorious UK-educated forty-nine-year-old lawyer with a PhD (see Section 4 for detailed discussion), whose parents were both teachers, for instance, we tend to believe that situational context impacts also on his usage of *while* vs. *whilst*, even though he reports to use only the latter. We do believe, however, that this rater indeed overwhelmingly uses traditionally BrE variants and that, over and above, our questionnaire results produce valuable data for charting variation and change, as the vast majority of raters do not report to use exclusively one variant in each case.

16 Apart from the items mentioned in this paragraph, only *forwards* and *whilst* show solidly negative ratings overall (see Figure 8.5). Both are special: the former has a nominal homograph *forward* used in sport(s); and *whilst* is a relatively formal variant that is probably less common than *while* in all present-day regional varieties.

References

Algeo, John (2006), *British or American English: A Handbook of Word and Grammar Patterns*, Cambridge: Cambridge University Press.

Braithwaite, David W. and Robert L. Goldstone (2013), 'Benefits of Graphical and Symbolic Representations for Learning and Transfer of Statistical Concepts', in Markus Knauff, Michael Pauen, Natalie Sebanz and Ipke Wachsmuth (eds), *Proceedings of the Thirty-Fifth Annual Conference of the Cognitive Science Society*, 1928–33, Berlin: Cognitive Science Society.

Central Intelligence Agency (2016), *The World Factbook*. Available online: https://www.cia.gov/library/publications/download/download–2016/index.html

Chambers, Jack K. (2009), *Sociolinguistic Theory: Linguistic Variation and its Social Significance*, Malden: Blackwell.

Dollinger, Stefan (2015), *The Written Questionnaire in Social Dialectology: History, Theory, Practice*, Amsterdam: Benjamins.

Fuchs, Robert and Ulrike Gut (2016), 'Register Variation in Intensifier Usage across Asian Englishes', in Heike Pichler (ed), *Discourse-pragmatic Variation and Change: Insights from English*, 185–210, Cambridge: Cambridge University Press.

Hilpert, Martin and Stefan Th. Gries (2009), 'Assessing Frequency Changes in Multistage Diachronic Corpora: Applications for Historical Corpus Linguistics and the Study of Language Acquisition', *Literary and Linguistic Computing*, 24 (4): 385–401.

HM Government of Gibraltar (2012), *Census of Gibraltar 2012*, Gibraltar: HM Government of Gibraltar. Available online: https://www.gibraltar.gov.gi/new/sites/default/files/HMGoG_Documents/Full%20Census%20Report%202012%20FINAL.pdf

Huson, Daniel H. and David Bryant (2006), 'Application of Phylogenetic Networks in Evolutionary Studies', *Molecular Biology and Evolution*, 23 (2): 254–67.

Kellermann, Anja (2001), *A New New English: Language, Politics, and Identity in Gibraltar*, Heidelberg: Books-On-Demand.

Kramer, Johannes (1998), *English and Spanish in Gibraltar*, Hamburg: Buske.

Krug, Manfred and Anna Rosen (2012), 'Standards of English in Malta and the Channel Islands', in Raymond Hickey (ed), *Standards of English: Codified Varieties around the World*, 117–38, Cambridge: Cambridge University Press.

Krug, Manfred, Ole Schützler and Valentin Werner (2016), 'Patterns of Linguistic Globalization: Integrating Typological Profiles and Questionnaire Data', in Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja and Sarah Chevalier (eds), *New Approaches to English Linguistics: Building Bridges*, 35–66, Amsterdam: Benjamins.

Krug, Manfred and Katrin Sell (2013), 'Designing and Conducting Interviews and Questionnaires', in Manfred Krug and Julia Schlüter (eds), *Research Methods in Language Variation and Change*, 69–98, Cambridge: Cambridge University Press.

Krug, Manfred and Lukas Sönning (2017), 'Language Change in Maltese English: The Influence of Age and Parental Languages', in Patrizia Paggio and Albert Gatt (eds), *The Languages of Malta*, 247–70, Berlin: Language Science Press.

Labov, William (1990), 'The Intersection of Sex and Social Class in the Course of Linguistic Change', *Language Variation and Change*, 2: 205–54.

Levey, Stephen (2008), *Language Change and Variation in Gibraltar*, Amsterdam: Benjamins.

Levey, Stephen (2015), 'Gibraltar English', in Jeffrey P. Williams, Edgar W. Schneider, Peter Trudgill and Daniel Schreier (eds), *Further Studies in the Lesser-known Varieties of English*, 51–69, Cambridge: Cambridge University Press.

Loureiro-Porto, Lucía and Cristina Suárez-Gómez (2017), 'Language Contact in Gibraltar English: A Pilot Study with ICE-GBR', *Alicante Journal of English Studies*, 30: 93–119.

McMahon, April and Warren Maguire (2013), 'Computing Linguistic Distances Between Varieties', in Manfred Krug and Julia Schlüter (eds), *Research Methods in Language Variation and Change*, 421–32, Cambridge: Cambridge University Press.

Neidig, Andrea (2008), *English und Spanisch im Kontakt: Das Yanito in Gibraltar*, Giessen: Herrmann.

Nevalainen, Terttu (1996), 'Gender Difference', in Terttu Nevalainen and Helena Raumolin-Brunberg (eds), *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*, 57–76, Amsterdam: Rodopi.

R Core Team (2016), *R: A Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing. Available online: http://r-project.org

Raumolin-Brunberg, Helena and Arja Nurmi (1997), 'Dummies on the Move: Prop-ONE and Affirmative DO in the 17th Century', in Terttu Nevalainen and Leena Kahlas-Tarkka (eds), *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, 395–417, Helsinki: Société Néophilologique.

Ruette, Tom, Katharina Ehret and Benedikt Szmrecsanyi (2016), 'A Lectometric Analysis of Aggregated Lexical Variation in Written Standard English with Semantic Vector Space Models', *International Journal of Corpus Linguistics*, 21 (1): 48–79.

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

Schützler, Ole (2015), 'Transforming Acoustic Vowel Data: A Comparison of Methods, Using Multidimensional Scaling', in Adrian Leemann, Marie-José Kolly, Stephan Schmid and Volker Dellwo (eds), *Trends in Phonetics & Phonology: Studies from German-speaking Europe*, 35–47, Bern: Peter Lang.

Suárez-Gómez, Cristina (2012), 'English in Contact with other European Languages (Italian, Spanish, Slavic)', in Alexander Bergs and Laurel J. Brinton (eds), *Historical Linguistics of English: An International Handbook*, *Vol. II*, 1738–53, Berlin: De Gruyter Mouton.

Szmrecsanyi, Benedikt (2011), 'Corpus-based Dialectometry: A Methodological Sketch', *Corpora*, 6 (1): 45–76.

Szmrecsanyi, Benedikt and Christoph Wolk (2011), 'Holistic Corpus-based Dialectology', *Brazilian Journal of Applied Linguistics*, 11 (2): 561–92.

Tagliamonte, Sali (2012), *Variationist Sociolinguistics: Change, Observation, Interpretation*, Malden: Blackwell.

Vessey, Iris (1991), 'Cognitive Fit: A Theory-based Analysis of the Graphs versus Tables Literature', *Decision Sciences*, 22 (2): 219–40.

Wainer, Howard (1992), 'Understanding Graphs and Tables', *Educational Researcher*, 21 (1): 14–23.

Wälchli, Bernhard and Benedikt Szmrecsanyi (2014), 'Introduction: The Text-Feature-Aggregation Pipeline in Variation Studies', in Benedikt Szmrecsanyi and Bernhard Wälchli (eds), *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech*, 1–25, Berlin: De Gruyter Mouton.

Werner, Valentin (2014), *The Present Perfect in World Englishes: Charting Unity and Diversity*, Bamberg: Bamberg University Press.

Werner, Valentin (2016), 'Overlap and Divergence – Aspects of the Present Perfect in World Englishes', in Elena Seoane and Cristina Suárez-Gómez (eds), *World Englishes: New Theoretical and Methodological Considerations*, 113–42, Amsterdam: Benjamins.

Weston, Daniel (2011), 'Gibraltar's Position in the Dynamic Model of Postcolonial English', *English World-Wide*, 32 (3): 338–67.

Weston, Daniel (2012), 'Code-switching Variation in Gibraltar', *International Journal of Bilingualism*, 17 (1): 3–22.

Weston, Daniel (2015), 'The Lesser of Two Evils: Atypical Trajectories in English Dialect Evolution', *Journal of Sociolinguistics*, 19 (5): 671–87.