



**University of
Zurich** UZH

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

**Social constraints on syntactic variation: The role of gender in Jamaican English
ditransitive constructions**

Röthlisberger, Melanie

DOI: <https://doi.org/10.1017/9781108696739.007>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-190220>

Book Section

Accepted Version

Originally published at:

Röthlisberger, Melanie (2021). Social constraints on syntactic variation: The role of gender in Jamaican English ditransitive constructions. In: Bernaisch, Tobias. Gender in World Englishes. Cambridge: Cambridge University Press, 147-175.

DOI: <https://doi.org/10.1017/9781108696739.007>

Title: Between social and linguistic constraints: The role of gender in Jamaican English ditransitive constructions (accepted version)

To be published in: Tobias Bernaisch (ed.), *Gender in World Englishes*. Cambridge: CUP.

Author: Melanie Röthlisberger

Affiliation: University of Zurich

Schönberggasse 2

8001 Zürich

Switzerland

Email: melanie.roethlisberger@uzh.ch

Social constraints on syntactic variation: The role of gender in Jamaican English ditransitive constructions

Abstract

The present study explores the effect of speakers' gender¹ in the well-known dative alternation (e.g. *Mary gives John an apple* vs. *Mary gives an apple to John*) and weighs the impact of this language-external factor against language-internal factors such as length of the constituents or semantics of the verb. Following up on previous research (Röthlisberger, Grafmiller & Szmrecsanyi 2017) that explored the dative alternation across nine varieties of English, the focus of the present work will be on Jamaican English, one variety where male and female speakers seem to use the two variants differently (see Röthlisberger 2018: 202). 615 variable dative tokens of acrolectal Jamaican English speech were annotated for 11 language-internal and three language-external factors and subjected to conditional random forest and mixed-effects logistic regression analyses. The results of these analyses indicate that the predictor GENDER plays only a marginal role vis-à-vis other language-external and -internal constraints. At the same time, if only the two most important language-internal predictors are considered, GENDER turns out to significantly affect dative choice with male speakers preferring the prepositional variant more than female speakers. These results not only highlight the potential of syntactic alternations to serve as sociolinguistic variables but also point to possibly different social dynamics between male and female speakers in Jamaica.

1 Introduction

Syntactic alternations such as the English dative alternation (see 1-2) have been researched to an almost excessive degree in recent decades.

(1) Double object variant

¹ Note that I use the term 'gender' here for two reasons: 1) The title of this volume is "Gender in World Englishes" and choosing 'gender' rather than 'sex' keeps this study within the streamline of the volume's topic. 2) The data for this analysis comes from the International Corpus of English. For some parts of this corpus (spoken informal conversations mostly), the background information of speakers was collected by way of a questionnaire where every informant could self-indicate their gender.

I'd given [Heidi]_{RECIPIENT} [my T-Shirt]_{THEME}. (ICE-GB:S1B-066:196)

(2) Prepositional variant

And I'd given [the key]_{THEME} to [Helen]_{RECIPIENT}. (ICE-CAN:S1A-058:183)

From these studies, we know that the choice between the double object and the prepositional variant is governed simultaneously and non-deterministically by several factors (see Bresnan et al. 2007; Bernaisch, Gries & Mukherjee 2014; Röthlisberger, Grafmiller & Szmrecsanyi 2017, among others). Attention has thereby been mostly restricted to language-internal factors, such as definiteness, givenness, length or animacy of the recipient and theme, verb semantics, and so on (see Röthlisberger 2018 for a lengthy list of such factors). In contrast, language-external factors such as style, register or regional background of speakers – while included in some studies – have been shown to play mostly a minor role, depending on the syntactic variable investigated (see Szmrecsanyi et al. 2016). On an even more fine-grained level, factors relating to the individual speakers (e.g. age, sex, education) have largely been omitted from the analysis of syntactic variation. As Cheshire (2003) points out, this omission could be due to the general argument among variationist linguists that syntactic variation is constrained mostly by language-internal, cognitive constraints rather than by social factors (Cheshire 2003: 245; see also Lavandera 1978). Recent studies, however, highlight that the choice of syntactic variants can be constrained by speaker-related factors even though these factors might not be as influential as language-internal ones. With regard to the dative alternation, current work has thereby focused mostly on genderlectal differences and language change across different age groups. This line of research shows that the choice of variant in the dative alternation is indeed socially variable across gender with male speakers preferring the prepositional variant more than female speakers in Australian English (e.g. Theijssen et al. 2011) and in British English (Jenset, McGillivray & Rundell 2018); in Canadian English, the prepositional variant is preferred more strongly by female speakers across all age groups (Tagliamonte 2014). The scarceness of research on the social dimensions of syntactic alternations and the potential correlations to be discovered highlights that this strand of research clearly merits further investigation. The present paper thus contributes to earlier work by analyzing the effect of gender in the choice of dative variant in another variety of English, namely acrolectal Jamaican English.

Acrolectal Jamaican English has only recently attracted more scholarly attention, mostly due to the completion of ICE-Jamaica (published 2008). Up until then, the main focus of research has been on Jamaican Creole, an English-lexified Creole that emerged through language contact

between the (mostly) African slaves and British settlers on the island, while Standard English in Jamaica was considered to be near-equivalent to British English, receiving only little attention (for exceptions, see, e.g., Miller 1987; Shields 1989; Irvine 2008; Deuber 2014; Irvine-Sobers 2018). With the publication of ICE-Jamaica, studies of linguistic variation in acrolectal Jamaican English have increased in numbers. These studies show that acrolectal Jamaican English contains lexical elements and grammatical rules that differentiate it consistently from the British English standard (see Section 2). However, none of these studies, at least to my knowledge, have looked at the extent to which this difference extends to the probabilistic factors driving speakers' choices of variable patterns. The aim of this chapter is thus to fill the gap in existing research by investigating the extent to which probabilistic constraints influence the choice between the double object and the prepositional dative variant in Jamaican English with a special focus on speaker's gender.

The paper is structured as follows: Section 2 provides an overview of the linguistic ecology in Jamaica. Section 3 discusses earlier work on the English dative alternation that has observed genderlectal differences in the choice of variant. Section 4 presents the case study that empirically tests the assumption of genderlectal differences in acrolectal Jamaican English. Section 5 situates the results and offers some final conclusions.

2 Linguistic variation and gender in Jamaica

Central to a description of linguistic variation in Jamaica is the concept of the creole continuum (DeCamp 1961; DeCamp 1971). This concept assumes that speakers continuously avail themselves of various linguistic forms from most basilectal Creole-like features to most acrolectal Standard Jamaican English features, moving fluidly between these two opposing poles along a continuum in everyday conversations. This concept has found empirical support in quantitative sociolinguistic studies that compared the frequency of basilectal and acrolectal linguistic variables across different conversation styles and that observed that the difference in basilectal and acrolectal forms between conversation styles was a matter of degrees rather than a categorical choice. Following this, Patrick (1999) proposes that the creole continuum should be considered a social rather than a linguistic phenomenon with fewer creole features in formal educated speech and more creole features in informal and less educated speech. Similarly, Winford (1997) states that linguistic variation in Jamaica is influenced by a complex pattern of social and situational factors where the boundaries between the creole and the acrolect are increasingly becoming

blurred. Opposing this view of a continuum, some scholars have argued for the existence of two separate linguistic systems – one creole and one English – referring to Ferguson’s concept of diglossia (Ferguson 1959). In such a diglossic situation, the creole is considered the ‘L’ or low variety used in informal conversations, at home and with friends, while English – Jamaica’s official language – constitutes the ‘H’ or high variety used by the media, the government, in court, etc. (see, e.g., Lawton 1980). These two perspectives on linguistic variation in Jamaica – creole continuum vs discreteness of two languages – have been mainly reconciled in Patrick (1999) and other, more recent, studies that adopt a sociolinguistic perspective on the continuum and acknowledge that Jamaican Creole and Jamaican English serve two functionally different purposes (see Deuber 2014: 10). Such a sociolinguistic perspective allows for a diglossic situation where Jamaican Creole and Jamaican English are used in socially different domains to varying degrees but pays due regard to the continuum to account for gradient linguistic variation. Viewing the linguistic ecology in Jamaica as a continuum rather than two separate systems makes further allowance for the fact that socially affluent, i.e. acrolectal, speakers seem to command a greater span of the continuum and thus display more linguistic variation than lower status speakers (Winford 1991; Patrick 1999).

The language contact situation in Jamaica is thus complex and very different from other Caribbean nations due to the existence of a robust basilectal creole (Schneider 2007: 227) whose use today is strongly associated with Jamaican national identity. For this very reason, the strict separation between domains of usage for the L and H variety has started to disintegrate in recent years. At the same time, we can now observe a re-orientation in the Jamaican speech community away from exonormative British English standards to local forms of speaking (Schneider’s phase 4, see Schneider 2007: 235). The use of creole-like features and lexical Jamaicanisms also in the speech of educated Jamaicans – as highlighted above – is just one of the symptoms. Increased acceptance of the creole and its association with Jamaican national identity has propagated Jamaican creole to the extent that it is now used in official and formal contexts to express solidarity and to associate with the local community – by politicians, on the radio, in newspaper writings, Jamaican literature and so on – also by acrolectal Jamaican speakers.

Studies analyzing the speech patterns of acrolectal speakers gained momentum in the early 2000s with the compilation and release of the Jamaican component of the International Corpus of English series (Rosenfelder et al. 2009; see also Sand 1999; Mair 2002; Jantos 2010). These early studies on the Jamaican acrolect tended to sample only the most standard variety, used by educated Jamaicans in formal conversations and (often written) contexts (see Deuber 2014: 16),

under the assumption that highly educated Jamaicans did not vary stylistically in written language and simply used Creole in spoken language (Deuber 2009: 3). The advent of new media forms (e.g. emails, internet discussion forums) revealed, however, that informal language use could also spread to the written genre (in unprecedented ways) indicating that even the highly educated acrolectal speakers might vary in their use of basilectal forms gradually between different styles. Analyzing spoken language, Deuber (2009), for instance, reports that informal conversations, as sampled in the ICE-Jamaica corpus, exhibit a higher rate of creole features than other speech styles even though the speakers are highly educated users of the acrolect. Moreover, formal salience of the creole feature impacts its frequency: The more salient the creole feature, the less often is it used in conversational language (Deuber 2009: 29–30). Similarly, Jantos (2010) reports a higher frequency of creole forms in the informal conversations in ICE with fewer creole forms in the more formal conversations (class lessons and broadcast news).

What about gender? Unfortunately, even though ICE-Jamaica has metadata on speakers' gender available, only few studies investigate or report gender effects. Similarly, studies that sampled their own data (for instance, in experimental settings or surveys) sought out a balanced gender distribution but then often ignored social factors in their analyses. Exceptions include Rosenfelder's (2009) dissertation, Irvin (2004; 2018) on phonological variation in the Jamaican acrolect, Patrick (1999) on urban mesolectal varieties, and Hinrichs and White-Sustaíta (2011) on spelling variation in computer-mediated communication. Rosenfelder (2009) tests the effect of social factors including age and speaker's gender on the realization of three phonological variables in acrolectal Jamaican English using data from spoken ICE-Jamaica. She observes that gender is only significantly influencing the choice of variant in some of the phonological variables, namely with regard to linking /r/ and glottalization in linking /r/ contexts, and in some non-high vowels, but not with regard to rhoticity. Irvin (2004) analyses variation of seven linguistic variables in the highly educated speech of staff working at a government agency and observes a significant influence of gender on the choice of some variants, notably h-dropping and diphthongization. Patrick (1999) reports relevant gender effects for two phonological variables, palatalization and (t,d)-deletion, and marginally also for preverbal past-marking. Finally, Hinrichs and White-Sustaíta (2011) observe gender effects in the proportion of non-standard spelling in emails and blogs. Their corpus results show that men use more non-standard spellings overall than female speakers even though the difference between the genders is not significant. A closer analysis of the spelling of four high frequency words subsequently contradicts the well-trodden notion that women use more standard forms than men. Their analysis shows that the difference in proportion of standard

vs non-standard spelling is more pronounced for men than for women highlighting that standard vs non-standard spelling is more equally distributed among female speakers. This discrepancy with previous sociolinguistic work aligns with the results in Rosenfelder (2009), who observes that female speakers of Jamaican English sometimes tend more towards the non-standard variant while male speakers prefer the standard variant. While Rosenfelder explains her findings with the possibility that the phonological variant traditionally considered 'standard' might not form part of Standard Jamaican English (Rosenfelder 2009: 112), Hinrichs and White-Sustaíta propose that the non-standard spelling variants enjoy more overt prestige which is why women tend to use them more than men (Hinrichs & White-Sustaíta 2011: 55–56). Common to all these studies is their focus on either phonological or spelling variation. Gender effects in lexical or (morpho-)syntactic variation seem to be either insubstantial or have just not been explored. The present study addresses this gap and contributes to our understanding of gender effects in acrolectal Jamaican English by extending the analysis to syntactic variation.

Further indications as to how gender might play out in linguistic variation in Jamaica can be obtained from the Jamaican Language Unit (JLU), an empirical survey conducted among 1,000 Jamaican informants. Results of this survey indicate that nearly half of the informants were able to conduct a conversation in both English and Creole. The rest considered themselves monolingual in Jamaican English or Jamaican Creole (The Jamaican Language Unit 2005: 5, 12). Bilingualism was higher among the professional and well-educated classes. With regard to gender, the JLU survey indicated that men were more likely than women to speak only Jamaican Creole while women were more likely than men to speak both languages (The Jamaican Language Unit 2005: 10). Hence, the effect of the Creole could potentially be stronger in female speakers' than in male speakers' acrolectal English due to the higher probability of female acrolectal speakers to also know Creole in contrast to male acrolectal speakers (see also Schneider 2007: 233 for a historical perspective on gender differences).

3 World Englishes, the dative alternation and gender

Starting off with Labov's work on rhoticity in New York City English (Labov 1966), studies in variationist sociolinguistics have abundantly shown that linguistic variation tends to correlate systematically with sociodemographic variables. While the majority of early research in this field was devoted to variation in phonological variables based on, for instance, speech style, social

class and gender (see, e.g., Labov 1972; Labov 2001), later studies extended their analysis to include lexical and morphosyntactic variation (see Sankoff 1973). There are crucial differences between the two: Variation between phonological variants does not entail difference in meaning, contrary to morphosyntactic or lexical variation where the variants, arguably, can express different functions and meanings depending on the context (see Tagliamonte 2012: 206). Research in morphosyntactic variation has thus often taken a keen interest in the language-internal factors that constrain the variation, mainly considering variation between a standard vs a non-standard variant where the choice between the two is highly socially marked. With respect to gender, for instance, results of these studies often concur that female speakers tend to use more standard and fewer stigmatized forms and orient themselves towards forms of overt prestige (e.g. Labov 1972: 243; Cheshire et al. 2002). In contrast, male speakers tend to opt for forms of covert prestige and non-standard or dialectal linguistic forms.

When it comes to patterns of variation between two (or more) standard forms, however, social factors were either ignored in the analyses or shown to have no significant effect (e.g. Weiner & Labov 1983; Jankowski & Tagliamonte 2014). Instead, language-internal factors take center stage. Among these studies we find research on variation in the use of relativizers (e.g. Hinrichs, Szmrecsanyi & Bohmann 2015), contractions (e.g. Barth & Kapatsinski 2014), complementizers (e.g. Tagliamonte & Smith 2005), future tense (e.g. Tagliamonte, Durham & Smith 2014), passive vs active voice (e.g. Seoane 2009), subject pronoun omission (e.g. Tamaredo 2018), the dative alternation (e.g. Röthlisberger, Grafmiller & Szmrecsanyi 2017), the genitive alternation (e.g. Heller, Szmrecsanyi & Grafmiller 2017), and particle placement (Grafmiller & Szmrecsanyi 2018). While these studies disregard speakers' sociodemographic background, other language-external factors are (sometimes) considered: Regional comparisons show, for instance, that language-internal constraints differ in their influence depending on locality (see, e.g., Bresnan & Hay 2008; Bresnan & Ford 2010; Röthlisberger, Grafmiller & Szmrecsanyi 2017). Recent stylistic comparisons further highlight that language-internal constraints impact linguistic variation differently depending on the register (e.g. Grafmiller 2014). At the same time, empirical evidence as to how the choice between two or more standard variants might depend on the sociodemographic background of the speaker remains rare.

The dative alternation is so far the only word-order alternation where researchers have taken an interest in speaker-related constraints on the variation – be it either in experimental setups or corpus-based approaches (as far as I am aware). One of the earliest studies in this line was Bresnan and Ford (2010) who observed in their experiment that male Australian speakers rated

the prepositional dative as more natural than female speakers in the same context. These findings were corroborated by Theijssen et al. (2011) who compared experimental and corpus data from Australian, British and American English. In their experiments, young Australian males tended to rate the prepositional dative variant as more natural than Australian female speakers and as more natural than older male speakers. At the same time, the authors observed a preference for the prepositional dative by young American speakers in a complementary corpus study. Taking a corpus-based approach, Jensen et al. (2018) report a similar gender difference for contemporary British English based on data from the newest BNC (BNC2014) as Theijssen et al. did in the experimental setups with Australian English speakers. They find that the speaker's gender influences dative choice significantly in spoken language with males using more prepositional datives than female speakers. The overall effect size, however, is comparatively small. Tagliamonte (2014), who compared spoken data from the UK and Canada, finds female speakers to overall prefer the prepositional dative across all age groups in both regions but gender as a predictor does not significantly influence dative choice in her modeling. The findings by Tagliamonte for Canadian English stand in opposition to all other varieties where genderlectal differences generally boil down to male speakers being more likely to use the prepositional dative variant than female speakers, i.e. in British English (Jensen, McGillivray & Rundell 2018), Australian English, American English (Theijssen et al. 2011) and New Zealand English (Bresnan & Hay 2008). Note, however, that these results might be not entirely comparable due to methodological differences between the studies as some of them restrict their data to the prototypical dative verb *give*, to specific registers or they use experimental setups vs corpus-based approaches. More large-scale comparative work is needed to shed light on genderlectal differences in the dative alternation from a global perspective. In the meantime, the present work continues the current enterprise by contributing one further case study of the dative alternation in one variety specifically.

4 Making syntactic choices in Jamaican English

In what follows, I will present a case study of the English dative alternation in acrolectal Jamaican English. Using conditional random forests and logistic regression analysis I will address the aim of the present chapter, namely how probabilistic constraints influence the choice of dative variant in Jamaican English with a special focus on speakers' gender.

4.1 Extracting and filtering dative tokens from ICE-Jamaica

The data for the present study comes from ICE-Jamaica (Rosenfelder et al. 2009) and forms part of a larger dataset used for a project that explored variation in probabilistic constraints across some nine national varieties of English (see Röthlisberger, Grafmiller & Szmrecsanyi 2017; Röthlisberger 2018). Compilation of the Jamaican component of the International Corpus of English series started in the 1990s at the Universities of Freiburg and of the West Indies (Mona) and the corpus was finally released in 2008 (see Rosenfelder et al. 2009 for details). ICE-Jamaica shares its corpus design with the other corpora of the ICE series: It comprises 500 texts of approximately 2000 words each totaling 1 million words, 60% of which come from spoken data and 40% from written data.

The dataset of variable dative tokens was retrieved following the procedure outlined in Röthlisberger et al. (2017) on the basis of a verb list which was composed of all possibly interchangeable dative verbs provided by earlier studies (e.g. Bresnan et al. 2007; Wolk et al. 2013; De Cuypere & Verbeke 2013), including those dative verbs that are traditionally thought to be restricted to one variant only (e.g. *donate*). Such a broader approach allows for the possibility that some verbs might alternate in non-British varieties. A given verb was considered alternating if it occurred at least five times in the alternating dative variant (double object or prepositional variant) in the nine ICE corpora studied or in an independent dataset (the Corpus of Global Web-based English [GloWbE]). The final list of verbs is provided in the appendix.

Next, prepositional and double object variants were retrieved from the part-of-speech-tagged version of the ICE corpus and the respective variety in GloWbE using a perl script. Through intensive manual coding, the dataset was then restricted to alternating tokens that were semantically similar and grammatically acceptable, excluding variants that involved particle verbs, coordinated verbs, elliptic structures, clausal or non-overt constituents, beneficiary constructions, constructions involving a spatial goal and idioms or fixed expressions (see Röthlisberger 2018: 55–58). The dataset was further delimited to exclude double object variants with extremely long themes (> 23 words) and prepositional datives with extremely long recipients (> 18 words) in order to dismiss those instances where the likelihood became increasingly zero of the other variant occurring. The final dataset covering nine national varieties of English contained 13,171 variable dative tokens that were then further annotated for numerous language-internal and language-external factors. While the majority of these factors had already been considered in the literature, hitherto neglected factors were also added to the annotation procedure (e.g.

themacity, type token ratio). Earlier work using that dataset (Szmrecsanyi et al. 2016; e.g. Röthlisberger, Grafmiller & Szmrecsanyi 2017; Röthlisberger 2018) adopted primarily a bird’s-eye perspective focusing on large-scale patterns across regions with some attention also paid to the influence of lexical considerations and register differences. Sociodemographic patterns were largely ignored and only mentioned in passing, mostly due to limits of space and the scope of the papers. However, since sociodemographic information is available for all nine ICE corpora, an analysis of the influence of social factors is easily attainable, particularly for the spoken parts where metadata is more substantial.

The data analyzed here comprises a small subset of the total dataset of nine varieties. Attention was restricted to Jamaican English and tokens from the spoken ICE-component, totaling 615 variable dative tokens. A consistent format of the metadata was generously shared by Martin Schweinberger via his website². The spoken component was selected not only for methodological reasons (availability of metadata) but also for theoretical reasons: As Deuber (2009) and Jantos (2010), among others, have pointed out, it is particularly in spoken language where linguistic variation between basilectal and acrolectal forms in Jamaica persist. And as results of the Jamaican Language Unit have shown, female and male speakers avail themselves differently of basilectal and acrolectal forms, making spoken Jamaican English a hotbed for an analysis of the effects of sociodemographic factors such as gender. The data thus samples from most informal speech styles (private conversations) to most formal speech styles (broadcast news) as illustrated in Table 1.

| Text category | Level of formality (REGISTER) | ICE-Jamaica label | number of variable tokens |
|---|--------------------------------------|---------------------------------------|----------------------------------|
| Private & distanced conversations | spoken informal | Private dialogues & phone calls (S1A) | 199 |
| Classroom lessons, Broadcast discussions & interviews, Parliamentary debates, legal cross-examinations, business transactions | spoken formal | Public dialogues (S1B) | 188 |
| Spontaneous commentaries, unscripted speeches, demonstrations, legal presentations | spoken informal | Unscripted monologues (S2A) | 79 |

² <http://www.martinschweinberger.de/blog/resources/>

| | | | |
|------------------------|---------------|---------------------------|-----|
| Broadcast news & talks | spoken formal | Scripted monologues (S2B) | 149 |
|------------------------|---------------|---------------------------|-----|

Table 1. Text categories and registers included in the dataset

4.2 Annotation of predictor variables

The full dataset covering nine national varieties of English had been annotated for roughly 20 different predictor variables. For the spoken component of ICE-Jamaica, the number of predictors was decreased due to the fact that some predictors (e.g. VARIETY or MODE) are completely monolithic in this subset. 11 language-internal and 3 language-external factors were retained for the analysis.

4.2.1 Language-internal predictors

- WEIGHTRATIO calculates the length of the recipient in the number of characters divided by the length of the theme in the number of characters and log-transforms this value following previous approaches (e.g. Bresnan et al. 2007).
- RECPRON/THEMEPRON accounts for recipient and theme pronominality and distinguishes between pronominal ('pron' incl. personal and impersonal pronouns) recipients/themes and non-pronominal ('non-pron') ones.
- RECGIVENNESS/THEMEGIVENNESS was coded as a binary predictor: If the recipient/theme occurred in the 100 preceding words or was a pronoun, it was coded as 'given'. All other constituents were coded as 'new'.
- RECANIMACY is a binary predictor that assesses whether the recipient is 'animate' (human, animal) or 'inanimate' following previous approaches (e.g. Wolk et al. 2013). (Note that the predictor THEMEANIMACY is excluded from this analysis because hardly any animate themes occur in the dataset)
- RECDEFINITENESS/THEMEDEFINITENESS accounts for definiteness of the constituents. This predictor was coded following the procedure outlined in Garretson et al. (2004): Any instance that allowed for an existential reading (opposed to a deictic reading) in the context of "There is/are ___" was considered indefinite ('indef'), all other instances were coded as definite ('def'). Additionally, constituents that started with a word marked as indefinite according to Garretson et al. were coded as 'indef' (e.g. *a, an, another, any, enough*).

Constituents that contained a proper noun or pronoun as their head or started with a definite article, demonstrative or any word tagged as definite by Garretson et al. were coded as 'def' (e.g. *the, this, those, that, her, his, its, my*)

- RECCOMPLEXITY/THEMECOMPLEXITY accounts for the absence/presence of post-head elements (i.e. postmodification) and distinguishes between recipients/themes that were not followed by a post-head element ('simple') and those that were followed by a post-head element ('complex').
- PRIMETYPE captures the effect of structural persistence which refers to speakers' tendency to reuse syntactic constructions that they have uttered or heard previously (see Branigan, Pickering & Cleland 2000; Gries & Wulff 2005; Szmrecsanyi 2006). PRIMETYPE was either a previous prepositional dative ('pd'), a previous double object variant ('do') or 'none' if neither of the two variants occurred in the previous discourse available in the corpus text file.
- VERBSEMANTICS distinguishes between five broad semantic classes that each dative verb can fall into depending on the context it occurs in (e.g. Bresnan & Hay 2008). Each dative token was thus classified as denoting an act of transfer (e.g. *They give everybody a piece of paper*; 't'), an act of communication (e.g. *She told me the whole story*; 'c'), an act of prevention of transfer (e.g. *They denied him entry to the country*; 'p'), and acts of future transfer (e.g. *Carl had promised her this car*; 'f'); all other dative instances were labelled as abstract (e.g. *You are paying me attention*; 'a')
- RECHEADFREQ/THEMEHEADFREQ accounts for the global lexical frequency of the recipient or theme head; lemma frequencies were retrieved from the respective complete component of the GloWbE corpus that represents the variety in question. This global frequency was then normalized as count per million words in the given variety in GloWbE.
- RECTHEMATICITY/THEMETHEMATICITY assesses the frequency of the recipient/theme in the text in which the dative token occurs and thus reflects the extent to which a constituent forms part of the central topic of a text. Thematicity is measured as the normalized text frequency of the head noun in the entire text in which the token occurs divided by the total number of words in the text (Hinrichs & Szmrecsanyi 2007: 450–451).
- TYPETOKENRATIO captures lexical density of the surrounding context of the dative token (~ 50 words) and is defined as the number of unique types in this 50-word window divided by the number of tokens. This predictor has been reported as influential in the genitive alternation (see Hinrichs & Szmrecsanyi 2007: 457) but the effect in the dative alternation is rather minimal, as illustrated below.

4.2.2 Language-external predictors

- GENDER: a binary distinction whether the token in question was produced by a female or male speaker. This information is available from the metadata accompanying the corpus.
- AGE: A three-way distinction between speakers aged around 20 ('18-24'), speakers in their late twenties and thirties ('25-45'), and speakers older than 46 years ('46+').
- REGISTER: Additionally, the spoken data was divided as either informal ('SpokInf') or formal speech ('SpokForm') on the basis of the 12-genre distinction in the ICE corpus structure and using the model proposed in Koch and Oesterreicher (1985) as a guideline (see also Table 1). Spoken formal includes public dialogues and scripted monologues while spoken informal covers private dialogues and unscripted monologues.

The distribution of the predictor GENDER by dative variant is shown in Figure 1. As illustrated there, female speakers use the double object variant more often than male speakers when it comes to making a choice between the two dative variants. All other distributional patterns are not shown here for reasons of space. The distributional patterns of the language-internal predictors highlight, however, that patterns of variation in the Jamaican dative alternation follows the global patterns observed in earlier studies (see, e.g., Bresnan et al. 2007; Röthlisberger, Grafmiller & Szmrecsanyi 2017). Overall, speakers tend to prefer that dative variant where the first constituent is animate, given, definite, short(er), more frequent, thematic, and simple compared to the other constituent. Also, while the double object variant is overall the most frequent variant, the prepositional dative becomes more likely if preceded by another variable prepositional dative. Double object variants become more frequent with increased lexical density and are preferred more by younger age groups compared to older age groups. Finally, double object variants are more frequent in spoken informal speech than in spoken formal speech.

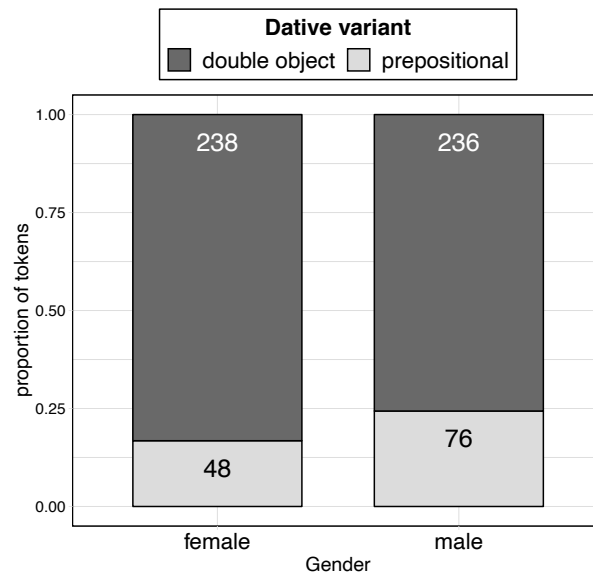


Figure 1. Distribution of double-object and prepositional variants by GENDER in spoken acrolectal Jamaican English

4.3 Statistical analyses

In order to address its aim, the present study makes use of conditional random forests in a first step, and mixed-effects logistic regression in a second step in a complementary fashion. Statistical analyses were carried out in R (R Core Team 2017) using the `cforest()` function in the `party` package (Hothorn et al. 2006; Strobl et al. 2008; Strobl, Hothorn & Zeileis 2009) for the forest and the `glmer()` function in the `lme4` package (Bates et al. 2015) to fit the mixed-effects model.

Conditional random forests (cf. Breiman 2001) are well suited to investigate the importance of probabilistic constraints vis-à-vis other constraints. In contrast to mixed-effects models that make predictions on the basis of a mathematical equation, random forests establish the usefulness of a predictor through trial and error by aggregating over numerous conditional inference trees using random subsampling and a permutation scheme (see Strobl et al. 2008 for details). While random forests are quite robust to common issues in linguistic analyses, such as data sparseness or predictor non-linearities (see also Tagliamonte & Baayen 2012: 158–161), they cannot deal very well with categorical predictors that have very large numbers of levels. Mixed-effects models on the other hand enable us to assess the direction of predictors in the model (in addition to strength and predictive importance). They allow us to account for the non-independence of our

observations via random effect adjustments (e.g. lexical items or speakers sampled). Mixed-effects models can thus provide more reliable generalizations beyond the specific dative tokens sampled in the dataset.

Due to the rather low number of tokens and the large number of language-internal predictors coded for, results from a conditional random forest fitted only with language-internal predictors were used to identify the most important language-internal predictors which were subsequently fed into the logistic regression as fixed effects.

4.4 Results

4.4.1 Variable importance of predictors

The conditional random forest fitted to the dataset included all predictors listed in Section 4.2. After tuning the hyperparameters of the model with the caret package (Kuhn et al. 2016), 3000 trees (*ntree*) were fitted with seven predictors (*mtry*) considered at each split. Variable importance of the predictors was then calculated with the `varimpAUC()` function following Janitza et al. (2013). Somer's *C* index for the final forest is an outstanding 0.985 indicating that the model is very well able to discriminate between the variants of the outcome (Levshina 2015: 259). Predictive accuracy of the model is 94.47% which is significantly better than the baseline of 79.19% when always selecting the most frequent variant ($p_{\text{binom}} < 0.0001$). The variable importance calculated for each predictor is shown in Figure 2.

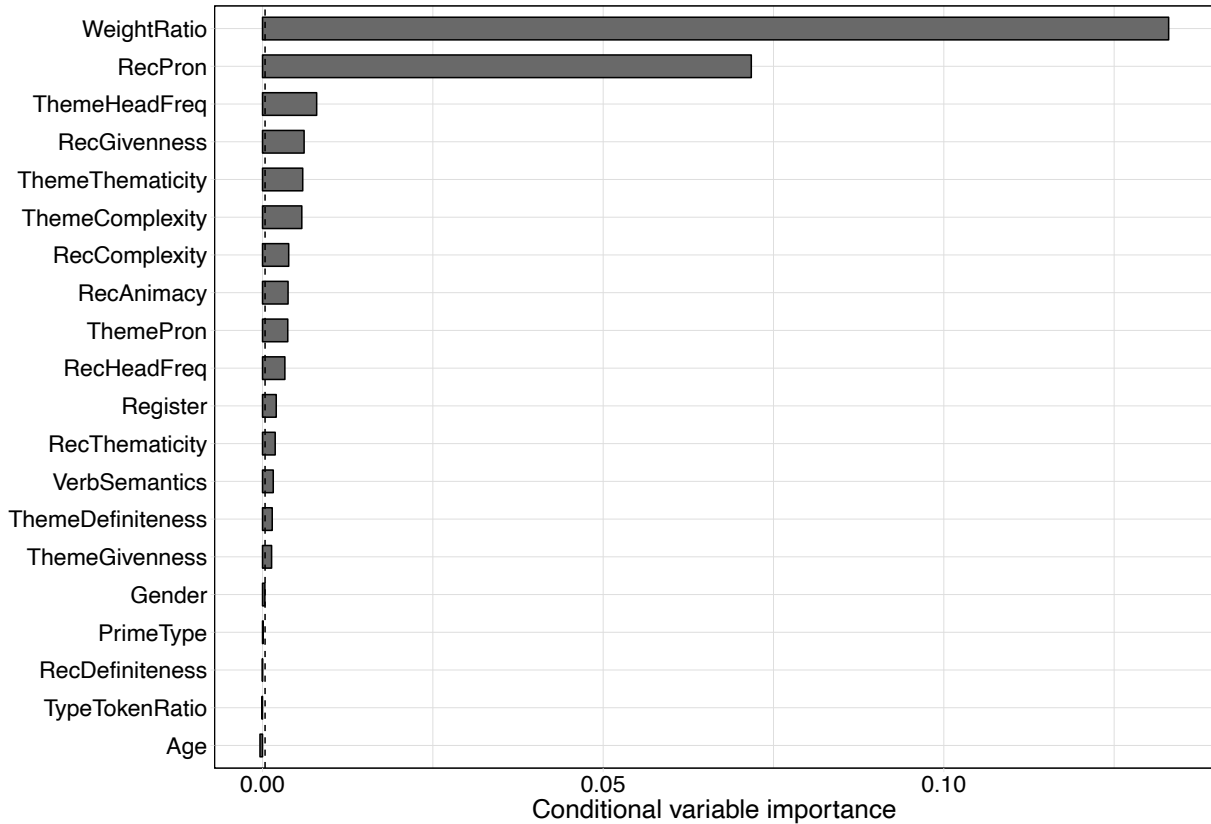


Figure 2. Conditional variable importance of language-internal and language-external predictors influencing the dative alternation in acrolectal spoken Jamaican English

Not surprisingly, the language-internal factors outweigh sociodemographic factors in importance, a finding repeatedly encountered in variationist research (see Szmrecsanyi et al. 2016). Weight ratio turns out to be the most important factor, followed by recipient pronominality. JamE thus closely falls in line with the overall trend observed in Röthlisberger et al. (2017) who report weight ratio and recipient pronominality to be the two most important factors across nine national varieties of English, with the exception of Indian English, where the order of importance of these two factors is reversed.

Note that the predictor REGISTER features as the most important language-external predictor in the conditional random forest, which is also consistent with previous studies that have observed that register or genre differences can play an important role in syntactic alternations (e.g. Grafmiller 2014; Szmrecsanyi et al. 2016).

4.4.2 Effect directions of predictors

Because weight ratio and recipient pronominality feature so importantly in the dative alternation in Jamaican English, only those two language-internal predictors were fed into the mixed-effects logistic regression analysis. Due to this study's interest in gender effects, 17 tokens had to be excluded from this subsequent analysis as the metadata did not provide information on speaker's gender for those specific tokens. The dataset for the mixed-effects model included 598 variable datives.

Two mixed-effects models were fitted: For once, I was interested in the extent to which the social predictors by themselves played a role in dative choice. The model thus included the two language-internal predictors, gender, age and register as fixed effects and speaker (individual) and recipient head, theme head and verb as random intercepts. Infrequent lexical effects were grouped together (recipient heads occurring less than 3 times, theme heads occurring less than 4 times, speakers contributing less than 4 tokens). Binary predictors (recipient pronominality, gender and register) were transformed into numbers and centered around the mean to reduce multicollinearity and ease model convergence. Model selection followed the top-down approach outlined by Zuur et al. (2009: 120–122). First, I fitted a maximal model including all fixed effects and random effects and then excluded those random effects that did not significantly improve model fit according to likelihood ratio tests. All fixed effects were, however, included in the final model irrespective of statistical significance unless model convergence and evaluation indicated that their removal from the model structure was necessary. The final model included the language-internal predictors RECPRON and WEIGHTRATIO (log-scaled), GENDER, REGISTER, and random intercepts for verb and theme head. The results of the model are shown in Table 2. Model predictions are for the prepositional dative. Predictors are given with the reference level to the left of the arrow and the predicted level to the right; coefficient estimates are given in the column labelled β : positive values indicate a preference for the prepositional dative, negative values indicate a preference for the double object dative. SE specifies standard error. Summary statistics indicate that the model fits the data well. The model can predict 96.15% of the data accurately which is significantly better than the baseline of 79.26% ($p_{\text{binom}} < 0.001$). C-index is an outstanding 0.990. Collinearity was assessed with the condition number κ (following Belsley, Kuh & Welsch 1980). Condition number κ equals 2.36 indicating no collinearity to speak of (see Baayen 2008: 182). Model evaluation indicated no overdispersion.

| Predictor | β | SE | p |
|--|----------|----------------|---------|
| (Intercept) | -2.5344 | 1.1027 | 0.2154 |
| weight ratio | 7.8082 | 1.4539 | <0.001 |
| recipient pronominality pron \Rightarrow non-pron | 2.2379 | 0.8257 | 0.00672 |
| gender female \Rightarrow male | 1.1772 | 0.5547 | 0.03381 |
| register informal \Rightarrow formal | 1.0452 | 0.5482 | 0.05657 |
| Random effects | Variance | Std. deviation | |
| Verb | 10.42 | 3.228 | |
| Theme | 5.53 | 2.352 | |

Table 2. Coefficient estimates of main effects and variance and standard deviations of random effects in the model – Model predictions are for the prepositional dative.

Visualizations of the effects of the four main effects were generated using the effects() package (Fox 2003) and are shown in Figure 3. Effects show that the prepositional dative becomes more likely when weight ratio increases, that is, when the recipient increases in length in comparison to the theme. Thus, the model predicts a higher probability for (3b) which has a higher weight ratio than for (3a) where recipient and theme are nearly equally long. The prepositional dative is also more likely if the recipient is non-pronominal as in (4b) compared to (4a). The trends of these two language-internal factors again follow the general consensus in the literature in that speakers opt for the variant where they can express the shorter and more accessible (i.e. pronominal) constituent first. Regarding social factors, the prepositional dative is also more likely when the speaker is male (example 5a) instead of female (5b) and when the dative token is used in spoken formal conversations (6a) rather than informal conversations (6b), which is not a significant effect in the model but confirms the proportional distributions in the data.

- (3) a. He is certain that he *gave the names to the police* [...] (ICE-JA:S2A-063:79)
b. I am going to suggest to you that is several days after the incident you *gave the name Scooter to the police officer Mr Bottom* (ICE-JA:S1B-069:233)
- (4) a. I have to *send them their itinerary* (ICE-JA:S1A-043:10)
b. Secondly I think this would also *send a direct message to the leaders of the political parties* [...] (ICE-JA:S1B-039:90)
- (5) a. Let us *give support to the security forces* to act professionally (ICE-JA:S1B-059:55, male)

- b. It looks at the way people living in particular locations *give those locations identity* [...] (ICE-JA:S2B-039:25, female)
- (6) a. [...] if they saw it fit to up *give the award to people* for breaking news (ICE-JA:S1B-046:81)
- b. [...] you don't want to *give people a false hope*. (ICE-JA:S2A-050:53)

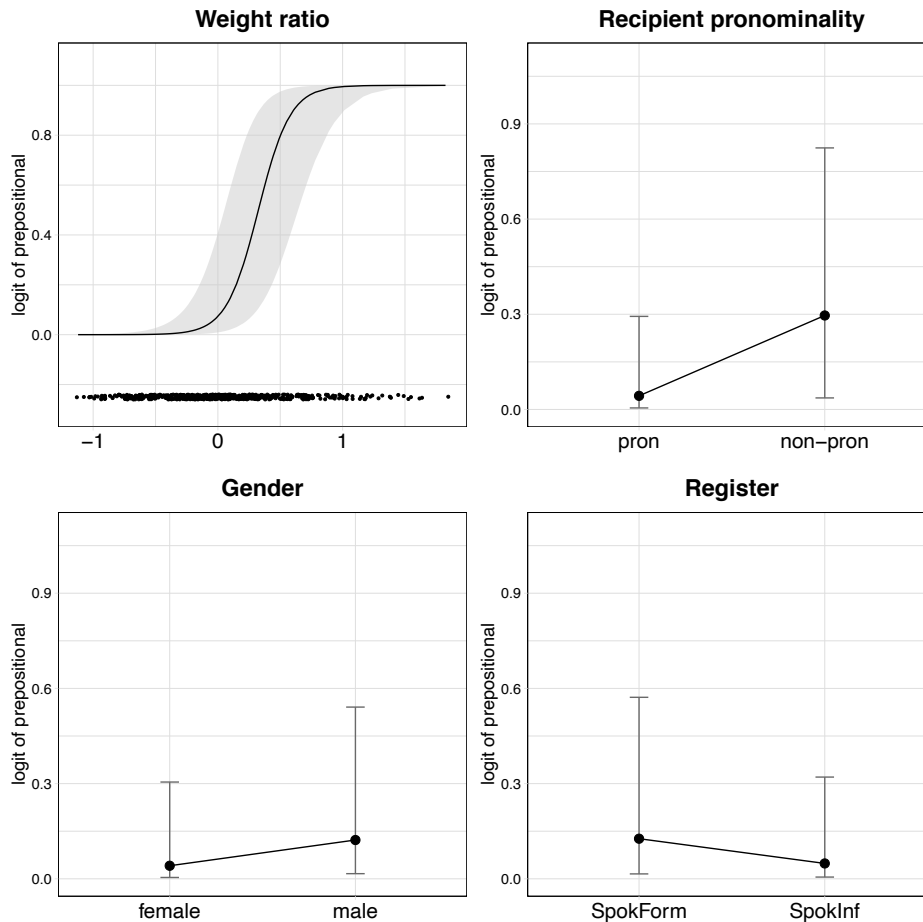


Figure 3. Effects plots for the four predictors *WEIGHTRATIO*, *RECPRON*, *GENDER* and *REGISTER* included in the final model with confidence intervals.

A second regression then aimed to test the assumption, observed in earlier work (e.g. Bresnan & Hay 2008; Bresnan & Ford 2010; Heller, Szmrecsanyi & Grafmiller 2017; Röthlisberger, Grafmiller & Szmrecsanyi 2017), that language-internal predictors can vary across different social dimensions. Hence, the second model tested interaction effects between the two language-internal predictors and *GENDER* and *REGISTER*. Due to convergence problems with the random effects, only random intercepts for the verb and the theme were included; a fixed effect for *AGE*

had to be excluded due to the evaluation procedure described in the first model. The final model showed that none of the interaction terms added to the model significantly, i.e. even though the effect of some language-internal predictors might vary across gender and register, the difference is very likely due to chance.

5 Discussion and conclusion

The present chapter has investigated the effect of gender in acrolectal Jamaican English on the choice between two syntactic variants, the double object vs prepositional dative. 615 variable dative tokens were sampled from the spoken component of ICE-Jamaica and analyzed with regard to the effect of 11 language-internal and three language-external predictors. Two methods were used in a complementary fashion: conditional random forests and mixed-effects logistic regression. The results of the conditional random forest showed that language-internal constraints outweigh language-external factors in their importance in dative choice. REGISTER (or formality) was found to be the most important language-external factor vis-à-vis gender and age of the speaker. The subsequent mixed-effects regression further illustrated that the speaker's gender influences variation between the double object and prepositional dative significantly. The influence of this sociodemographic predictor, however, remained rather marginal (as could also be discerned from the effects plot and the coefficient estimate). An additional mixed-effects regression testing for interaction effects indicated that none of the language-internal predictors differed in their effect size significantly by the predictors GENDER and REGISTER.

The findings of the present work are insofar consistent with earlier studies as the effects of the two language-internal predictors fitted in the logistic regression, WEIGHTRATIO and RECPRON, align with previous observations: Speakers tend to prefer that dative variant in which the first constituent is pronominal and/or short. The proportional distributions of the other language-internal predictors also confirm earlier work in that regard. Additionally, recipient pronominality and weight ratio are again the two most important predictors when it comes to dative choice in Jamaican English, which is also consonant with earlier work (see Röthlisberger, Grafmiller & Szmrecsanyi 2017). With regard to the language-external predictors – REGISTER, GENDER and AGE – REGISTER turned out to be more important than the other two predictors in the random forest, which would be expected given the literature: For one, sociodemographic predictors are consistently shown to be least important if fitted in a model together with language-internal predictors, while register often turns out to be an influential predictor (see, e.g., Szmrecsanyi et

al. 2016: 119). Note, however, that the effect of REGISTER is not significant in the mixed-effects model while GENDER is. Zooming in on the sociodemographic predictor GENDER (since AGE could not be considered in the logistic regression), the results of the present study point to females using fewer prepositional variants than male speakers – a pattern that has also been observed for Australian English and British English (Theijssen et al. 2011). At the same time, Tagliamonte (2014) observed that male speakers use more double object datives than female speakers across all age groups in Canadian English in her distributional analysis, but no genderlectal differences emerged in her regression modeling. This potential inconsistency among varieties of English with regard to the effect that GENDER has on dative choice is consonant with other fine-grained differences in the probabilistic domain observed in earlier macro-social studies (e.g. Bresnan & Hay 2008; Bresnan & Ford 2010; Wolk et al. 2013; Heller, Szmrecsanyi & Grafmiller 2017; Röthlisberger, Grafmiller & Szmrecsanyi 2017). These studies highlight gradient differences between regional varieties of English with regard to the probabilistic constraints influencing syntactic variation. While the present work could not detect any statistically significant differences between the two genders with regard to the two language-internal predictors, weight ratio and recipient pronominality, fitted in the logistic regression, the fact that female and male speakers might make different dative choices given the otherwise same set of predictors points to subtle differences in the linguistic repertoire of these two speech groups.

This line of reasoning fits in neatly with what these earlier regional studies have observed. Even though probabilistic grammars seem to be quite robust on a global scale, discernable differences can be detected on a local scale (see Szmrecsanyi et al. 2016; Heller, Szmrecsanyi & Grafmiller 2017; Röthlisberger, Grafmiller & Szmrecsanyi 2017; Grafmiller & Szmrecsanyi 2018). From a usage-based perspective, such local acculturation of probabilistic constraints seems self-evident: To the extent that speakers fluidly construct and adapt their probabilistic grammar based on the linguistic input that they receive – as usage-based approaches propose (e.g. Bybee 2010) – we can assume to find subtle probabilistic differences between speech communities or speech groups. These gradient localized acculturations of probabilistic constraints in the grammar of speakers from different communities has been termed *probabilistic indigenization* (Szmrecsanyi et al. 2016; see also Röthlisberger, Grafmiller & Szmrecsanyi 2017). Probabilistic indigenization is thus a by-product of the eternal fluid adaptation of one's probabilistic grammar through day-to-day interactions. If we assume that community-specific social forces, such as language attitudes or stylistic preferences, can shape biases in individual speakers' production and comprehension (as argued in, e.g., Röthlisberger, Grafmiller & Szmrecsanyi 2017; Grafmiller et al. 2018), we should find innovation and greater variability between groups where these forces exert divergent

influence. Against the background of the present study, it seems then that for male and female speakers to differ in their choice of variant, they are highly likely to also differ with regard to social context of language usage (e.g. stylistic choices, attitudes). It thus might be the case that the different social groups in these localities adapt linguistically to diverging social contexts or are exposed to diverging linguistic input.

But why should female and male speakers of acrolectal Jamaican English differ in linguistic exposure (as speaker groups in other localities do too, e.g. Canada, UK, Australia)? While I would like to remain cautious in making any definite claims about probabilistic differences in Jamaican genderlects on the basis of just one syntactic variable and only a few hundred tokens, the present study's results suggest that male and female speakers might avail themselves differently of the linguistic repertoire available to them in the complex linguistic ecology of Jamaica.

One obvious difference concerns the degree of (mostly intra-speaker) contact with Jamaican Creole which is presumably higher with female acrolectal speakers than with male speakers (as suggested in The Jamaican Language Unit 2005; but see The Jamaican Language Unit 2007: 13). As Farquharson (2013) points out, the only available pattern to order recipient and theme in ditransitive constructions in Jamaican creole is by using the double object variant as in (7). Speakers of Jamaican Creole can use double object variants with verbs of physical transfer of possession without any additional grammatical marking on the recipient contrary to Standard English where speakers can also explicitly mark it using *to* in the prepositional variant (see also Bruyn, Muysken & Verrips 1999: 330).

(7) Jamaican Creole

Di uman gi di bwai di fuud.

DET woman give DET boy.RECIPIENT DET food.THEME.

'The woman gave the boy the food.' (Farquharson 2013)

A usage-based perspective is concordant with the view that speakers in language contact situations not only transfer lexical elements or grammatical structures from one language to the other, but also probabilistic constraints that they have derived from linguistic input in one language (MacWhinney 1997). Transfer effects of probabilistic constraints can thus lead to differences between speaker groups if they differ in the degree of language contact. Data obtained from the Jamaican Language Unit (2005) shows that Jamaican women are more likely than male speakers to use both Jamaican Creole and acrolectal Jamaican English in their everyday conversations depending on style while Jamaican men tend more towards only using Creole. This greater female

likelihood of JamE/JamC bilingualism potentially also holds when only speakers of JamE are considered. Female Jamaican speakers are hence more likely to shift between Creole and English and might conversely also be more likely to display transfer effects of Jamaican Creole – where the double object variant is prevalent – to acrolectal Jamaican English. While a possible influence from Jamaican Creole could thus be responsible for the gender effect observed in the present analyses, other more general restructuring processes as outlined in Röthlisberger et al. (2017) or Sand (2005) need to be considered as well (see Deuber 2009: 47). While the present study has certainly tried to investigate the effect of speakers' gender in dative choice, more in-depth analysis and experimental methods are needed to explore the extent to which genderlectal differences emerge in syntactic variation, not only in Jamaica but also elsewhere. Especially comparing the effect of probabilistic constraints in this study to a similar study with data sampled from the basilectal end of the creole continuum would be revealing about the extent to which transfer effects from Jamaican Creole can account for the gender differences observed in the current dataset. Such a comparative approach would enable us to explore more fully the effect of probabilistic constraints in Jamaican Creole and the extent to which these are mirrored in acrolectal Jamaican English by speaker groups and speech styles.

From the perspective of World Englishes, the present work not only advances our understanding of the effect of genderlects in syntactic variation, it also aligns with recent proposals regarding the advancement of models of World Englishes (e.g. Schneider 2007; Mair 2013) by highlighting the importance of an encompassing analysis of the linguistic ecology of a variety of English. These recent proposals (see the contributions in Deshors 2018) stress the importance of intravarietal heterogeneity when modeling a variety of English. In other words, a variety of English should not be simply regarded as a holistic whole, but any theoretical perspective should allow for variety-internal variability. Acknowledging variation within the variety would move the focus away from a traditional view of World Englishes that aims to categorize varieties vis-à-vis other varieties based on their structural make-up but would make room for the diversity to be found within individual speech communities. Such a perspective does not only take the complex language contact situation in Jamaica into account but allows for further diversification within the creole continuum along the gender dimension, paving the way for Jamaican English to move to phase 5 in Schneider's model and to exert regional dominant influence beyond its national borders (see Mair 2013: 263).

References

- Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barth, Danielle & Vsevolod Kapatsinski. 2014. A multimodel inference approach to categorical variant choice : construction , priming and frequency effects on the choice between full and contracted forms of am, are and is. *Corpus Linguistics and Linguistic Theory*. doi:10.1515/cllt-2014-0022.
- Bates, Douglas, Martin Mächler, Benjamin M. Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1--48.
- Belsley, David A., Edwin Kuh & Roy E. Welsch. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley and Sons.
- Bernaisch, Tobias, Stefan Th. Gries & Joybrato Mukherjee. 2014. The dative alternation in South Asian English(es): Modelling predictors and predicting prototypes. *English World-Wide* 35(1). 7–31.
- Branigan, Holly P., Martin J. Pickering & Alexandra Cleland. 2000. Syntactic co-ordination in dialogue. *Cognition* 75(2). B13-25.
- Breiman, Leo. 2001. Random forests. *Machine Learning* 45. 5–32.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Boume, Irene Kraemer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1). 168–213.
- Bresnan, Joan & Jennifer Hay. 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118(2). 245–259.
- Bruyn, Adrienne, Pieter Muysken & Maaïke Verrips. 1999. Double-object constructions in the creole languages: Development and acquisition. In Michel DeGraff (ed.), *Language creation and language change: Creolization, diachrony and development*, 329–373. Cambridge, MA: MIT Press.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Cheshire, Jenny. 2003. Social dimensions of syntactic variation: The case of *when* clauses. In David Britain & Jenny Cheshire (eds.), *IMPACT: Studies in Language and Society*, vol. 16, 245–261. Amsterdam: John Benjamins Publishing Company.
- Cheshire, Jenny, J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes. 2002. Sex and gender in variationist research. *The handbook of language variation and change*, 423–443. Malden and Oxford: Blackwell.
- De Cuypere, Ludovic & Saartje Verbeke. 2013. Dative alternation in Indian English: A corpus-based analysis. *World Englishes* 32(2). 169–184.
- DeCamp, David. 1961. Social and geographical factors in Jamaican dialects. In Robert B. Le Page (ed.), *Creole language studies*, vol. 2, 61–84. London: Macmillan.
- DeCamp, David. 1971. Toward a generative analysis of a post-creole speech continuum. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 349–370. Cambridge: Cambridge University Press.
- Deshors, Sandra C. (ed.). 2018. *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties* (Varieties of English Around the World). Vol. G61. Amsterdam: John Benjamins Publishing Company.

- Deuber, Dagmar. 2009. 'The English we speaking': Morphological and syntactic variation in educated Jamaican speech. *Journal of Pidgin and Creole Languages* 24(1). 1–52.
- Deuber, Dagmar. 2014. *English in the Caribbean: Variation, style and standards in Jamaica and Trinidad*. Cambridge: Cambridge University Press.
- Farquharson, Joseph T. 2013. Jamaican structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/8> (20 January, 2017).
- Ferguson, Charles A. 1959. Diglossia. *WORD* 15(2). 325–340. doi:10.1080/00437956.1959.11659702.
- Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15). 1–27.
- Garretson, Gregory, M. Catherine O'Connor, Barbora Skarabela & Marjorie Hogan. 2004. Coding practices used in the project 'Optimality Typology of Determiner Phrases'. MS. <http://npcorpus.edu/documentation/index.html>.
- Grafmiller, Jason. 2014. Variation in English genitives across modality and genres. *English Language and Linguistics* 18(3). 471–496.
- Grafmiller, Jason & Benedikt Szmrecsanyi. 2018. Mapping out particle placement in varieties of English: A study in comparative sociolinguistic analysis. *Language Variation and Change* 30(3). 385–412.
- Grafmiller, Jason, Benedikt Szmrecsanyi, Melanie Röthlisberger & Benedikt Heller. 2018. General introduction: A comparative perspective on probabilistic variation in grammar. *Glossa: a journal of general linguistics* 3(1). 94. 1–10.
- Gries, Stefan Th. & Stefanie Wulff. 2005. Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics* 3. 182–200.
- Heller, Benedikt, Benedikt Szmrecsanyi & Jason Grafmiller. 2017. Stability and fluidity in syntactic variation world-wide: The genitive alternation across varieties of English. *Journal of English Linguistics* 45(1). 3–27.
- Hinrichs, Lars & Benedikt Szmrecsanyi. 2007. Recent changes in the function and frequency of Standard English genitive constructions: A multivariate analysis of tagged corpora. *English Language and Linguistics* 11(3). 437–474.
- Hinrichs, Lars, Benedikt Szmrecsanyi & Axel Bohmann. 2015. Which-hunting and the Standard English relative clause. *Language* 91(4). 806–836.
- Hinrichs, Lars & Jessica White-Sustaíta. 2011. Global Englishes and the sociolinguistics of spelling: A study of Jamaican blog and email writing. *English World-Wide* 32(1). 46–73.
- Hothorn, Torsten, Peter Buehlmann, Sandrine Dudoit, Annette Molinaro & Mark Van Der Laan. 2006. Survival ensembles. *Biostatistics* 7(3). 355–373.
- Irvine, G. Alison. 2004. A good command of the English language: Phonological variation in the Jamaican acrolect. *Journal of Pidgin and Creole Languages* 19(1). 41–76.
- Irvine, G. Alison. 2008. Contrast and convergence in Standard Jamaican English: The phonological architecture of the standard in an ideologically bidialectal community. *World Englishes* 27. 9–25.
- Irvine-Sobers, G. Alison. 2018. *The acrolect in Jamaica: The architecture of phonological variation*. Zenodo. <https://zenodo.org/record/1306618> (18 April, 2019).
- Janitza, Silke, Carolin Strobl & Anne-Laure Boulesteix. 2013. An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics* 14(119). 1–11.
- Jankowski, Bridget L. & Sali A. Tagliamonte. 2014. On the genitive's trail: data and method from a sociolinguistic perspective. *English Language and Linguistics* 18(02). 305–329.
- Jantos, Susanne. 2010. Agreement in educated Jamaican English: A corpus-based study of spoken usage in ICE-Jamaica. In Heidrun Dorgeloh & Anja Wanner (eds.), *Syntactic Variation and Genre*, vol. 70, 305–332. Berlin, New York: Mouton de Gruyter.

- Jenset, Gard B., Barbara McGillivray & Michael Rundell. 2018. The dative alternation revisited: Fresh insights from contemporary British Spoken Data. In Vaclav Brezina, Robbie Love & Karin Aijmer (eds.), *Corpus approaches to contemporary British speech: Sociolinguistic studies of the spoken BNC2014*, 185–208. New York: Routledge, Taylor and Francis.
- Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe - Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36, 15–43. Berlin/New York: Mouton de Gruyter.
- Kuhn, Max, With Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, et al. 2016. *caret: Classification and regression training* (R Package Version 6.0-73). <https://CRAN.R-project.org/package=caret>.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press.
- Labov, William. 2001. *Principles of Linguistic change: Volume 2: Social factors*. Cambridge and Oxford: Blackwell.
- Lavandera, Beatriz. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7(2). 171–183.
- Lawton, David. 1980. Language attitude, discreteness, and code-shifting in Jamaican Creole. *English World-Wide* 1. 211–226.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam/Philadelphia: John Benjamins.
- MacWhinney, Brian. 1997. Second language acquisition and the Competition Model. In Anette M. B. De Groot & Judith F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 113–142. Mahawa, NJ: Lawrence Erlbaum Associates.
- Mair, Christian. 2002. Creolisms in an emergent standard: Written English in Jamaica. *English World-Wide* 23(1). 31–58.
- Mair, Christian. 2013. The world system of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide* 34(3). 253–278.
- Miller, Faye. 1987. Acrolectal Jamaican English: some aspects of phonological and morpho-syntactic variation. Mona: University of the West Indies MPhil thesis.
- Patrick, Peter L. 1999. *Urban Jamaican Creole: Variation in the Mesolect* (Varieties of English Around the World). Vol. G17. Amsterdam: John Benjamins Publishing Company.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rosenfelder, Ingrid. 2009. Sociophonetic variation in educated Jamaican English: An analysis of the spoken component of ICE-Jamaica. Freiburg, Germany: University of Freiburg PhD thesis.
- Rosenfelder, Ingrid, Susanne Jantos, Nicole Höhn & Christian Mair. 2009. Manual for the Jamaican component (ICE-JA). <http://ice-corpora.net/ice/download.htm>.
- Röthlisberger, Melanie. 2018. Regional variation in probabilistic grammars: A multifactorial study of the English dative alternation. Leuven: KU Leuven PhD dissertation.
- Röthlisberger, Melanie, Jason Grafmiller & Benedikt Szmrecsanyi. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28(4). 673–710.
- Sand, Andrea. 1999. *Linguistic variation in Jamaica: A corpus-based study of radio and newspaper usage*. Tübingen: Narr.
- Sand, Andrea. 2005. Angliversals? Shared morpho-syntactic features in contact varieties of English. Freiburg, Germany: University of Freiburg Habilitation thesis.

- Sankoff, Gillian. 1973. Above and beyond phonology in variable rules. In Charles-James N. Bailey & Roger W. Shuy (eds.), *New ways of analyzing variation in English*, 44–62. Washington, DC: Georgetown University Press.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Seoane, Elena. 2009. Syntactic complexity, discourse status and animacy as determinants of grammatical variation in Modern English. *English Language and Linguistics* 13(03). 365.
- Shields, Kathryn. 1989. Standard English in Jamaica: A case of competing models. *English World-Wide* 10. 41–53.
- Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9(307). <http://www.biomedcentral.com/1471-2105/9/307>.
- Strobl, Carolin, Torsten Hothorn & Achim Zeileis. 2009. *Party on! A new, conditional variable importance measure for random forests available in the party package*. Technical Report. Munich: Department of Statistics.
- Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English* (Trends in Linguistics. Studies and Monographs [TiLSM]). Berlin/New York: Mouton de Gruyter.
- Szmrecsanyi, Benedikt, Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger. 2016. Around the world in three alternations: Modeling syntactic variation in varieties of English. *English World-Wide* 37(2). 109–137.
- Tagliamonte, Sali A. 2012. *Variationist sociolinguistics: Change, observation, interpretation*. Malden, MA: Wiley-Blackwell.
- Tagliamonte, Sali A. 2014. A comparative sociolinguistic analysis of the dative alternation. In Rena Torres-Cacoullos, Nathalie Dion & André Lapierre (eds.), *Linguistic variation: Confronting fact and theory*, 297–318. London/New York: Routledge.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.
- Tagliamonte, Sali A., Mercedes Durham & Jennifer Smith. 2014. Grammaticalization at an early stage: Future *be going to* in conservative British dialects. *English Language and Linguistics* 18(1). 75–108.
- Tagliamonte, Sali A. & Jennifer Smith. 2005. No momentary fancy! The zero in English dialects. *English Language and Linguistics* 9(2). 289–309.
- Tamaredo, Iván. 2018. Pronoun omission in high-contact varieties of English: Complexity versus efficiency. *English World-Wide* 39(1). 85–110.
- The Jamaican Language Unit. 2005. *The language attitude survey of Jamaica: Data analysis*. Online publication. Department of Language, Linguistics & Philosophy, The University of the West Indies, Mona, Jamaica. (29 April, 2019).
- The Jamaican Language Unit. 2007. *The language competence survey of Jamaica: Data analysis*. Online publication. Department of Language, Linguistics & Philosophy, The University of the West Indies, Mona, Jamaica. (29 April, 2019).
- Theijssen, Daphne, Joan Bresnan, Marilyn Ford & Lou Boves. 2011. In a land far far away... A probabilistic account of the dative alternation in British, American, and Australian English. Manuscript.
- Weiner, Judith & William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19(1). 29–58.
- Winford, Donald. 1991. The Caribbean. In Jenny Cheshire (ed.), *English around the world: Sociolinguistic perspectives*, 565–584. Cambridge: Cambridge University Press.
- Winford, Donald. 1997. Re-examining Caribbean English Creole Continua. *World Englishes* 16(2). 233–279.

Wolk, Christoph, Joan Bresnan, Anette Rosenbach & Benedikt Szmrecsanyi. 2013. Dative and genitive variability in Late Modern English: Exploring cross-constructural variation and change. *Diachronica* 30(3). 382–419.

Zuur, Alain F., Elena N. Ieno, Neil J. Walker, Anatoly A. Saveliev & Graham M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. New York: Springer.

Appendix: List of verbs considered in the extraction of dative variants:

accord, advise, allocate, allot, allow, answer, appoint, ask, assign, assure, award, bequeath, bid, bring, call, carry, cause, cede, charge, concede, convey, cost, deal, deliver, demonstrate, deny, develop, drop, entrust, explain, extend, feed, flick, flip, forward, get, gift, give, grant, guarantee, hand, impart, inform, issue, keep, lease, leave, lend, loan, lose, mail, name, offer, owe, pass, pay, permit, play, pose, post, prescribe, present, promise, propose, provide, quote, read, recommend, refuse, render, sell, send, serve, set, show, sing, slip, submit, suggest, supply, take, teach, tell, throw, toss, vote, wish, write, yield