

John Benjamins Publishing Company



This is a contribution from *English World-Wide* 33:2
© 2012. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

The hypothetical subjunctive in South Asian Englishes

Local developments in the use of a global construction

Marianne Hundt, Sebastian Hoffmann and Joybrato Mukherjee
University of Zurich / University of Trier / University of Giessen

This paper studies the distribution and usage patterns in hypothetical *if*-clauses in a set of South Asian Englishes (SAEs), namely Indian, Pakistani, Bangladeshi and Sri Lankan English on the basis of web-derived newspaper data. Comparative evidence comes from newspaper texts in the British National Corpus (BNC). It looks at the use of subjunctive *were*, indicative *was* and modal *would* as variant verb forms in the *if*-clause. The qualitative analyses also consider tense sequencing in the main and subordinate clause. In terms of overall frequencies, the SAEs do not cluster together in their use of the subjunctive but form a gradient or cline with British English at one end. Similarities between the SAEs emerge from the qualitative analyses. An additional, serendipitous result of the study concerns the local use in SAEs of the subordinator *on if* meaning ‘whether’, a pattern that is likely to have its origin in Indian English.

Keywords: South Asian Englishes, web-based newspaper corpora, subjunctive, indicative, nativization, exonormative influence

1. Introduction

The vast majority of linguistic features in regional varieties of standard(izing) English around the world are shared with other varieties. It is received wisdom that divergence from the global norm, especially at the syntactic and morphological level, tends to be minimal. But globally available constructions are used slightly differently — with respect to overall frequency, stylistic preferences, divergent variants — in regional varieties of English. Such local usages and developments of global grammatical patterns are interesting because they provide us with insights

into norm development and the dynamics of different Englishes in a globalizing linguistic market (see Hundt 2009).

The focus in this paper is on variation between subjunctive *were* and indicative *was* in hypothetical *if*-clauses across South Asian Englishes (SAEs). Variability in this particular area of English grammar has been studied for varieties of English as a first language (L1-varieties; see Johansson and Norheim 1988; Peters 1998; Hundt 2006; Leech *et al.* 2009); as for second-language (L2) varieties, Indian English (IndE) has been studied by Sayder (1989) and Schneider (2000, 2005) on the basis of the Kolhapur Corpus and ICE-India.¹ With respect to the Englishes that we will focus on, i.e. English in India, Pakistan, Bangladesh and Sri Lanka, standard corpora are only available for IndE. The evidence for our study therefore comes from large web-derived corpora of English-medium daily newspapers in India, Pakistan, Bangladesh and Sri Lanka.

In the first part of our paper (Sec. 2), we will briefly review the results of previous studies from which we will derive our hypotheses. In Section 3, we will define the variable, describe the composition of our web-derived corpora and comment on the retrieval process. We will then discuss the results against the notion of ongoing 'glocalization' (Sec. 4). The analysis will take both quantitative and qualitative differences into account. In Section 5, we turn to a serendipitous finding of a local SAE innovation.

2. Previous research

The first corpus-based approach to *were*-subjunctives in British English (BrE) and American English (AmE) is the article by Johansson and Norheim (1988) who found that, in the LOB and Brown corpora, the subjunctive was preferred to the indicative in conditionals. This applies even more to AmE than to BrE. Diachronic data (Peters 1998; Hundt 2006; Leech *et al.* 2009) show that the *were*-subjunctive is decreasing in hypothetical adverbial clauses in the 20th century. One of the reasons that have been given for this ongoing change is the redundancy of the subjunctive: *if*, for instance, already expresses counterfactuality.

Among varieties of English as a first language (L1-varieties), AmE turns out to be the conservative variety, whereas Australian (AusE) and New Zealand English are even more advanced in losing the *were*-subjunctive than BrE. This unusual pattern of American conservatism most likely has to be ascribed to a stronger impact of

1. Note that Schneider (2000) criticizes Sayder's analytical accuracy. In the following, we have therefore focussed on Schneider's rather than Sayder's results for our hypotheses. Schneider (2005) has been slightly updated and republished as Schneider (2011).

prescriptivism in the US in written published texts. Prescriptive influence may also account for the use of hypercorrect subjunctives in non-counterfactual contexts, such as *I wonder if it were possible* or in the following example from the Frown Corpus:

- (1) ... he simply didn't care if she *were* a snake charmer from India or the whore of the island, which he supposed she was. (Frown, P09 142–3)

Anecdotal evidence points at another possible change, namely the spread of *would* to the subordinate clause of conditional sentences. This option is also attested in earlier stages of the language, especially in non-standard usage (see Denison 1998: 300; Hundt *fc.* for evidence from Late Modern English). But we also find it in the speech of middle class characters, for instance Molly Gibson in *Wives and Daughters*:

- (2) 'Papa doesn't care what he has, if it's only ready. He would take bread-and-cheese, if cook *would* only *send* it in instead of dinner'. (E. Gaskell. 1866. *Wives and Daughters*. Penguin Edition 1996, p. 127)

However, there is no corpus evidence for BrE and AmE of a growing use of *would* as a variant for subjunctive *were* in counterfactual *if*-clauses — a usage that still seems to be more or less non-standard; there is even only a sprinkling of examples in spoken AmE (Leech *et al.* 2009: 63, 66).

The results of previous corpus-based studies (Schneider 2000, 2005, 2011) suggest that second-language (L2) varieties (along with more recent post-colonial L1-varieties like AusE) use fewer subjunctives than BrE and AmE. As possible reasons for this, Schneider discusses the complexity of the subjunctive, its semantic redundancy, and the dynamics of the contact situation (dialect levelling in the case of AusE and L2-acquisition phenomena in the case of IndE). He hypothesizes that “[c]olonial varieties tend to reduce grammatical complexity if it is not functionally required, i.e. meaningful” (Schneider 2000: 130). On the other hand, research into New Englishes has shown that they show a tendency for more transparent or explicit marking of grammatical meaning (see Hartford 1989).

In the light of previous research into the subjunctive on the one hand and New Englishes on the other we add the following hypotheses with regard to the use of the subjunctive in SAEs:

- a. SAEs in general and IndE in particular display a potential to develop their own norms so that differences between BrE and SAEs are to be expected.
- b. In particular, we expect the non-standard use of *would* in counterfactual *if*-clauses — as a variant of subjunctive *were* — to be more frequent, in general, in SAEs than in BrE.

- c. Other deviant² patterns may be typical of the nativization process in the standard(izing) L2-varieties.
- d. We assume that IndE as the largest and most institutionalized variety serves as a lead variety for the other varieties, so that by and large all SAEs cluster together with IndE.

3. Definition of the variable and retrieval of data

The focus in our paper is on the past subjunctive *were*. More precisely, we are concerned with the variation between subjunctive and indicative in the subordinate clause of a conditional sentence. The examples in (3) illustrate the subjunctive pattern³, those in (4) the indicative; the examples in (5) illustrate the variant with a modal periphrasis.

- (3) a. There is nobody to beat the politician in doling out sheer twaddle as if it *were* the profoundest truths of life! (*The Statesman*, 24/05/2004)
- b. “If I *were* you I wouldn’t go to the press with that statement.” (*The Statesman*, 06/02/2003)
- (4) a. Among other nations of like mind many Indians would fear mass unemployment if a man *was* asked to do more than one job? (*The Statesman*, 13/07/2003)
- b. Our cricketers endorse products as if advertising *was* going out of fashion tomorrow. (*The Statesman*, 22/2/2003)
- (5) a. ... a top UN official would not hesitate ... to go further if that *would* achieve the desired result? (*The Statesman*, 02/06/2003)
- b. “I will even give up my chief ministership if it *would* save the territory of Manipur from being divided”, Mr Ibobi Singh said. (*The Statesman*, 12/01/2003)

2. The term ‘deviant’ is not used with a negative evaluative connotation in this context. A possible synonym would be ‘divergent’. For a discussion of the notion of ‘deviation’ or ‘divergence’ in the context of New Englishes and learner Englishes, see Mukherjee and Hundt (2011).

3. In theory, the *were*-subjunctive could also be used in conditional sentences with an inverted protasis (*Were she a little more friendly, I would certainly call her more often*), but this is an extremely formal option. A search for *I, he, she, it, this* and *that* preceded by directly adjacent *were(n’t)* in four corpora of Present-Day English only yielded a total of 15 *were*-subjunctives of this type (seven American, six British and two Australian). Emphasis in the examples has been added throughout.

The data used for our study comes from the electronic archives of four South Asian newspapers. Thousands of HTML documents were automatically downloaded and post-processed with Perl scripts to isolate the actual news items from its surrounding boilerplate code and other irrelevant text elements (e.g. advertisements, links). For BrE, which — historically — represents the input variety for South Asia, we made use of the newspaper texts contained in the British National Corpus (BNC). Table 1 presents an overview of the data used for the present study.

Table 1. Data used for the present study

Country	Newspaper	N words
Bangladesh	Daily Star	33,167,503
India	The Statesman	31,064,290
Pakistan	Daily Times	55,961,050
Sri Lanka	Daily News	26,138,250
Britain	various sources (BNC)	9,412,174

The data were searched for the relevant subordinating conjunctions *as if*, *as though*, *even if* and *if* followed either by *were(n't)* or *was(n't)*, allowing for a space of up to ten tokens (including punctuation marks) to the right of the conjunction.⁴ To verify the hypothesis on the use of *would* in conditional clauses, we also retrieved instances of the same set of conjunctions that were followed by *would*, again with a maximum of ten intervening tokens.

We defined a number of conditions to automatically discard irrelevant constructions. For this purpose, the data was first part-of-speech tagged with CLAWS (see Garside 1987), as this allowed us to discard a number of irrelevant constructions on the basis of abstract tag-based patterns. For example, the Perl script automatically removed instances where an existential *there* was followed by *was / were / would* and a plural noun phrase (defined by an optional quantifier — e.g. *more*, *any* —, optional determiners and adjectives and a plural noun). Instances automatically excluded in this way are *if there were more such losses*, *if there were friendly governments* and *if there were universities*. The resulting concordances of the searches were then randomly reduced to 500 instances each and manually post-edited. We discarded instances of *if + was* in contexts of back-shifting as these would have been ambiguous between an ordinary indicative and one that was used instead of a subjunctive. Clauses with collective nouns or otherwise

4. The subordinating conjunction *like* as a colloquial variant of *as if* was not included for two reasons: previous studies as those by Johansson and Norheim (1988) or Peters (1998) had not included it, and it was not attested in the Frown Corpus, the most likely standard-corpus source of a possible occurrence, containing one million words of written American English from 1992. It is thus an unlikely variant in the type of written texts that we analyzed.

ambiguous subjects were not included in the data set, either. Similarly, instances where *if* was used in the sense of ‘whether’ were not included, among them instances of what Ryan (1961) called pseudo-subjunctives:

- (6) a. The one safeguard squeezed from a reluctant home secretary is that a judge may ask if he, that is his security advisers, *were* “reasonable” in their suspicions. (*The Statesman*, 24/02/2005)
- b. He asked if it *were* possible for the academy to “please defer the honour until I am 80?” (*The Statesman*, 16/02/2003)

Finally, as Peters (1998: 88) points out, “just how hypothetical a given condition is in the minds of those communicating may be impossible to ascertain”. One problem for a study of hypothetical *if*-clauses is that the main clause does not necessarily have to contain a modal verb to make a subjunctive grammatical, as the following examples from *The Statesman* show:

- (7) a. Mr Bush turned every electoral fight into a national issue, as if he *were* fighting the presidential election again. (*The Statesman*, 13/11/2002)
- b. Saudi Arabia has to be treated differently from Saddam Hussain’s Iraq or Kim Jong II’s North Korea, even if it *were* in the same league, as most Americans believe. (*The Statesman*, 13/08/2003)

It is the failure to use *would* in the main clause in these sentences that poses a problem and not the use of the subjunctive. We therefore decided to include them as instances of *were*-subjunctives in conditional adverbial clauses.

4. Results and discussion

4.1 Overall frequencies of subjunctive contexts

Between 41% (India) and 48% (BNC) of all retrieved instances proved to be relevant for our investigation. Figure 1 shows the overall frequency of all three hypothetical constructions (i.e. the conflated frequencies of *was*, *were* and *would*) in the five varieties. (Raw frequencies are given in Table 2 below.)

With 70 instances per million words (pmw), the British data contains more than twice as many hypothetical *if*-clauses than the data from Bangladesh and Sri Lanka (31 pmw and 28 pmw, respectively). For the remaining analysis, we will concentrate on the proportions of our three variant constructions within each variety of English and ‘deviant’ patterns typically found in L2-varieties.

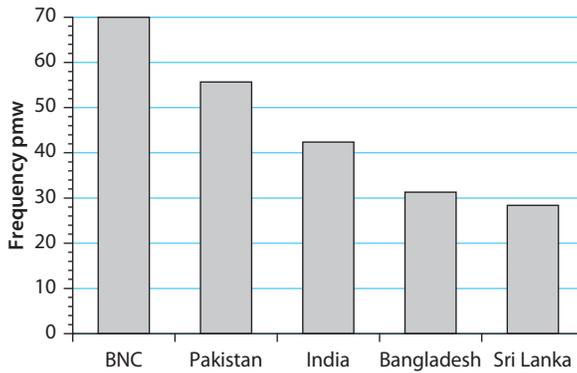


Figure 1. Overall frequencies of investigated hypothetical constructions in the five varieties; extrapolated from random subsets of 500 instances

4.2 Variability in subjunctive contexts

As Figure 2 and the corresponding Table 2 show, considerable differences were found between the five investigated varieties.

Table 2. The use of *was*, *were* and *would* in the *if*-clauses of hypothetical sentences, raw frequencies and percentages

	Pakistan		Sri Lanka		India		Bangladesh		GB	
	N	%	n	%	n	%	N	%	N	%
<i>was</i>	38	19.3	53	31.0	70	38.7	75	41.0	118	49.8
<i>were</i>	153	77.7	101	59.1	100	55.2	97	53.0	115	48.5
<i>would</i>	6	3.0	17	9.9	11	6.1	11	6.0	4	1.7
total	197	100	171	100	181	100	183	100	237	100

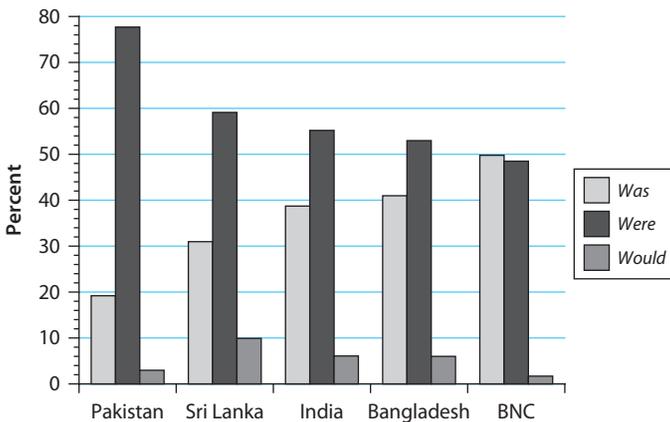


Figure 2. Proportions of *was*, *were* and *would* in the *if*-clauses of hypothetical sentences

In the BrE newspaper data, the subjunctive pattern is about as frequent as its indicative variant. In contrast, Pakistani English clearly prefers the *were*-subjunctive, which accounts for 78% of all cases. IndE covers the middle ground between those two extremes: *were*-subjunctives account for 55% of all hypothetical sentences and the indicative pattern for 39%. If we compare the frequencies of the indicative and the subjunctive only, we find that usage patterns in our newspaper data do not divide the varieties into a neat Asian-Englishes vs. L1-variety dichotomy: quantitative differences between usage patterns in newspapers in Britain, Bangladesh and India are below the level of statistical significance (see Table 3); journalists in Sri Lanka use the subjunctive significantly more often than British journalists, and the marked preference for the subjunctive in our Pakistani data sets the variety off from all others studied here. The statistics support the visual impression in Figure 2, namely that in quantitative terms, we are dealing with a gradient rather than two clusters. This also means that there is no evidence to support the hypothesis that IndE serves as a lead variety for the other SAEs.

Table 3. Chi-square significance values for *was* vs. *were* in hypothetical conditional clauses⁵

	Pakistan	Sri Lanka	India	Bangladesh	GB
Pakistan		0.00234	0.00001045	0.00000111	0
Sri Lanka	0.00234		NS (0.21)	NS (0.09)	0.00165189
India	0.00001045	NS (0.21)		NS (0.65)	NS (0.06)
Bangladesh	0.00000111	NS (0.09)	NS (0.65)		NS (0.16)
GB	0	0.00165189	NS (0.06)	NS (0.16)	

Similarly, our hypothesis on the frequency of *would* as a variant in the *if*-clause is not borne out by the data, either: it is not generally more frequent in our SAEs than in the BNC data. The difference between the BNC data and the Pakistani newspaper did not prove significant. *Would* is a low frequency variant in hypothetical *if*-clauses in all data sets. It is used most frequently in the Sri Lankan newspaper (significantly more often than in the BNC data and the Pakistani paper). With the exception of the latter paper, all SAEs do use *would* significantly more often in hypothetical *if*-clauses than BrE (see Table 4).

Furthermore, a closer look at the BNC data shows that the four occurrences of *would* in the *if*-clause are not instances of prototypical counterfactual uses:

5. A zero-value refers to an extremely low p-value and therefore very high statistical significance.

Table 4. Chi square significance values: *was* + *were* together vs. *would*

	Pakistan	Sri Lanka	India	Bangladesh	GB
Pakistan		0.0064181	NS (0.16)	NS (0.16)	NS (0.35)
Sri Lanka	0.0064181		NS (0.18)	NS (0.17)	0.00019695
India	NS (0.16)	NS (0.18)		NS (0.97)	0.01682056
Bangladesh	NS (0.16)	NS (0.17)	NS (0.97)		0.01791952
GB	NS (0.35)	0.00019695	0.01682056	0.01791952	

- (8) a. Mandy feels *as though she would* be prostituting herself to sell her story, she said. (BNC: CBF: 10903)
- b. Beardsley, meanwhile, produced a performance of such craft, guile and wisdom that it could yet make him the saviour of England manager Graham Taylor — *if only he would pick* him. (BNC: CH3: 2731)
- c. *If only he would drop* all this ridiculous pretence, stop acting the fool and raise his game one more time. (BNC: CEP: 4678)
- d. Videotaped interviews should only be excluded *if admitting them would be* contrary to the interests of justice. (BNC: AA5: 561)

In (8a), *would* is used after the complex subordinating conjunctions *as though* and *if only* rather than simple *if*; in (8b) and (8c) it is part of a complex verb phrase and *would* thus replaces *were to* rather than simple *were*; in (8d), finally, a de-verbal noun immediately follows the subordinator which might camouflage the combination of *if* + *would*; one could further argue that *would* in (8d) is not really counterfactual. In other words, while there is no statistical, quantitative evidence that all SAEs investigated here use *would* more often in *if*-clauses, there is evidence of a qualitative difference. We find *would* used in the *if*-clause as a true variant of *were* or *was* in our SAEs:

- (9) a. The Belgian law against war crimes is a step in this direction, and I hope that many other countries will follow suit. Of course, it would be better *if the International Criminal Court in The Hague would fulfil* this duty, but much time will pass before it will be able to. (*Daily Times*, 3/2/2003)
- b. *Even if there would be a general election* the results will not bring a major change in the parliamentary composition. In the national interest, a general election should be the last resort. (*Daily News*, 20/01/2004)
- c. As we parted without having the slightest hint that it was to be for the last time, he leaned forward and murmured, “we might die, but I won’t regret *if our people would be happy and content*”. (*Daily Star*, 16/12/2003)

On the whole, however, we do not find a more or less uniform distributional pattern in all SAEs. In fact, some SAEs are much more markedly different from each other than from BrE, as for example the data for Pakistan and Bangladesh

in Figure 2 show. This complex picture is similar to what we found for light-verb constructions across SAEs (see Hoffmann, Hundt and Mukherjee 2011) — and presumably we will find more or less marked differences between SAEs with regard to other lexicogrammatical patternings as well.

Our data initially seem to contradict Schneider's findings, which were based on an analysis of a range of databases including the Kolhapur Corpus (with data from the 1970s) and LOB (with data from 1961), which share the same size and design. Schneider (2000: 130) found a relative frequency of 49% *were*-subjunctives compared with about 65% *were*-subjunctives in BrE. In other words, for data from the 1960s / 1970s, IndE had a lower relative frequency of past subjunctives than BrE. Diachronically, the *were*-subjunctive has been losing ground in BrE, where it is now at about 52% (see Hundt 2006 and Leech *et al.* 2009). Even though they are not directly comparable with the Kolhapur Corpus, our *Statesman* data suggest that the subjunctive might have increased slightly in IndE. Interestingly, AmE has a relatively stable proportion of about 74% *were*-subjunctives in hypothetical contexts (see Leech *et al.* 2009) — and thus a similar proportion of past subjunctives as the one found in our Pakistani data. In other words, with a suspected growth of subjunctives in IndE and a high incidence of subjunctives in Pakistani English, an additional external factor might be the global influence of AmE on local usage patterns in SAEs.

4.3 Deviance from the L1-model

In the process of post-editing the concordances we found that the SAEs often diverged from the BrE norm, for instance in the sequencing of tenses in the main and subordinate clause. Typical examples are shown in (10):

- (10) a. I had to laugh, but it might not have been funny if I *was* [had been] still based in India. (*The Statesman*, 02/08/2004)
- b. Caste-based reservation wouldn't have been needed if the caste system *was* [had been] wiped out. (*The Statesman*, 18/10/2004)
- c. If Shakespeare *were* [had been] born in India, he would have written: "All the world's a urinal ..." (*The Statesman*, 16/07/2004)
- d. But the director of the George Telegraph Group mightn't have sounded half as credible if he *weren't* [hadn't been] ensconced in his southern Kolkata corporate enclave surrounded by the sort of electronic gadgetry that marks the blue-chip executive out. (*The Statesman*, 19/02/2004)

These deviations from the BrE norm add support to Schneider's (2000: 130) hypothesis that post-colonial L2-varieties might prefer less complex patterns — both indicative *was* and subjunctive *were* are less complex than the past perfect usually

found in BrE in these contexts.⁶ Similarly, Agnihotri, Khanna and Mukherjee (1988:94) use L2-acquisition strategies to account for the underuse of past perfect constructions in their student essay data. However, this explanation is probably somewhat simplistic. Sharma (2001:355), for instance, found an overall higher number of past perfect constructions in the press sections of the Kolhapur than in those of the Brown and LOB corpora. Furthermore, Sharma found that the functions of the past perfect in IndE were markedly different from those in the American and British data: the past perfect was used with preterite and present perfect meanings and the pragmatic function of a general marking of remoteness and completion in narrative contexts. Comrie (1985:69) advances a substrate-based explanation for the apparent overuse of the past perfect in IndE. Sedlatschek (2009:260–1), finally, observes diachronic change in the use of past perfect in journalese. In addition, the use of past perfect VPs appears to be highly register- and medium-dependent in his corpus of IndE (260–1, 264). Schneider and Hundt (2012) compared the use of the past perfect across different L1- and L2-Englishes and found that, due to high dispersion, it is quite possible to obtain datasets that suggest opposite results, especially on the basis of small corpora. It is therefore necessary to combine purely quantitative with more qualitative approaches.

The SAEs we investigated also occasionally use the subjunctive in non-conditional contexts, as illustrated in (11):

- (11) a. Prepaying loans and stopping future loans will seriously harm India's social sector programmes if and only if money from these sources *were* [was] efficiently used to create viable public goods. (*The Statesman*, 29/06/2003)
- b. He asked if it *were* [was] possible for the academy to “please defer the honour until I am 80?” (*The Statesman*, 16/02/2003)
- c. “Why did Zia legalise Mushataq's ordinance if there *were* [was] no relation between them ?” (*Daily Star*, 08/2005)

This rare ‘overuse’ of the subjunctive in non-conditional contexts might indicate a (recessive) exonormative pressure that occasionally results in hypercorrect patterns. However, deviation in tense sequencing — an ‘innovative’ pattern — is much more common in our data.

First of all, it is interesting to note that slightly ‘deviant’ patterns are also attested in our BNC data, as illustrated in (12):

6. Note that the indicative occurs more often in these past perfect contexts than the subjunctive (for instance with a ratio of 20:3 in our IndE data).

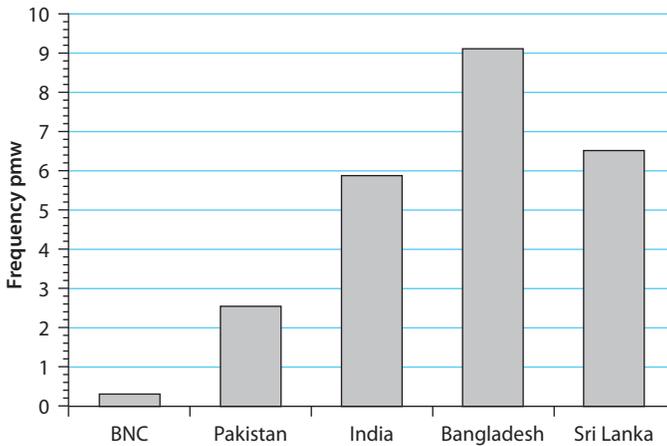


Figure 3. Deviant patterns in SAEs and the BNC

- (12) a. *If it were possible* for the Football League to kick him out of the club, that's what I want. (BNC: CEP: 8868)
- b. He said: *If I were to single out* any one project as a model of its kind, it is the car park in Chichester. (BNC: CH2: 11058)

It is not the use of the subjunctive itself that is somewhat unusual, but the combination with a present tense verb instead of a modal in the main clause. Deviant patterns of the type illustrated in examples (10) and (11) are only attested in our SAEs. Moreover, we find the same grouping of varieties as in the overall distribution of conditional contexts: the deviation from the L1-norm is less common in our Pakistani data. Our Indian and Sri Lankan newspapers yield slightly more deviant patterns pmw. The *Daily Star* (Bangladesh) is the paper with the highest relative frequency of patterns typical of L2-varieties of English.

4.4 Discussion

The question arises how best to account for the patterns of variation found in the SAEs investigated, both vis-à-vis BrE and amongst each other. Firstly, as far as similarities between the SAEs are concerned, one factor might be the influence that IndE is likely to have as a regional lead variety on smaller, less institutionalized neighbouring varieties. Alternatively, exonormative influence from AmE may play a role. Secondly, the observed heterogeneity of SAEs may be due to different degrees of institutionalization of the SAEs studied here. Finally — and linked to different degrees of institutionalization —, the historical input variety, i.e. BrE, may exert different degrees of exonormative pressure on individual SAEs. We

would like to sketch out by referring to Figure 2 to what extent the three factors can be linked to the trends and findings for the subjunctive.

— Factor 1: regional lead variety or exonormative influence from AmE?

What Figure 2 shows is that the SAEs yield (1) a higher proportion of *were*-subjunctives and (2) a higher proportion of counterfactual *if*-clauses with *would*. In these two regards, the SAEs are similar, even though they form a gradient towards BrE rather than a neat group that sets them off from their historical input variety. Two different factors may account for these trends: either the exonormative influence of the largest and most stable variety in the region, namely IndE, or exonormative influence from the largest and most dominant variety worldwide, namely AmE. AmE influence is a more likely influential factor that may account for a higher use of *were*-subjunctives in the SAEs than for the higher proportion of counterfactual *if*-clauses with *would*. Evidence from the Brown family of corpora indicates that *were*-subjunctives are used significantly more frequently in the 1990s AmE texts than in the parallel BrE texts (Leech *et al.* 2009: 65). Even though prescriptivists frequently comment on the (growing) use of *would* in counterfactual *if*-clauses in the US, this could not be corroborated by corpus evidence (Leech *et al.* 2009: 66). That AmE as such might exert some influence on SAEs, especially if we base our investigation on newspaper data, is corroborated by Sedlatschek: he observed that IndE journalese was converging with AmE newspaper usage in the area of past perfect constructs (Sedlatschek 2009: 262–3). Further evidence is needed to solve this puzzle, e.g. questionnaire data or information from house-internal style guides that may have influenced the journalists writing for the newspapers our study is based on.

— Factor 2: degree of institutionalization

As we have already argued with regard to the use of light-verb constructions (see Hoffmann, Hundt and Mukherjee 2011), the degree of institutionalization — in the sense of status, stability and acceptance of the local variety of English — is a factor which may explain some of the intervarectal differences between SAEs. While in L1-varieties of English there is a development towards the indicative, in L2-varieties the subjunctive seems to be retained or even on the increase. One could argue, for example, that the higher proportions of *were*-subjunctives and of *if*-clauses with *would* in the Pakistani and Sri Lankan data are exponents of a trend towards a more explicit and overt lexical marking of counterfactuality in varieties of English which are less institutionalized than IndE. It has been shown elsewhere that other deviations or innovations in L2-varieties of English may be caused by a need for a more explicit and transparent marking of linguistic functions. Consider, for example, the prepositional verbs *discuss about sth.* and *visit to sth.* with the additional — and semantically redundant — prepositions, which are attested in many

L2-varieties worldwide. Our data only partly support Schneider's (2000: 130) claim that New Englishes in general reduce grammatical complexity if it is not needed: our SAEs do not use the more complex past perfect but prefer either the subjunctive or indicative in counterfactual *if*-clauses after perfective VPs in the main clause.

— Factor 3: exonormative pressure from BrE

Linked to a very low degree of institutionalization is another factor, namely exonormative pressure from the historical input variety. In Bangladesh, a largely monolingual country, English has no official status, is not a medium of instruction and does not fulfil the function of a link language as in the other multilingual South Asian countries — English in Bangladesh is, in essence, a foreign language. The fact that Bangladesh is closest to BrE in the distribution of hypothetical variants is in line with a higher degree of exonormative pressure from native English on the unstable variety of English in Bangladesh. Note, for example, that while for the other SAEs, the frequency of *were*-subjunctives decreases with a higher degree of institutionalization (Pakistan à Sri Lanka à India), Bangladesh opts out from this cline because it is oriented towards the exonormative British model.

Given that English in South Asia displays a certain degree of internal variation, it is unrealistic to attribute the observed patterns of variation to any single factor alone, e.g. influence of IndE on the neighbouring varieties. In other words, quantitative and qualitative corpus evidence shows that the local use of global constructions produces a picture of convergence and divergence that defies a monocausal explanation: "... there is no monolithic explanation for the nature of variability at large" (Sedlatschek 2009: 274).

5. Excursus: the corpus linguist's "lucky dip"

The close analysis of large web-derived corpora is not only useful for the description of quantitative differences between varieties, but also for the identification of low-frequency innovations in lexicogrammar. In this section, we will briefly mention one such innovation.

While manually post-editing the concordances, one of the authors came across an example that had to be eliminated from the list of relevant hits but nevertheless proved particularly interesting for a discussion of glocalization in SAEs. In the *Statesman*, the subordinator *if* is used directly after the preposition *on* in contexts where BrE requires the use of *whether*:

- (13) *On if* the man, indeed, was the DM's security guard, he said it can be determined only after a probe. (*The Statesman*, 16/01/2002)

A subsequent search in our databases showed that this pattern is used quite regularly in the *Statesman*:

- (14) a. *On if* he was interested to play in Kolkata or in JCT, the ace striker said: “...” (*The Statesman*, 29/04/2003)
- b. *On if* he had received any instructions from the Indian camp to prepare pitches facilitating the Indian bowlers, he said: “...” (*The Statesman*, 27/09/2003)
- c. He was replying to a reporter’s question *on if* India would raise the issue at the WTO, on the sidelines of a seminar here. (*The Statesman*, 04/04/2003)
- d. Meanwhile, to a query *on if* he was interested in a second stint in county cricket, after turning out for Lancashire in 2000, the Indian skipper said he wished to turn out for Kent next season. (*The Statesman*, 18/07/2003)
- e. He, however, was evasive *on if* KMC had given him permission to use the space. (*The Statesman*, 01/03/2004)
- f. “Obviously there are two opinions *on if* the players are overworked”, he added. (*The Statesman*, 12/03/2002)
- g. The Prime Minister, his deputy and the NDA convener are expected to meet to decide *on if* there is a need to review Mr Nitish Kumar’s decision on ER bifurcation. (*The Statesman*, 14/07/2002)

The *Statesman* archive yields a total of 23 relevant instances. This indicates that *on + if* is a nativized syntactic pattern in IndE. A search in our Bangladesh newspaper produced seven hits (see examples in [15]), whereas only one example each could be retrieved from the Sri Lankan and Pakistani newspaper archives (examples in [16] and [17], respectively).

- (15) a. Each sounds like the other, depending *on if* her party is in power or in the opposition. (*Daily Star*, 10/2003)
- b. Dwelling *on if* sculptor was being valued as it should be today in Bangladesh, Hamiduzzaman said, “...” (*Daily Star*, 11/2003)
- c. *On if* a bilateral free trade agreement (FTA) was off the table because of difficulties it has run into, Morshed shot back saying: “...” (*Daily Star*, 06/2004)
- d. Some alleged they were being spied *on if* they were casting their votes in favour of the alliance nominee. (*Daily Star*, 07/2004)
- e. The court also issued a warrant of arrest against Saifuddin and formed a three-member commission to assess the 84-year-old’s health condition, asking it to submit a report by yesterday *on if* he is fit for deposition. (*Daily Star*, 10/2004)

- f. Police sources also said the investigators are focusing *on if* any feuds over the formation of Kafrul thana unit committee of the AL led to the killing. (*Daily Star*, 05/2005)
 - g. “We’ve no evidence to comment *on if* it has international links”, he said, but hastened to add that such serious incident will not go unnoticed by “international terrorists”. (*Daily Star*, 08/2005)
- (16) “That’s going to be the key, depending *on if* I do play that game on the 17th, and maybe finish off the one-day series”, he said (*Daily News*, 08/2001)
- (17) The device is expected to be in use at festivals this summer in Britain, the Seattle Post-Intelligencer reported Tuesday. There’s no word *on if*, or when, the iLoo will make its way across the pond. (*Daily Times*, 2003/0/2)

The occurrence of *on if* in new contexts in SAE varieties can be explained by an extension of the synonymous use of *on if* and *on whether* from many other contexts in which both sequences are attested in native varieties, e.g. in *it depends on if* and *it depends on whether*. The emergence of the *on if* pattern in IndE can thus be traced back to what Christian Mair calls a “pre-determined breaking point” in the English language system (personal communication; see also Hoffmann and Mukherjee 2007), meaning that users of IndE extend the interchangeability of *whether* and *if* to new syntactic contexts. It is quite possible that this lexicogrammatical innovation has been spreading from IndE to the neighbouring varieties. In other words, while we did not find clear-cut evidence of IndE being an epicentre for SAEs the use of nativized *on if* indicates that it may well play this role. It also indicates that the norm providing status of IndE is more likely to be linked to local than to global patterns.

6. Conclusion

For our concluding remarks, we would like to get back to our hypotheses and assess them in the light of our findings.

Firstly, we did find differences between SAEs on the one hand and BrE on the other hand that show a potential for the development of a regional norm. These differences in the use of the subjunctive are, for the most part, qualitative in nature rather than quantitative. A greater propensity for the SAEs than BrE to use *would* in counterfactual *if*-clauses, for instance, is not confirmed by our results. But at the same time, it was only the SAEs that produced instances of *would* as a true variant for *were / was* in our data.

Secondly, we expected that IndE as the largest and most institutionalized variety would serve as a lead variety for the other varieties. But the overall quantitative results for the use of the subjunctive in counterfactual *if*-clauses in our web-derived newspaper data do not show all SAEs clustering together with IndE. Influence from AmE, the different degrees of institutionalization and, potentially, different grades of exonormative pressure from BrE on the individual varieties may be at the root of the gradient that we observe. It is at the qualitative level (e.g. tense sequencing) that the SAEs diverge from their parent variety and other L1-varieties.

Finally, our web-derived corpora provide us with a serendipitous result, namely a local use of a complex subordinating conjunction *on if*, which is more widely used in IndE than in the neighbouring varieties. It is possible that it has its origin in IndE and has begun to spread to other SAEs from there, but we lack the diachronic evidence to test this hypothesis. Moreover, there is the (remote) possibility that the innovation may have arisen independently in the other SAEs.

References

- Agnihotri, Rama Kant, Amrit Lal Khanna and Aditi Mukherjee. 1988. *Tense in Indian English. A Sociolinguistic Perspective*. New Delhi: Bahri.
- Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.
- Denison, David. 1998. "Syntax". In Suzanne Romaine, ed. *The Cambridge History of the English Language*. Vol. 4: 1776–1997. Cambridge: Cambridge University Press, 92–329.
- Garside, Richard. 1987. "The CLAWS word-tagging system". In Richard Garside, Geoffrey Leech and Geoffrey Sampson, eds. *The Computational Analysis of English: A Corpus-based Approach*. London: Longman, 30–41.
- Hartford, Beverly S. 1989. "Prototype effects in non-native English: object-coding in verbs of saying". *World Englishes* 8: 97–117.
- Hoffmann, Sebastian and Joybrato Mukherjee. 2007. "Ditransitive verbs in Indian English and British English: a corpus-linguistic study". *Arbeiten aus Anglistik und Amerikanistik* 32: 5–24.
- , Marianne Hundt and Joybrato Mukherjee. 2011. "Indian English — an emerging epicentre? A pilot study on light-verbs in web-derived corpora of South Asian Englishes". *Anglia* 129: 258–80.
- Hundt, Marianne. 2006. "A pale shadow? — The *were*-subjunctive in four 'major' or inner-circle varieties of English". Paper presented at the 27th ICAME conference, Helsinki.
- . 2009. "Global feature — local norms? A case study on the progressive passive". In Lucia Siebers and Thomas Hoffmann, eds. *World Englishes — Problems, Properties and Prospects*. Amsterdam, Philadelphia: Benjamins, 287–308.
- . fc. 2013. "Heterogeneity vs. homogeneity". To appear in Anita Auer, Daniel Schreier and Richard J. Watts, eds. *Letter Writing and Language Change*. Cambridge: Cambridge University Press.

- Johansson, Stig and Else H. Norheim. 1988. "The subjunctive in British and American English". *ICAME Journal* 12: 27–36.
- Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in Contemporary English. A Grammatical Study*. Cambridge: Cambridge University Press.
- Mukherjee, Joybrato and Marianne Hundt, eds. 2011. *Exploring Second-Language Varieties of English and Learner Englishes. Bridging a Paradigm Gap*. Amsterdam, Philadelphia: Benjamins.
- Peters, Pam. 1998. "The survival of the subjunctive: evidence of its use in Australia and elsewhere". *English World Wide* 19: 87–103.
- Ryan, William M. 1961. "Pseudo subjunctive *were*". *American Speech* 36: 48–53.
- Sayder, Stefan. 1989. "The subjunctive in Indian, British and American English: a corpus-based study". *Linguistische Arbeitsberichte* 69: 58–66.
- Schneider, Edgar W. 2000. "Corpus linguistics in the Asian context: exemplary analyses of the Kolhapur Corpus of Indian English". In Ma. Lourdes S. Bautista, Teodoro A. Llamzon and Bonifacio P. Sibayan, eds. *Parangal Cang Brother Andrew: Festschrift for Andrew Gonzalez on His Sixtieth Birthday*. Manila: Linguistic Society of the Philippines, 115–37.
- . 2005. "The subjunctive in Philippine English". In Danilo T. Dayag and Stephen Quakenbusch, eds. *Linguistics and Language Education in the Philippines and Beyond. A Festschrift in Honor of Ma. Lourdes S. Bautista*. Manila: Linguistic Society of the Philippines, 27–40.
- . 2011. "The subjunctive in Philippine English — an updated assessment". In Maria Lourdes S. Bautista, ed. *Studies in Philippine English: Exploring the Philippine Component of the International Corpus of English*. Mandaluyong City: Anvil Publishing, 159–73.
- Schneider, Gerold and Marianne Hundt. 2012. "'Off with their heads' — profiling TAM in ICE corpora". In Marianne Hundt and Ulrike Gut, eds. *Mapping Unity and Diversity World-Wide: Corpus-based Studies of New Englishes*. Amsterdam, Philadelphia: Benjamins, 1–34.
- Sedlatschek, Andreas. 2009. *Contemporary Indian English. Variation and Change*. Amsterdam, Philadelphia: Benjamins.
- Sharma, Devyani. 2001. "The pluperfect in native and non-native English: a comparative corpus study". *Language Variation and Change* 13: 343–73.

Authors' addresses

Marianne Hundt
Englisches Seminar
Universität Zürich
Plattenstrasse 47
CH-8032 Zürich
Switzerland
m.hundt@es.uzh.ch

Joybrato Mukherjee
Department of English
Justus Liebig University Giessen
Otto-Behagel-Str. 10B
D-35394 Giessen
Mukherjee@uni-giessen.de

Sebastian Hoffmann
Department of English
University of Trier
D-54296 Trier
hoffmann@uni-trier.de