Progressive or simple? A corpus-based study of aspect in World Englishes

Hundt, Marianne

Edinburgh University Press
Progressive or simple? A corpus-based study of aspect in World Englishes

Marianne Hundt, University of Zurich
Paula Rautionaho, University of Eastern Finland
Carolin Strobl, University of Zurich

Abstract

Previous corpus-based research on the progressive (BE + Ving) investigated it from a diachronic point of view or from the angle of World Englishes (WEs). However, factors such as its propensity to occur with animate subjects or its preference for dynamic verbs have not been studied in relation to the choice between progressive and simple aspect. As the progressive has been extended to stative verbs, we argue that a variationist study of the construction in WEs needs to systematically take simple VPs into account, too, and investigate whether there is interaction between predictor variables underlying the progressive:simple choice. We use a probabilistic grammar approach to study progressives in newspaper writing across a broad range of WEs. We apply a tree and forest analysis to gauge the relative strength of the predictor variables ‘regional variety’, ‘animacy’, ‘tense/modality’, ‘verb type’ and ‘voice’. Our results show that the core grammar for the progressive:simple choice is shared across all Englishes. The extension of progressives to stative verbs, in particular, does not result in statistically detectable
effects. We argue that they nevertheless serve to give a very ‘local’ flavour to contact varieties as they are salient against the backdrop of the core grammar.

**Keywords:** progressive: simple alternation, World Englishes, *International Corpus of English*, probabilistic modelling

1. **Introduction**

The progressive construction (*BE + V-ing*) has been a source of keen interest for researchers of first- and second-language varieties of English (see section 2). Researchers’ attention has been aimed at the increase in the frequency of the progressive since the 17th century, the reasons propelling the increase, and the emergence of novel uses such as *You’re always having to get there first*, where speaker stance is more important than aspectual meaning (e.g. Leech et al., 2009). With regard to progressives in WEs, linguists have been particularly interested in the identification of nativized patterns in language contact situations (see section 2). The extension of the progressive to non-standard stative situations or “verbs with strictly stative meanings” (Levin, 2013: 188), as in (1), has been said to be characteristic of Indian English, in particular (e.g. Mesthrie, 2005: 322).

(1) That means *you will not be knowing* that a carbon atom has four valency isn't it (ICE-IND, S1b-001).
Most previous research on the progressive uses normalised frequencies of the progressive to study a variety of factors that play a role in its spread and current usage. This approach was taken because the retrieval of simple VPs from unannotated corpora was not really feasible. An entirely different approach is to compare variation between progressive and simple VPs. Rogers (2002: 197) sees it as “… a next logical step” because of such extended uses as the one in (1). Moreover, treating the choice between progressive and simple as an alternation opens up the possibility of applying probabilistic modelling and thus the detailed analysis of the determinants of variation across a broad range of varieties. Such an approach allows to describe not only the core grammar of varieties but also variety-specific probabilistic peculiarities, which Szmrecsanyi et al. (2016: 133) refer to as ‘probabilistic indigenization’, i.e.

… the process whereby stochastic patterns of internal linguistic variation are reshaped by shifting usage frequencies in speakers of post-colonial varieties. To the extent that patterns of variation in a new variety A, e.g. the probability of item x in context y, can be shown to differ from those of the mother variety, we can say that the new pattern represents a novel, if gradient, development in the grammar of A.

It is these gradient patterns of variation that we aim to model in our study of the progressive. Syntactically annotated corpora of WEs, namely the PoS-tagged and parsed version of the International Corpus of English (ICE)\(^1\), allow us to retrieve both simple

---

\(^1\) The annotation was done with Schneider’s (2008) probabilistic dependency parser (Pro3Gres) and searched via the ICE Dependency Bank, i.e. the interface provided by
and progressive verb phrases (VPs). This makes it possible to study variation in WEs in terms of a progressive:simple alternation and with a multifactorial approach. The factors we investigate are tense, modal auxiliaries, voice, animacy of the subject and verb type.

In part two of the paper, we briefly review previous research on the progressive in different WEs. Part three provides background information on the corpus, the definition of the progressive:simple choice as an alternation, and how we coded for predictor variables that potentially influence the choice between variants of this alternation. In part four, we first present descriptive statistics of the predictor variables and then explore the relative importance of the different predictors for the progressive:simple alternation. In the final section, we discuss the relevance of our results for the field of WEs research.

2. Previous research into the progressive in World Englishes

Previous studies typically aimed to uncover regional differences in the overall frequency of the progressive. While initial research (e.g. Mair & Hundt, 1995) did not reveal any significant differences in the use of progressives between newspapers published in Great Britain and the US, Hundt (1998: 75-77) found New Zealand (NZE) and Australian English (AusE) to be more advanced in the ongoing spread of the progressive. Collins (2008) corroborates this regional difference on the basis of a broader range of both spoken and written texts. A series of recent papers have extended the research from first-language (ENL) varieties to contexts where English is used as an

---

Hans Martin Lehmann and Gerold Schneider at the University of Zurich (http://es-iceonline.uzh.ch).
institutionalised second-language (ESL) variety and also included learner Englishes (e.g. Hundt & Vogel, 2011; Meriläinen et al., 2017; Paulasto, 2014; Rautionaho, 2014; Van Rooy, 2014; Schilk & Hammel, 2015).

Previous research has also looked into the impact that individual factors may have on the spread and/or regional distribution of the progressive.

(a) Tense: institutionalised second-language (ESL) varieties use an even higher proportion of present tense progressives than first language (ENL) varieties (Rautionaho, 2014: 104; Salles Bernal, 2015).

(b) Modals: modal progressives seem to be slightly more frequent in SgE and particularly more frequent in IndE than in other varieties (Rautionaho, 2014; Salles Bernal, 2015)

(c) Voice: progressive passives show no regular distribution with respect to the ENL-ESL distinction, and while Hundt (2009) looks at other factors (such as speech:writing, tense, animacy) that might play a role for the use of progressive passives, the interaction of these factors cannot be studied unless the variable is defined as a choice context.

(d) Animacy: regional variation has been found to affect the proportion of animate subjects with the progressive in BrE, AmE and NZE (Hundt, 2004, Hundt & Szmrecsanyi, 2012).

(e) Verb type: The progressive typically combines with dynamic verbs, which refer to activities and events rather than to mental or physical states (see Biber et al., 1999: 471–75). Stative verbs are possible in the progressive when they refer to temporary events. Previous research shows that there is variability with regard to the use of extended stative (and habitual) progressives: in contact varieties of
English, the progressive combines with stative verbs without resulting in a change in meaning (see e.g. Paulasto, 2014; Rautionaho, 2014 and example (1)). In order to be able to gauge the relative importance that the factors listed above have on the use of the progressive, the variable needs to be defined as a choice context (i.e. the progressive:simple alternation). Only a few studies so far focus on the alternation and are thus able to employ advanced statistical analyses to assess how co-occurring linguistic factors affect the constructional choice. Different sets of predictor variables (such as semantic domain, register, Aktionsart) are studied in Rautionaho & Deshors (2018), Rautionaho et al. (2018), and Deshors & Rautionaho (2018). These studies use statistical approaches such as binary regression analysis, distinctive collexeme analysis, and hierarchical cluster analysis, and are able to show that the choice between progressive and simple VPs is often influenced by the combined influence of two factors (such as the interaction between AKTIONSART and GENRE; see Rautionaho & Deshors, 2018: 238). Our study aims to add to this novel line of research on progressives in WEs, by answering Rautionaho & Deshors’ call for “further applications of regression-based approaches such as classification trees and random forests” (2018: 248).

3. Data and methodology

In this section, we describe the rationale for choosing to restrict our study to newspaper data, and give some background information on the components ICE. We provide detail on data retrieval, the definition of our variable and the predictor variables. We also give a brief description of our approach to statistical modelling.
3.1 Choice of ICE components and annotation

ICE components are compiled using the same sampling frame to provide the empirical base for comparative studies across different WEs (Greenbaum, 1996). The first corpora were collected from materials produced in the late 1980s and early 1990s, but new corpora keep being added to the original eighteen sub-corpora. More recent ICE corpora use material that stem from the second decade of the current century, thus introducing a potential diachronic bias into the analysis (see Hundt, 2015). For language-internal predictor variables that influence the choice between a progressive and a simple VP, this slight diachronic bias should have little effect, however. In other words, while frequency developments might have occurred in this time span, language internal constraints on the use of the simple:progressive choice are less likely to have undergone dramatic change in the very recent past.

As the progressive is particularly frequent in spoken language (see Leech et al., 2009: 125) we would ideally have used the spoken components of ICE. However, as only the written part has been completed for some varieties (including AmE), we decided to opt for a broad regional coverage (see Table 1) and focus on the newspaper section, only. Newspaper writing is the most suitable of the written genres as it is open to ongoing change and the trend towards colloquialisation (Hundt & Mair, 1999). We extracted our data from the twenty 2,000-word samples of news reporting included in the fifteen currently available ICE corpora (i.e. a total of about 600,000 words). For a detailed study of the progressive:simple alternation, this provides more than enough data. We include ICE components from countries where English is the first language for the majority of speakers (i.e. Great Britain, Ireland, the US, Canada, New Zealand and
Australia), countries where English is an institutionalized second language (i.e. Singapore, India, Sri Lanka, Hong Kong, the Philippines, Ghana, Nigeria and Fiji), and finally, a country where English is used as a second dialect alongside an English-based creole (i.e. Jamaica).

The search algorithm used to retrieve progressives\(^2\) includes a form of the auxiliary \textit{be}, followed by possibly intervening adverbial elements, followed by a present participle verb form. This also retrieves variants with modal auxiliaries. The algorithm for simple VPs targets the base form, past tense forms, and present tense forms (including both 3\textsuperscript{rd}-person and non-3\textsuperscript{rd} person). Since the combination of a perfect with the progressive is generally rare (Leech et al., 2009: 124), we focus on past and non-past VPs, including combinations with modals. Since the total number of relevant VPs proved too numerous for qualitative analysis, we limit our study to samples of two hundred hits per regional variety (100 progressives and 100 simple VPs). This has the disadvantage that it precludes ‘variety’ to emerge from the statistical modelling as a main effect. But we thoroughly investigated whether variety moderates the effect of the other independent variables by means of interaction effects.

\(^2\) For the retrieval of progressives, we used the following regular expression:

\[\text{((am|is|are'|m'|s'|re|was|were)}\_VB. (\[^\_]+\_RB)*\[^\_]+\_VBG)\text{(be} \_VB} (\[^\_]+\_RB)*[^\_]+\_VBG)\text{(be} \_VB} (\[^\_]+\_RB)*[^\_]+\_VBG). \text{Simple VPs were retrieved by searching for } \_VB | \_VBD | \_VBP | \_VBZ.\]
3.2 Defining the variable: progressive vs. simple

We aim to compare usage of the progressive with that of simple VPs, i.e. variation between *He was driving along the road* vs. *He drove along the road*. This may seem unusual at first, particularly regarding standard grammatical descriptions of the progressive (e.g. Quirk et al., 1985), but we are not the first to do so: Smitterberg (2005), for instance, explored the option of measuring the increase of progressives against simple VPs. While standard reference grammars (e.g. Quirk et al., 1985: 198-199) point out that the progressive is generally ungrammatical with stative verbs, previous research (among others Leech et al., 2009) has shown that this is not the case (as e.g. in *I’m loving it*). Add to these the instances where the progressive is used to express speaker stance rather than aspectual meaning, the foregrounding function in narrative contexts and the extended uses found in ESL varieties, treating the progressive:simple choice as an alternation begins to make more and more sense. Importantly, instead of investigating the semantic aspects, we focus on the linguistic contexts of use of the two constructions, and aim to investigate if and how the factors described below explain the alternation (see Rautionaho & Deshors, 2018 for further discussion).

As Aarts et al. (2013: 20, 21) point out, identifying the variants of an alternation is not necessarily straight forward and partly subjective. Following Smitterberg (2005: 45-48), we excluded all imperative and non-finite simple VPs (e.g. example (2)) from

---

3 Couper-Kuhlen (1999) finds that the progressive can also be used in narrative contexts with a foregrounding function. An example from our ICE data would be “JYOTI Basu, [...], is suddenly finding the rug being pulled from under his feet in his own fiefdom” (ICE-IND, W2e-010). Mesthrie (2013) also observes the foregrounding use in his South African Indian English data.
our set of ‘progressivisable’ cases. In a similar vein, instances of going to that were future time expressions as in (3), were excluded because they are not progressives in the narrow sense and do not have simple counterparts.

(2)  It looked to be working when Greg Cooper cut through after the restart, and unloaded to Brewer who found Steve Cottrell for the try. (ICE-NZ, W2c-015)

(3)  “It’s going to be a fun week, I’m going to enjoy every day of it,” he said… (ICE-CAN, W2c-008)

From our retrieved datasets, we additionally excluded all instances that only superficially resemble progressives, e.g. copula be followed by an adjectival present participle, as in (4). We also removed instances such as (5), where the participle is ambiguous with respect to voice, seeing that an active reading would fall under the same conditions for exclusion as example (4) and be impossible to code for with respect to the predictor variable ‘voice’. Occasionally, the ing-form is not a participle but functions as a gerund: eradicating in sentence (6) forms a clausal constituent with the following noun rather than constituting part of a progressive.

(4)  THE message just outside the Gallowgate End of St James Park is enticing. (ICE-GB, W2c-004)

(5)  The place was actually covered with weeds. (ICE-NIG, W2c-001)

(6)  The ultimate goal of safety ropes is eradicating poverty. (ICE-SL, W2c-003)
We also excluded constructions where progressive equivalents were missing from our data, such as mediopassives, non-finite *being to V* and existentials with *there*.

In the case of ICE-Ghana, the total number of genuine progressives retrieved with our search algorithm was lower than 100 (i.e. 92), and we therefore only included 92 simple VPs from ICE-Ghana, giving us a total of 1492 progressives and 1492 simple VPs.

### 3.3. Predictor variables

The resulting 2984 instances were annotated for the following predictor variables and values:

(i) **Voice** (encoded as a factor with categories active ‘a’ vs. passive ‘p’)

(ii) **Tense & Modality** (encoded as an ordered factor with past ‘1’ vs. non-past ‘0’ vs. modal ‘2’, including future time expression with *will*)

(iii) **Animacy** (encoded as a factor with animate ‘a’ vs. inanimate ‘i’)

(iv) **Verb type** (encoded as a factor with dynamic ‘0’ vs. stative ‘2’)

(v) **Variety** (encoded as a factor with individual varieties but also by ENL vs. ESL/ESD)

The third predictor variable, semantics of the subject, needs some further comment, since it might not be straightforward as to how certain examples should be coded. ‘Animacy’ is, strictly speaking, not a binary choice but rather a gradient. Strang (1982), the first to investigate ‘animacy’ as a factor in the spread of the progressive, distinguishes between ‘human’, ‘quasi-human or animal’ and ‘inanimate’ subjects,
others have used more fine-grained distinctions. Zaenen et al. (2004), who distinguish ‘human’, ‘organisation’, ‘animal’, ‘place’, ‘time’, ‘concrete’, ‘nonconcrete’, ‘machines’, ‘vehicles’, and two codes for instances where the annotator was unsure of the categorisation. As there is no universal solution to the problem, we decided to code this variable in a binary way despite the gradient nature of ‘animacy’. It is therefore all the more important to comment on how we dealt with problematic cases. Seemingly ‘inanimate’ nouns that are used metonymically for humans (e.g. chair when used to refer to a professor) were coded as ‘animate’, for instance. Collective nouns like army are also problematic. A frequent co-referent singular pronoun for these is it, but they were coded as ‘animate’ on the rationale that the referents of these nouns are human beings, and thus animate (see example (7)):

(7)  The Callan Bacon company is also trading extremely well … (ICE-IRE, W2c-015)

As soon as collective nouns are used as subjects in a progressive passive, however, they appear to be less agentive, and we therefore decided to code them as ‘inanimate’; in the following example, the meaning of the verb, additionally, suggests that the subject is inanimate, since human beings cannot be eased out, as a general rule:

(8)  Filipino companies are deliberately being eased out … (ICE-PHI, W2c-002)

Example (8), then, illustrates that coding for animacy cannot only be done on the basis of the semantic interpretation of the subject NP. Occasionally, it is the interplay of the
verb and the subject that plays a role for the coding, as predicted by a construction grammar approach to language (see e.g. Goldberg, 1995 or Traugott & Trousdale, 2013). This approach also guided the coding of other instances. In concrete instantiations of a construction, elements can inherit properties from the construction. Thus, while the subject pronoun *it* in (9) was coded as ‘inanimate’, we took it to inherit ‘animacy’ from the verb *hope* in the impersonal construction in (10). If a verb that typically requires an ‘experiencer’ subject was used metaphorically, on the other hand, as in (11), the subject was coded as inanimate.

(9) There is a world supply glut and *it is not helping* that Singapore with its more efficient and modern refineries is cutting prices. (ICE-PHI, W2c-088)
(10) *It is hoped* that he will make that dream a reality today. (ICE-SL, W2c-011)
(11) She, however, expressed frustrations at the delays *corruption cases were suffering* at the courts. (ICE-NIG, W2c-001)

### 3.4 Statistical modelling

There are various ways to analyse the relative importance of predictor variables in a probabilistic grammar approach. Tagliamonte & Baayen (2012) show that traditional regression analysis (whether a generalised or a mixed model) is not suitable for highly correlated predictor variables and runs the risk of overfitting. We decided to adopt the
alternative approach they suggest and conduct the main analysis using a conditional inference tree (ctree) and a conditional random forest analysis in R.\footnote{All tree and tree ensemble analyses were conducted with the R-package party. For technical background information on how best to conduct such analyses, see Strobl, Malley & Tutz (2009) and Strobl, Hothorn & Zeileis (2009).}

Conditional inference trees use recursive partitioning and predict outcomes on binary splits of the data. The ctrees have the advantage that their graphical illustration can be interpreted straightforwardly (see Figure 6 in section 4.2). However, this may also lead to over-interpretation because individual trees are known to be instable (see below). Therefore, in addition to an individual ctree, we also present results on the stability of this tree as well as results for a random forest.

The random forest analysis provides more stable predictions, but is more difficult to interpret than a single tree. It provides only a descriptive measure of the importance of each predictor variable, summarizing its effect in predicting the progressive:simple choice. We therefore illustrate the way the predictors work together in complex interactions by means of plotting the predicted probability against all factor combinations. This kind of graphical illustration is very helpful for interpreting the underlying effects of the predictor variables.

To enable readers unfamiliar with trees to better understand the potential pitfalls in the interpretation of individual trees, we use an illustration from Philipp et al. (2016). Individual trees are known to be instable, i.e. small changes in the learning data can lead to a substantially different looking tree (see Strobl, Malley & Tutz, 2009, Bernaisch et al., 2014 and Philipp et al., 2016 and references therein). It is therefore important to caution against over-interpretation of single ctrees. What is particularly important to
Note is that the tree structure itself is only a representation of a partitioning of the underlying space of the predictor variables. Figure 1 (from Philipp et al., 2016) shows two trees that look very dissimilar at first sight. When we interpret the meaning of the trees, however, we find that both trees make exactly the same prediction in that the observations, characterized by their values of the predictor variables \((x_1 \text{ and } x_2)\), belong to the same class of the response variable (pink or green) (i.e. the equivalent of the variants in our choice context). This means that logically, and with respect to their substantial meaning (as the partition chart shows), the two superficially different trees are actually identical.

[FIGURE 1 HERE]

Note in particular that the position of the variables in the two trees with the same substantive meaning is not necessarily the same. The entire tree structure represents a complex interaction effect of the two variables. It therefore does not make sense to say, for example, that variable 1 was more important than variable 2, just because it appears in a higher position in tree 1.

For the analysis of real data this means that, while the variable first selected for splitting will be the one with the strongest main effect in the current sample, another variable might have had an only slightly weaker effect. However, due to the recursive nature of the tree where only one variable is selected at a time, this slightly weaker variable might not be selected again until lower in the tree. In other words, this does not mean that the latter variable was substantially less important than the former.
To account for both the instability of single trees and for the difficulty of comparing the importance of variables based on single trees, we perform a new form of stability analysis for single trees as well as a random forest variable importance analysis.

In order to assess the predictive quality of the ctree and random forest, we report the prediction accuracy as well as the concordance or ‘C-index’ (see Tagliamonte & Baayen, 2012). For random forests, besides computing these measures based on the learning sample, it is also possible to compute them using only those observations in the prediction that were not used for fitting the current tree, the so called out-of-bag observations.5 Since both response classes are of the same size in this data set, the baseline level (corresponding to random assignment for both the C-index and the prediction accuracy) is 0.5, and values approaching 1 indicate a good performance.

4. Results

4.1 Variables co-occurring with the progressive

Before we turn to the tree and forest analysis, we present summary statistics on the individual predictor variables, starting with voice. Figure 2 shows that both progressive and simple aspect prefer the active in newspaper language across the Englishes studied here: overall, only around 10% of all tokens are passives. In other words, even though the progressive passive (as in (9)) is a relatively recent structural innovation and met with considerable prescriptive resistance (see van Bergen, 2013 and Anderwald, 2014,

---

5 Out-of-bag performance measures are more realistic estimates for the performance on new data than measures computed from the learning data, and thus are a better indicator of the generalisability of the results (see Strobl et al., 2009).
respectively), passives in the ICE newspaper data are similarly dispreferred with both the simple and the progressive aspect. And while individual varieties show a slight preference for simple passives (NZE, NigE, and PhilE) and others for progressive passives (AusE, IrE, GhE and JamE) there is no clear divide along the ENL-ESL divide in this (see Hundt, 2009).

(12) Others indicted by the commission are being sought by the police already. (ICE-NIG, W2c-001)

With regard to our second predictor variable, Table 2 shows the proportion of present tense VPs as opposed to past tense or combinations with a modal verb.

Our data reveal that progressive and simple behave differently with respect to their propensity to combine with present tense (59% vs. 29%) across our data. The progressive’s preference for present tense is only slightly more pronounced in the contact varieties (ENL 57% : ESL & ESD 60%) and there is considerable variation within both variety types (see Figure 3) with BrE, AmE and CanE favouring present tense more clearly than AusE and NZE, for instance.
With respect to the semantics of the subject, the ICE newspaper material shows that, again, progressives and simple VPs differ: the affinity to co-occur with animate subjects is more marked for the progressive in all WEs (as in (13); see Figure 4).

(13) *We are advising* anyone affected to contact their insurers. (ICE-GB, W2c-012)

[FIGURE 4 HERE]

Finally, with respect to verb type, we see that while dynamic verbs are preferred over stative verbs throughout, the prevalence of progressives to combine with a dynamic verb is much more pronounced at 89% against 64% for the simple aspect. Figure 5 shows that the progressive’s marked preference for dynamic verbs holds across all varieties in our newspaper data.

[FIGURE 5 HERE]

On closer inspection, standard stative progressives, i.e. ones referring to temporary states, are found in all 15 varieties, whereas extended stative progressives, i.e. ones referring to permanent states, (as in (14) and (15)) were found in five varieties only.

(14) Presently, the group *is having* operations in more than 50 countries. (ICE-NIG, W2c-001)

(15) A bogus telephone caller who threatens to shoot the women he speaks to *is concerning* Christchurch police. (ICE-NZ, W2c-010)
Interestingly, the ICE data show that extended use of stative progressives is also occasionally found in ENL varieties (one occurrence each in ICE-NZ and ICE-IRE), but they are more frequent in our ESL data (four occurrences in SLE, two in NigE and one in GhE), which conforms results obtained in previous studies (e.g. Meriläinen et al., 2017). However, the overall proportion of extended stative progressives in the present data is low compared to findings from spoken data (see Rautionaho, 2014). This indicates that the spread of the progressive to permanent situations is more advanced in casual conversation than in newspaper language.

In sum, we see that the contextual variables affect the choice between progressive and simple VPs to varying degrees. Only a multifactorial analysis will reveal, however, which of the predictors has the strongest effect on the choice and, importantly, whether there is any interaction with the variable ‘variety’ (i.e. whether we can see probabilistic indigenization at work).

4.2 Tree and forest analysis: predicting the progressive across World Englishes

To investigate any differences between the English varieties, ‘variety’ was offered as a predictor variable to the conditional inference tree algorithm both in its original form (with 15 different categories) and in a binarised form (ENL vs. ESL). However, ‘variety’ was not selected into the tree in either form, indicating that it was not significant for the progressive:simple alternation in the ICE newspaper data. The other four variables were selected into the tree (see Figure 6).

[FIGURE 6 HERE]
As pointed out in section 3.4, the graphical representation of this tree needs to be interpreted with some caution. The interpretation of Figure 6 that follows is therefore only a preliminary result that we aim to validate by means of the stability analysis and the random forest below.

The dark grey bars in Figure 6 illustrate the predicted probability of PROGRESSIVE (as opposed to SIMPLE in light grey) in the different combinations formed by the predictor variables. When we look at the terminal nodes with the highest proportions of PROGRESSIVE, we find that these correspond to paths with dynamic and present tense verbs; ‘stative verb’ leads to nodes with lower probabilities of PROGRESSIVES, as do past tense VPs and non-past with a modal. Animate subjects, on average, lead to nodes with a slightly higher probability of PROGRESSIVE than inanimate ones. Voice is selected only in some of the terminal nodes and the direction of its effect depends on the levels of the other variables, indicating a high-order interaction effect. The findings from the single tree are summarized in Table 3. We will validate these interpretations with the random forest analysis (see Figure 7).

[TABLE 3 HERE]

The predictive performance of the individual tree on the learning sample shows a prediction accuracy of 0.714 and a C-index of 0.780 (notably above the random assignment level of 0.5, even though a C-index of 0.8 is usually considered as the threshold for a good performance, see e.g., Tagliamonte & Baayen, 2012). We compare these values to those from the random forest below.
In order to assess the stability of trees on empirical data sets, Philipp et al. (2016) have developed a toolkit for assessing how stable the variable and cutpoint selection of a tree is with respect to random changes in the learning data. The results of this stability analysis are described in Appendix 1. They indicate the variables ‘tense/modality’, ‘animacy’ and ‘verb type’ are selected in all trees generated on random samples of the data, whereas the variables, ‘variety’ and ‘voice’, were selected in many but not all trees. An additional visual cutpoint analysis also indicated that the variable ‘variety’ does not interact systematically with the other predictor variables.

In order to further investigate the relevance of the individual variables and their possible interaction effects, we conducted a random forest analysis. We used an ensemble size of 500 trees. The number of randomly pre-selected predictor variables in each split (mtry) was set to the value 2 based on the out-of-bag C-index (see Appendix 2 for details). Default settings were used for all other parameters, including the settings for unbiased variable selection (cforest_unbiased).

Often random forests with mtry smaller than the number of predictor variables show the best predictive performance, because in this case variables that dominate the tree structure (for example because they have a strong main effect) are suppressed in some splits, so that other variables have a chance to prove their predictive power (see e.g. Strobl et al., 2009 and Bernaisch et al., 2014). However, as shown in Appendix 2, in this particular data set, the influence of mtry on the prediction accuracy was not very strong. Therefore, while in the following we will present the results for a random forest with mtry = 2, very similar results can be found for other values.

While increasing the stability and quality of the predictions, a random forest loses the interpretability that a single tree could offer. It is important to understand that
there is no way of directly visualizing the entire forest, but we will present two means of exploring the random forest: a variable importance ranking and a plot of the predicted response probabilities illustrating the effects of the factors.

The variable importance plot (see Figure 7) illustrates that – in accordance with the results for the single tree – the variables ‘tense/modality’, ‘animacy’ and ‘verb type’ are the most relevant for predicting the response. As the variable importance in a random forest analysis does not distinguish between main effects and interaction effects of a variable, and also depends on other characteristics of the data set as well as on tuning parameter choice, it should only be interpreted as a ranking of the overall importance of a variable in a descriptive way (Strobl, Malley & Tutz, 2009). Since negative values can only result from random variation, it is fair to say that any positive values of the same magnitude as negative ones can be considered as noise (Strobl, Malley & Tutz, 2009). However, this is merely a conservative rule of thumb and should not be over-interpreted as a significance test. Note that here we report the unconditional variable importance, but the conditional importance leads to the same variable ranking in this case. We have also checked the stability of these results against different random starting seeds and found that the variable importance pattern is very stable.

What is visually striking in the random forest variable importance analysis (Figure 7) is that the importance of the variables ‘tense/modality’ and ‘verb type’ are notably higher than that of ‘animacy’. This is in accordance with the results from the single tree, where ‘animacy’ was also selected in the original tree and all replications of our tree stability analysis, but corresponded to smaller differences in the class proportions than the other two predictor variables (as summarized in Table 3).
The results displayed here are for a random forest with the binarised version of ‘variety’ following the argumentation in Appendix 2. This will also enhance the interpretability of the following interaction plots. However, if the multi-categorical version of variety is used, its variable importance is still below zero (results not shown for brevity), indicating that the variable does not contribute to predicting the response in either format, and the ranking of the remaining variables also remains the same.

The predictive performance of the random forest presented here with mtry = 2 and the binarised version of ‘variety’ was highest with respect to the out-of-bag C-index (see also Appendix 2): the in-sample C-index and prediction accuracy were 0.779 and 0.714. When compared to the in-sample values of the single tree (0.780 and 0.714), we find only a slight improvement, indicating that the tree already captured all relevant patterns in the data. As is to be expected, the more realistic and thus more conservative out-of-bag estimates of C-index and prediction accuracy for the random forest were even slightly lower, with 0.772 and 0.711.

The results in Appendix 2 show only slightly higher out-of-bag predictive performance measures for models including the variable ‘variety’ as opposed to leaving it out of the analysis, which supports our earlier findings from the individual tree, that did not select variety at all, and the low variable importance for this predictor in Figure 7.

With respect to interpretation, the variable importance plot in Figure 7 gives only a very broad idea of the effect that each predictor has on the progressive: simple choice. The variable importance plot does not, however, provide any insight into the
form and direction of the effects. Since a random forest captures potentially complex interactions of several variables, it is often termed a ‘black box’ in the sense that it gives very good predictions but the underlying functional form may be so complex that it is not directly interpretable (unlike for example in a parametric regression model, where the functional form may be an over-simplification of the true association structure, but the coefficients for main effects and interactions are directly interpretable).

One way of obtaining at least some insight into the black box of a tree ensemble is through partial dependence plots (see also Bernaisch et al., 2014 and Szmrecsanyi et al., 2016). These visualize the effect of one variable on the predicted probability of the alternation while holding all others constant at their average value. In the presence of interaction effects, looking at one variable at a time may lead to misinterpretations because different effects of one variable between the levels of another variable may cancel out when the second variable is averaged over. Luckily, the number of predictor variables in this study as well as the number of levels that they can take is small enough to plot the predicted response probability in all combinations of the levels of the predictor variables. This type of illustration was inspired by Szmrecsanyi et al. (2016), who used a similar illustration for a combination of some of their predictor variables.

[FIGURE 8 HERE]

The patterns in Figure 8 show that the predicted probability of PROGRESSIVE over SIMPLE depends on an interaction of ‘tense/modality’, ‘verb type’ and ‘animacy’. Only small differences in the predicted probabilities show in some combinations between
‘active’ and ‘passive’ (blue and pink lines). Crucially for a study aiming to verify whether WEs show subtle variations in the factors predicting the progressive: simple choice, we see that there are virtually no differences between ENL and ESL varieties (upper and lower row). This is in accordance with the variable importance ranking of ‘tense/modality’ > ‘verb type’ > ‘animacy’, with ‘voice’ and ‘variety’ displaying values close to or even below zero, as well as with the results of the single tree, where ‘voice’ was only selected in some of the end nodes and ‘variety’ was not selected at all.

With respect to ‘tense/modality’ (on the x-axis in Figure 8), present tense (0) again shows the highest predicted probability of PROGRESSIVE. In most combinations with other predictor variables, the predicted probability decreases with past (VerbType 1) and modal (2) VPs, which is also in accordance with the single tree. Sentence (16) thus is the most highly prototypical progressive in that it involves the present tense, a dynamic verb and an animate subject. Only in combination with dynamic verbs (0) and animate subjects (column 1 in Figure 8; see (17)) do we find a slight u-shaped pattern with past (1) showing the lowest probability of PROGRESSIVE. This pattern indicates a three-way interaction between the variables and is also represented in the tree.6

(16) We are going through a major change in work practices and culture. (ICE-SING, W2c-003)

6 Note that this is not very easy to see: The average predicted probability in nodes 5 and 6 in Figure 6, weighted by the node size n, for the combination ‘dynamic’, ‘animate subject’, and ‘present’ is high; the predicted probability in node 10 for ‘dynamic verb’, ‘animate subject’, and ‘past’ is low; the average predicted probability in nodes 12 and 13, again weighted by the node size n, for ‘dynamic verb’, ‘animate subject’ and ‘modal VP’ is medium.
(17) Mr. Kufour was commenting on the allegations... (ICE-GH, W2c-001)

For dynamic verbs (VerbType 0, columns 1 and 3 in Figure 8) we find on average higher predicted probabilities of PROGRESSIVE than for stative verbs (VerbType 2), which indicates a positive main effect of dynamic verbs. For animate subjects (columns 1 and 2 in Figure 8) we find slightly higher predicted probabilities than for inanimate subjects, which also indicates a slight positive main effect of animate subjects. These results support those of the single tree summarized in Table 3.

The overall highest predicted probabilities of PROGRESSIVE are found for the combination of present tense, dynamic verbs with animate subjects (highest predicted probability in leftmost point in column 1), which is again in accordance with the single tree result.

The single tree for all observations also split in ‘voice’ in some of its lower nodes. This is somewhat surprising since this predictor shows a negligibly low variable importance in the random forest. However, the large overall sample size induces a high statistical power even for splits at lower levels of the tree, and the direction of the effects is in accordance with our interaction plot from the ensemble: For example for dynamic verbs with a modal and an animate subject (rightmost points in column 1 in Figure 8, corresponding to nodes 12 and 13 in Figure 6), active voice (as in (18)) has a higher predicted probability for PROGRESSIVE than passive, while for stative verbs in the present tense following an inanimate subject (leftmost points in column 4 in Figure 8, corresponding to nodes 21 and 22 in Figure 6), active voice has a lower predicted probability for PROGRESSIVE than passive (as in (19); see also Table 3).
(18) He will be directing a production of Hamlet for this season. (ICE-CAN, W2c-012)

(19) She bemoaned the economic status of the teachers and said the neglect of education is being reflected in the high level of poverty in the nation. (ICE-JA, W2c-006)

5. Discussion: progressives, probabilistic indigenization and structural nativization

The careful statistical modelling of the choice between progressive and simple as a grammatical alternation returned ‘tense’ and ‘verb type’ as important predictor variables as well as ‘animacy’ of the subject. Co-occurrence with active voice could not be validated as an important predictor in all of the models (i.e. it only occurred at lower levels in the single tree). While previous grammatical description mentions these factors as relevant for the use of the progressive as such, our multifactorial analysis adds to the picture in that we are now able to show that tense (present) and verb type (dynamic) are more important than animacy (animate subject) and voice, which fits in well with historical accounts of the grammaticalisation of the progressive (e.g. Strang, 1982; Hundt, 2004).

With respect to variation across WEs, our multifactorial analysis does not support the notion of probabilistic indigenization in this area of grammar. While we retrieved our data in a way the precluded ‘variety’ to emerge as a main effect, it could have shown up as a variable interacting with other predictors in our analysis. Since this was not the case in either the ctree nor the random forest analysis, we can safely assume that for the progressive:simple choice the core grammar of English (at least with respect to newspaper language) is, indeed, shared across ENL, ESL/ ESD varieties. Note,
however, that among the stative verbs included in the modelling, we did not distinguish between standard and more ‘novel’ extensions of the progressive, as the latter are very rare (see section 4.1). Moreover, non-standard stative progressives occasionally occur in ENL newspapers. We also found them to be rare in our ESL sample. In addition to the occasional use of stative progressives, we also found a smattering of other extended uses, such as habitual progressives (as in (20)), and so-called perfective progressives\(^7\) (as in (21); see e.g. Paulasto, 2014).

(20) One man in particular pleaded: We’re playing into the hands of the media, over and over again. (ICE-JA, W2c-002)

(21) It is for the 21st time the Railways is winning the trophy compiling 391 points. (ICE-IND, W2c-020)

Surprisingly, our ICE data did not yield any instances of extended stative or habitual progressives for IndE, even though these are said to be established in and characteristic of this variety. This, together with the overall low frequency of extended uses in all varieties under investigation, indicates that newspaper language is not likely to yield patterns of indigenization. Editorial practices may be the key here: awareness of the potential unacceptability of a stative or habitual progressive referring to permanent situations, for instance, may lead to such features being omitted (see Kruger & van Rooy, 2017).

Taken together, the evidence from the probabilistic modelling and the qualitative analysis of the progressives shows that frequency of use is not necessary for a feature to

\(^7\) In (21), it is evident from the context that the competition is already over.
be perceived as typical of a variety or group of varieties: despite their low discourse frequency, these patterns are obviously still salient. This finds support from psycholinguistic research: a feature in a contact variety of English does not necessarily have to occur with a statistically higher frequency than in ENL varieties to be salient. On the contrary, it is typically infrequent phenomena that stand out against the backdrop of the expected and thus are noticed (in this case by linguists). Ellis (2017: 79), for instance, explicitly links salience with surprisal and low frequency: “Surprisal is inversely related to probability. Research operationalizations of surprisal in language involves computing norms in corpora of usage, and then looking for violations of those norms.” When applied to the present case study, this enables us to account for the fact that, while extended progressives do not occur frequently enough in the ESL varieties to result in a statistically significant effect in a multivariate model, extension of the progressive beyond the standard contexts is a phenomenon typically associated with varieties like IndE precisely because they violate the norms of what speakers normally do. Thus, while we do not see probabilistic indigenization at work in the grammar of the progressive: simple choice across WEs newspaper subcorpora, we can observe the structural nativization of low-frequency, salient patterns.

APPENDICES

These appendices provide additional information about details of the statistical analysis.

Appendix 1: Stability Analysis
In order to assess the stability of trees on empirical data sets, Philipp et al. (2016) have developed a toolkit for assessing how stable the variable and cutpoint selection of a tree is with respect to random changes in the learning data. This toolkit is used here to assess the stability of the original ctree. The underlying framework uses repeated random sampling to mimic what would happen if new data could be drawn from the population. By default, a tree is fit to each of 500 bootstrap samples and the variable and cutpoint selection is summarized in the following tables and plots.

[TABLE A1 HERE]

When a variable’s relative selection frequency is 1.0 over all 500 bootstrapped trees, this means that the variable was selected in every single tree. This means that we can reliably conclude that the variable is relevant for predicting the aspect choice. This is the case for the variables ‘tense/modality’, ‘animacy’ and ‘verb type’. The other two variables, ‘variety’ and ‘voice’, were selected in many but not in all trees.

[FIGURE A1 HERE]

The cutpoint selection plots illustrate the locations of the cutpoints (corresponding to the grouping of the categories in categorical predictors) over the 500 trees. The cutpoints/groupings chosen in the original tree are marked in red. The red numbers indicate at which level the split occurred in the original tree. Cutpoints/groupings can be selected more than 500 times (the number of repetitions in this example, indicated by
the dashed black line), when splits in this variable appeared several times in different positions in some or each of the 500 trees (for example in parallel branches).

Those variables that are binary (i.e. ‘animacy’, ‘verb type’ and ‘voice’) offer only one possible cutpoint and thus each split creates the same two groups. For the ordinal variable ‘tense/modality’ both possible cutpoints, between values ‘present’ (0) and ‘past’ (1) and between ‘past’ (1) and ‘modal’ (2), appear almost equally often. In the multi-categorical variable ‘variety’, we would see bigger blocks of colour if certain categories always ended up in the same branch after splitting. This is not the case here, as the larger areas for the leftmost categories are only due to the fact that the colouring starts on the left hand side. If two or more categories were always in the same branch, they would form bars of the same colour over the entire height of the plotting region, where each line stands for a split in the variable over all splits in all 500 trees. This indicates that ‘variety’ shows no indication of interacting systematically with other predictor variables.

Appendix 2: Tuning parameter choice for random forest analysis

The following tables (Tables A2 and A3) list the in-sample and out-of-bag C-indices and prediction accuracies (rounded to three digits) for tree ensembles with different values of mtry and different treatments of the variable variety. For comparison: The in-sample C-index and prediction accuracy for the single tree were 0.780 and 0.714. However, the out of sample estimates provided by tree ensembles are more realistic estimates of the performance on new data. The highest value in each column is underlined, the overall highest out-of-bag value is printed in boldface. The highest out-of-bag C-index is achieved by a random forest with mtry=2 using the binarised version
of the variable variety. The results for the prediction accuracy are not as conclusive, as several values of mtry lead to the same or similar accuracies. This indicates that with respect to prediction accuracy (and we have checked that this also applies to the variable importance ranking) the choice of mtry (as long as it is greater than one, which corresponds to the random selection of splitting variables) does not have a strong effect in this data set. Since the rounded off square root of the number of variables (here $\sqrt{5} = 2.4$, rounded off to 2) is also the standard suggestion for mtry in classification problems, the results for a random forest with mtry=2 using the binarised version of the variable variety are presented in the main part of this paper.

[TABLES A2 and A3 HERE]

REFERENCES


Schneider, G. 2008. Hybrid Long-Distance Functional Dependency Parsing. PhD, University of Zurich.


Tagliamonte, S. and H. Baayen. 2012. ‘Models, forests, and trees of York English:
Was/were variation as a case study for statistical practice’, Language Variation and Change 24, pp. 135–178.


Table 1. The International Corpus of English (ICE) – components used in the study.

<table>
<thead>
<tr>
<th>Variety status</th>
<th>Variety acronym</th>
<th>Corpus Acronym</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENL</td>
<td>BrE</td>
<td>ICE-GB</td>
<td>Great Britain</td>
</tr>
<tr>
<td></td>
<td>IrE</td>
<td>ICE-IRE</td>
<td>Ireland</td>
</tr>
<tr>
<td></td>
<td>NZE</td>
<td>ICE-NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td></td>
<td>AusE</td>
<td>ICE-AUS</td>
<td>Australia</td>
</tr>
<tr>
<td></td>
<td>CanE</td>
<td>ICE-CAN</td>
<td>Canada</td>
</tr>
<tr>
<td></td>
<td>AmE</td>
<td>ICE-US</td>
<td>United States of America</td>
</tr>
<tr>
<td>ESL</td>
<td>SingE</td>
<td>ICE-SING</td>
<td>Singapore</td>
</tr>
<tr>
<td></td>
<td>HKE</td>
<td>ICE-HK</td>
<td>Hong Kong</td>
</tr>
<tr>
<td></td>
<td>PhilE</td>
<td>ICE-PHI</td>
<td>The Philippines</td>
</tr>
<tr>
<td></td>
<td>IndE</td>
<td>ICE-IND</td>
<td>India</td>
</tr>
<tr>
<td></td>
<td>SLE</td>
<td>ICE-SL</td>
<td>Sri Lanka</td>
</tr>
<tr>
<td></td>
<td>GhE</td>
<td>ICE-GH</td>
<td>Ghana</td>
</tr>
<tr>
<td></td>
<td>NigE</td>
<td>ICE-NG</td>
<td>Nigeria</td>
</tr>
<tr>
<td></td>
<td>FijE</td>
<td>ICE-FJ</td>
<td>Fiji</td>
</tr>
<tr>
<td>ESD</td>
<td>JamE</td>
<td>ICE-JAM</td>
<td>Jamaica</td>
</tr>
</tbody>
</table>
The response class is indicated by the colouring (pink for one class and green for the other). The terminal nodes of each tree are indicated with a capital T followed by the number of the terminal node; the same numbering is used to point out the respective section in the partition charts.
Figure 2. % passives in simple vs. progressive VPs across newspaper section of 15 ICE corpora.
Table 2: % Tense & modal with progressive:simple VPs across newspaper section of 15 ICE corpora.

<table>
<thead>
<tr>
<th></th>
<th>Progressive</th>
<th></th>
<th></th>
<th>Simple</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-past</td>
<td>past</td>
<td>modal</td>
<td>non-past</td>
<td>past</td>
<td>modal</td>
</tr>
<tr>
<td>ENL</td>
<td>56.7%</td>
<td>36.0%</td>
<td>7.3%</td>
<td>30.8%</td>
<td>55.3%</td>
<td>13.8%</td>
</tr>
<tr>
<td>ESL+ESD</td>
<td>59.8%</td>
<td>33.6%</td>
<td>6.6%</td>
<td>28.5%</td>
<td>55.5%</td>
<td>16.0%</td>
</tr>
</tbody>
</table>
Figure 3. % present VPs and the progressive:simple alternation across newspaper section of 15 ICE corpora.
Figure 4. % animate subjects in progressive:simple VPs across newspaper section of 15 ICE corpora.
Figure 5. % dynamic verbs and aspect choice across newspaper section of 15 ICE corpora.
Figure 6. Conditional inference tree analysis predicting the use of PROGRESSIVE across 15 varieties of English (ICE newspaper section).
Table 3. Summary of effects visible in the tree. In this table, the > sign means that for one level of a variable the proportion of progressive is higher than for another; >> stands for a particularly strong difference, >< stands for effects that vary in direction. The node numbers in the tree refer to the little numbers in boxes for internal nodes and above the terminal nodes in Figure 6. Average proportions over several nodes can be judged by weighting the proportions with the respective node sizes n.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Indication in the tree in Figure 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>dynamic verbs &gt;&gt; stative verbs</td>
<td>Left arm of tree beneath node 1 has higher proportions on average</td>
</tr>
<tr>
<td>present &gt;&gt; past and modal</td>
<td>Left branches of tree beneath nodes 2 and 19 have higher proportions on average</td>
</tr>
<tr>
<td>animate subject &gt;</td>
<td></td>
</tr>
<tr>
<td>inanimate subject</td>
<td></td>
</tr>
<tr>
<td>active &gt;&gt; passive</td>
<td>&gt; nodes 5 vs. 6 and 12 vs. 13</td>
</tr>
<tr>
<td></td>
<td>&lt; nodes 21 vs. 22</td>
</tr>
</tbody>
</table>
Figure 7. Random forest permutation of progressive:simple alternation data – variable importance plot.
Figure 8. Visualization of predicted probabilities of progressive:simple in all combinations of predictor variables.
Table A1. Summary of variable selection stability for the tree analysis.

<table>
<thead>
<tr>
<th>‘tense/modality’</th>
<th>Relative selection frequency in bootstrap samples</th>
<th>Selection in original tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘animacy’</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>‘verb type’</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>‘variety’</td>
<td>0.946</td>
<td>0</td>
</tr>
<tr>
<td>‘voice’</td>
<td>0.928</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure A1. Graphical summary of cutpoint selection stability for the tree analysis.
Table A2. In-sample and out-of-bag C-indices for tree ensembles

<table>
<thead>
<tr>
<th>C-index</th>
<th>Variety original</th>
<th>Variety binarized</th>
<th>Variety not included</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In-sample</td>
<td>Out-of-bag</td>
<td>In-sample</td>
</tr>
<tr>
<td>mtry = 1</td>
<td>0.776</td>
<td>0.754</td>
<td>0.775</td>
</tr>
<tr>
<td>mtry = 2</td>
<td>0.793</td>
<td>0.769</td>
<td>0.779</td>
</tr>
<tr>
<td>mtry = 3</td>
<td>0.802</td>
<td>0.771</td>
<td>0.782</td>
</tr>
<tr>
<td>mtry = 4</td>
<td>0.809</td>
<td>0.763</td>
<td>0.782</td>
</tr>
<tr>
<td>mtry = 5</td>
<td><strong>0.813</strong></td>
<td>0.754</td>
<td>0.783</td>
</tr>
</tbody>
</table>

Table A3. In-sample and out-of-bag prediction accuracies for tree ensembles.

<table>
<thead>
<tr>
<th>accuracy</th>
<th>Variety original</th>
<th>Variety binarized</th>
<th>Variety not included</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In-sample</td>
<td>Out-of-bag</td>
<td>In-sample</td>
</tr>
<tr>
<td>mtry = 1</td>
<td>0.649</td>
<td>0.648</td>
<td>0.650</td>
</tr>
<tr>
<td>mtry = 2</td>
<td>0.714</td>
<td>0.709</td>
<td><strong>0.714</strong></td>
</tr>
<tr>
<td>mtry = 3</td>
<td>0.714</td>
<td>0.705</td>
<td><strong>0.714</strong></td>
</tr>
<tr>
<td>mtry = 4</td>
<td>0.720</td>
<td>0.699</td>
<td><strong>0.714</strong></td>
</tr>
<tr>
<td>mtry = 5</td>
<td><strong>0.726</strong></td>
<td>0.683</td>
<td>0.713</td>
</tr>
</tbody>
</table>