*Journal of Linguistic Geography* (2019), **7**, 33–45
doi:10.1017/jlg.2019.4

CAMBRIDGE UNIVERSITY PRESS

Article

# When Wurst comes to Wurscht: Variation and koiné formation in Texas German

Luke Lindemann

Department of Linguistics, Yale University, New Haven, CT, USA

**Abstract**

Texas German is a new world language variety that shows some evidence of koiné development but also presents with substantial variation at many levels of structure. I present a case study on the variant pronunciation of sibilants in Texas German consonant clusters. This feature is fairly frequent and found throughout the regions of German settlement in Central Texas. After a discussion of the presence of this feature in the donor dialects, I investigate the factors that correlate with variation in the modern language. From an analysis of local and global spatial autocorrelation, I argue that variation is not significantly associated with particular geographic regions and is compatible with stable and homogenous variation. This provides insight into our understanding of new dialect emergence and the mechanisms by which dialect features are leveled over multiple generations.

**Keywords:** Texas German; German; koinéization; dialect formation; variation; dialectology; geospatial analysis; heritage varieties; spatial autocorrelation; sibilants

(Received 25 October 2017; accepted 15 August 2018; first published online 5 August 2019)

## 1. Introduction

Recent advances in the ability to collect and categorize large amounts of linguistic data provide new opportunities for understanding the mechanisms of language variation and shift (Boas, 2002; Zanuttini & Horn, 2014; Wood et al, 2015). Investigation into Texas German, an endangered but geographically widespread heritage language, is particularly useful for the investigation of koinéization, the process by which related dialects brought together into a new area mix together to create a new stable and homogenous variety. A koiné is a language contact phenomenon, a "stabilized contact variety which results from the mixing and subsequent leveling of features of varieties which are similar enough to be mutually intelligible, such as regional or social dialects" (Siegel, 2001:175). Sophisticated analysis of the variation and distribution of Texas German has been made possible by the ongoing Texas German Dialect Project (Boas et al, 2010). Boas (2009) utilizes Trudgill's (2004) model of new dialect formation to argue that the presence of substantial variation in modern Texas German suggests that the koinéization process is incomplete. Alternatively, it could be understood as stable variation of a coherent unary dialect. To distinguish between these hypotheses, it is necessary to carefully examine the factors that are associated with this variation.

The focus of the current investigation is the variant pronunciation of sibilants in *rst*-clusters, which are pronounced as [s] in Standard German and either [s] or [ʃ] in Texas German. I examine the pronunciation of *Wurst* 'sausage,' *Donnerstag* 'Thursday,' and *Haarbürste* 'hairbrush' in the archives of the Texas German Dialect Project.

The goal of the current study is twofold. First, I supplement our knowledge of the wider pattern of variation with data from interviews collected by the TGDP throughout Texas. This will allow for a comparison between the previously studied pattern of the New Braunfels settlement with other communities to explore the extent to which the New Braunfels pattern is representative of Texas German as a whole. Secondly, I explore in more detail the patterns of variation found in modern Texas German by looking at social factors that could correlate with the expression of the variant: age, gender, and place of birth. I run a statistical clustering analysis on the geographical regions to answer the question of whether or not communities differ significantly in the expression of this variation, and the consequences this has for the development of a new world dialect.

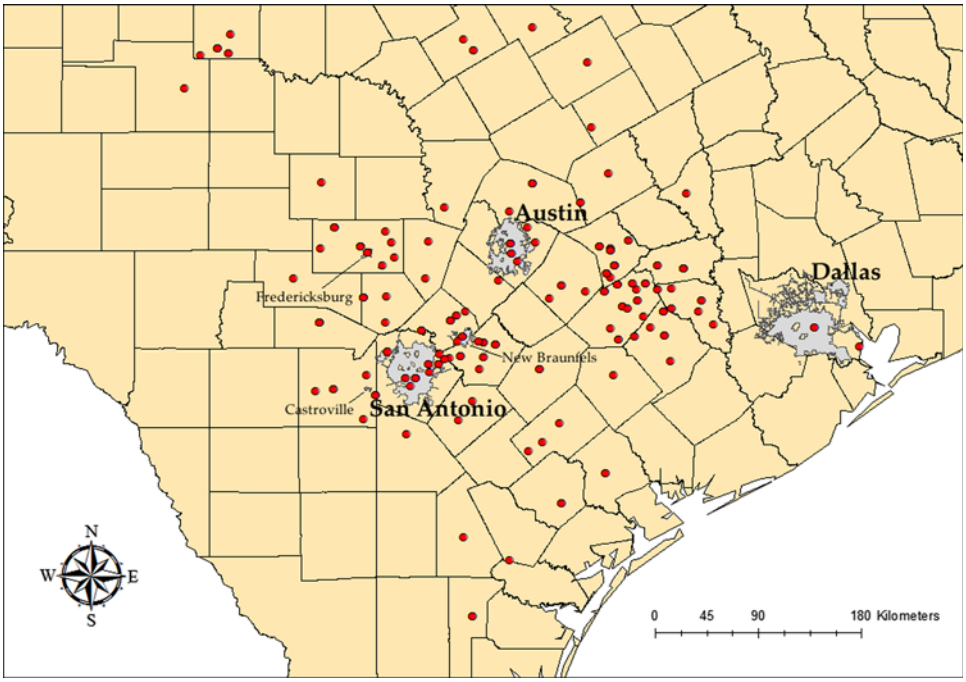
I argue that the lack of a positive correlation for these factors, particularly for geographic region, provides evidence for the homogeneity of this feature. This paper is a case study of one particular phonological alternation found in Texas German, and therefore I cannot here make any definitive claims about the larger pattern of Texas German koiné formation. Rather, this study represents an important step in characterization of the variation that exists within Texas German as a whole.

The next section provides more background on the history and development of Texas German and an overview of previous scholarship on the region. This is followed by a discussion of Trudgill's model of dialect formation as it relates to Texas German, and variation and dialect convergence more broadly, before an analysis of the data collected from the Texas German Dialect Project.

**Author for correspondence:** Luke Lindemann, Email: luke.lindemann@yale.edu
**Cite this article:** Lindemann L. (2019) When Wurst comes to Wurscht: Variation and koiné formation in Texas German. *Journal of Linguistic Geography* **7**: 33–45, https://doi.org/10.1017/jlg.2019.4

© Cambridge University Press 2019.

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4




**Map 1.** TGDP respondent birthplaces

## 2. Texas German History and Development

Map 1, which depicts the reported birthplaces of every Texas German Dialect Project respondent as of 2017, provides a good picture of the overall spread of German settlement in Texas.

The map was created with ESRI's Geographic Information Systems (GIS) software, ArcGIS.[1] Each token in the dataset is associated with the geographical coordinates of a speaker's birthplace. These have been plotted on a map of Texas counties and major cities created from US Census data.

Large-scale German immigration to Texas began in the 1840s and continued until around 1890 (Biesele, 1930; Jordan, 1977:2). The earliest immigrations were to settled land in the eastern part of the Texas German belt. Settlers came primarily from the Low German areas of present-day north-central and northwestern Germany, particularly the Münsterland, Holstein, Mecklenburg, and Oldenburg (Jordan, 1966:63). Western settlement onto the Texas frontier began in the mid-1840s. Settlers came primarily from the Mid- and High German areas of present-day central Germany, particularly the Hessen-Nassau region, southern Hannover, Württemberg, and Alsace (Jordan, 1966:120), and established the major settlements Fredericksburg and New Braunfels. Compared to the eastern settlements, the settlements in the west were more isolated from the settlements of English-speaking Anglo-American settlers.

By 1890, there were roughly 145,000 German-Americans in the state, representing more than 6% of the state's population (Jordan, 1977:2). Due to the general isolation of German-speaking communities along the Western frontier, a robust enclave of German speakers persisted well into the 20th century (Boas, 2009:47).

Despite the advent of English-only education laws in the 1890s that were relatively unenforced, the German belt represented a robust linguistic enclave, or Sprachinsel, until World War I. German was the language of education, the press, religion, and commerce. This situation shifted dramatically between World War I and II due to increased mobility and the stigma of the German language. In almost all cases, children born after World War II were raised entirely in English. While Texas German may have as many as 6,000 modern-day speakers, almost all of them are over the age of seventy (Salmons, 1983; Wells, 1985; Boas, 2009:74).

Much of the research on Texas German in the 20th century focused on a variety of German spoken in single settlements (Eikel 1949, 1966a, 1966b, 1967; Clardy, 1954; Gilbert, 1965). The settlement of New Braunfels in Comal County was the research site for Eikel (1954) and Gilbert (1963), while Wilson (1960) focuses on the German of Lee and Fayette Counties, and Guion (1996) looks at the settlement of Fredericksburg and surrounding Gillespie County. The noteworthy exception to this pattern is Gilbert (1972), presenting an extensive linguistic atlas of Texas German compiled through interviews with 286 informants throughout Texas. The atlas maps the geographical distribution of various phonological, morphological, and syntactic features which vary in Texas German.

Boas (2009) is a study of Texas German based on the earliest interviews conducted by the Texas German Dialect Project (TGDP), a central database of Texas German interviews at the University of Texas. The data from Boas (2009) comes from the settlement of New Braunfels and surrounding Comal County; the German of this community has been relatively homogenous, allowing a comparison of Boas' results with the results of other studies from the previous century (Boas, 2009:16). Boas (2009) is based on the interviews from 52 speakers around New Braunfels, using open interviews and samples of the same questionnaires used by Gilbert (1972).

Boas corroborates the presence of variant features found in Gilbert (1972) for present-day Texas German. Although Gilbert maps extensive variation in Texas German, he conceives of Texas German as a whole: "The trend is clearly toward a dissolution of the old, fragmented, mutually unintelligible (or at best partially intelligible) dialects, either by outright replacement or by gradual modification to form a new type of speech, which,

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

although far from uniform, enjoys sufficient common characteristics to merit the generic name, Texas German." (Gilbert 1972:4). Boas, conversely, concludes that the process of new dialect formation was incomplete when it was interrupted in the middle of the 20th century. The presence of continued inter- and intraspeaker variability indicates that the speech has not converged: "we find a broad spectrum of dialectal mixtures with considerable English admixture. What has traditionally been called 'Texas German' should thus be regarded as a collection of various subvarieties that share a limited set of linguistic features" (Boas, 2009:98).

These two viewpoints represent different hypotheses about the nature of the variety (or varieties) of German spoken in Central Texas. The extent to which variation can be taken as evidence for the non-convergence of a group of dialects into a koiné will depend upon the precise nature of the variation and the model used to describe how a group of dialects converge to form a new dialect.

### 2.1 Variation in Texas German

Boas' (2009) analysis of new dialect formation in Texas German is based on Trudgill's (2004) model, which describes how a dialect mixture will cohere into a new dialect. This occurs over the course of three stages which correspond roughly to generations of settlement in a homogenous community without a previously established language population. Trudgill's model is deterministic and predicts the survival of dialectal variants based on strict proportionality at the feature level.

The first stage of Trudgill's model consists of rudimentary leveling and interdialect development of the adult immigrant generation. Accommodation between adults begins en route to the destination, as the most salient features of traditional dialects are leveled in favor of intermediate forms. The second stage is characterized by extreme intra-individual and inter-individual variability of forms in the first native-born generation. Minority variant forms that exist below a certain rate are leveled out because they are so rare as to be unnoticed by young speakers. Children are exposed to a wide degree of variation in the environment. During the third and final stage, a focused, stable, and coherent dialect emerges. The majority variant in the original dialect mixture will survive in the new dialect. Table 1, adapted from Boas (2009:109), shows the successful leveling of a particular variant among Eikel's 24 New Braunfels German speakers: rounded/unrounded front vowels ([y:]/[i:] and [ø]/[e:]).

Of the nine phonological features described by Boas, three appear to have gone through all three stages and exhibit little variation in the modern language. Other morphological and phonological features described by Boas (2009) show at least some evidence of leveling in the direction of one particular variant (Boas et al, 2004; Pierce et al, 2015). However, there is still substantial variation in the modern language. This includes the variant pronunciation of *rst*-clusters discussed in the next section. Boas takes this variation as evidence that the process of focusing was never completed: some features reached Stage Three, leaving only one surviving variant, while other features never left Stage Two.

The argument against convergence is that the koinéization process was interrupted by large-scale language shift to English, and variation in the modern language reflects this lack of convergence. There were at least three generations of Texas Germans before this shift began, so the process could conceivably have been completed in this timeframe. Kerswill (2002) notes that koinéization "typically takes two or three generations to complete, although it is achievable within one." However, Kerswill and Trudgill (2005) present case studies for which koinéization has not occurred over multiple generations, typically due to sociological factors such as isolation and continuing in- and out-migration, although there are a finite number of these factors. Trudgill (2011) argues for two factors in particular that determine the speed of linguistic change: relative degree of isolation and relative social stability of the community. Focusing happens more quickly in isolated, socially homogenous communities. Baxter, Croft, and McKane (2009) criticize Trudgill's (2004) model on the basis that solely deterministic processes cannot account for the emergence of a new dialect in an appropriate timeframe. They conclude that social and other non-deterministic factors must also be involved in order for the convergence to happen in a realistic timeframe.

**Table 1.** Age-graded Distribution of Rounded and Unrounded Front Vowels

| Speakers | Percentage (n = 24) |
|---|---|
| Generation 1 (b. 1855–1875) | |
| rounded | 33.3% |
| unrounded | 33.3% |
| mixed | 33.3% |
| Generation 2 (b. 1880–1910) | |
| rounded | 0.0% |
| unrounded | 0.0% |
| mixed | 100.0% |
| Generation 3 (b. 1910–1930) | |
| rounded | 0.0% |
| unrounded | 100.0% |
| mixed | 0.0% |

For Texas German, these factors may include influence from surrounding English speakers and the influence of Standard German through media and education. Both of these factors are discussed by Boas, although he argues that the influence of Standard German on Texas German is overstated (contra Salmons & Lucht, 2006). Boas (2009) argues that the loss of the dative case is due to the leveling process of dialect contact rather than influence from English. Another factor is the continuing immigration from Germany during the latter half of the 19th century, and, to a lesser degree, the 20th. Nützel and Salmons (2011) note that ongoing immigration may be a factor in slowing the process of koinéization.

Trudgill's model describes the process by which variant features in the dialect mixture come to stabilize such that a single variant prevails. This question of dialect convergence has recently been taken up by others working on Germanic heritage varieties in the United States (Johanneson & Salmons, 2015). The presence of significant intra- and interspeaker variability is taken as evidence against the formation of a new dialect in these works. Nützel (1998, 2009) is a description of Haysville East Franconian, a dialect of German spoken in Indiana that has avoided the leveling processes of koinéization that have affected other varieties of German in Indiana. Nützel and Salmons (2011) argue that other varieties of German have had little effect on Haysville East Franconian because of the tight-knit nature of the community and their linguistic and

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

religious separation with their neighbors. In contrast, a fully-fledged dialect is "relatively uniform," meaning that the majority of feature variants have been leveled out (Baxter et al, 2009).

On the other hand, all natural languages exhibit feature variation across levels of structure. This may particularly be the case for unwritten language varieties that differ from a perceived standard, as is the case for Texas German. Dorian (1973, 1983, 1994) investigates the presence of variation in East Sutherland Gaelic and concludes that a substantial level of variation cannot be attributed to language shift, prestige, social factors, or attrition. It is simply stable variation in the language.

East Sutherland Gaelic and Texas German have more in common than might be expected. There is, in both cases, the effect of English as a dominant language as well as influence from a perceived (Gaelic and German) standard. While East Sutherland Gaelic is not considered to be a koiné, some of the variation is very likely due to dialect mixture that occurred during the forced migration of inlanders to the coast in the early 19th century. Dorian writes that "nineteenth-century population mixture can thus be invoked to account for the appearance of some multiple variants in ESG, as can local language history" (Dorian, 1994:691).

If a dialect mixture exhibits substantial variation, it may be the case that it has not yet focused into a unary dialect, but it may also be the case that the variation is stable and representative of a homogenous variety. In order to distinguish these cases, it is necessary to test which factors correlate with this variation. If there remain regions of statistically high proportionality of a particular variant that correspond to particular communities or groups of communities, then this provides evidence that a group of dialects has not yet become a koiné. If multiple variant features cannot be correlated with an expected geographic distribution, this may constitute evidence that a new dialect has formed. The expected geographic distribution would be a significant clustering of particular variants corresponding to the boundaries of particular communities or homesteads.

As to the expected geographic distribution for the Texas German case, Boas (2009) and Gilbert (1972) mention a distinction between the western part of the German belt, where isolation from English-speaking communities was larger, homesteads were more isolated, and there were large German-speaking towns, and the eastern edge, in which Germans homesteads coexisted with more heterogenous communities of Eastern European and Anglo-American settlers. In these eastern Texas settlements, Wilson (1960) also notes small deviations in dialect due to influence from a settlement of Wends, a group of settlers from Saxony who came in the 1850s and shifted from Slavic Wendish to German. Additionally, the western settlement of Castroville was settled by speakers of French-influenced Alsatian German, and Roesch (2012) provides a description of the unique form of Texas Alsatian spoken there. If leveling of features has not occurred, then we can expect to find different patterns of variation corresponding to particular large German communities like New Braunfels, Fredericksburg, and Castroville, and perhaps a general differentiation between the east and west.

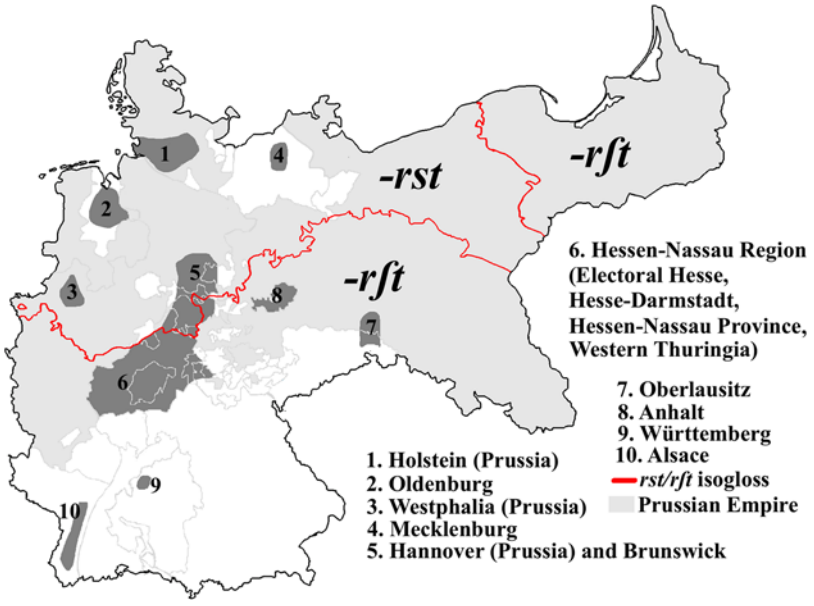
### 2.2 Sibilant pronunciation in rst-clusters

For this study, I analyze one of the variants discussed in Gilbert (1972): the pronunciation of sibilants in *rst*-clusters. I have chosen this variant because it is well-represented in the TGDP database, both in the questionnaires and the open interviews.

Gilbert (1972) first documented the variable distribution of [s] and [ʃ] in consonant clusters in the words *Donnerstag* ('Thursday'), *Wurst* ('sausage'), and *Haarbürste* ('hairbrush'). The s-variant is Standard German, and the ʃ-variant existed primarily in Middle and High German dialects.

Map 2 depicts the German states in 1871. The red line represents an isogloss of the pronunciation of the word *Wurst* ('sausage'), traced from the *Digitaler Wenkeratlas*[2], a digitized database of dialect maps which were compiled in the late 19th century (Schmidt et al, 2008; Wenker, 1888–1923). The isogloss for *Wurst* is nearly identical to the isogloss for *Durst* ('thirst'), the only other *rst*-cluster word for which there is data in the Wenker Atlas.[3]

The dark grey areas represent the major sources of immigrants from Germany to Texas according to Jordan (1966:33). Jordan uses




**Map 2.** Sibilant pronunciation in *Wurst* and the origin of German settlers

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

**Table 2.** Pronunciation of *Wurst* in immigrant donor dialects

| Regional Pronunciation | Map Regions | Eastern TX settlements | Western TX settlements |
|---|---|---|---|
| Primarily s-variant | 1,2,3,4,5 | 26% | 23% |
| Primarily ʃ-variant | 6,7,8,9,10 | 19% | 41% |
| "Unspecified Prussia" | 1,3,5,6,7 | 54% | 36% |

1860 and 1870 US Census data in central Texas counties for his analysis. For the eastern Texas settlements, the bulk of immigration was from northwestern and north-central Germany. For the western Texas settlements on the frontier (Jordan, 1966:123), an immigration society with roots in the Hessen-Nassau region of central Germany sponsored initial immigration. Based on Jordan's figures and the isogloss from the Wenker atlas (Jordan, 1966:64, 123), Table 2 shows the estimated percentage of immigrants from regions in which the s-variant for *Wurst* is dominant (Hannover, Brunswick, Mecklenburg, Oldeburg, Lippe-Detmold, Hamburg, Hannover, Bremen) and regions in which the ʃ-variant is dominant (Saxony, Anhalt, Baden, Württemberg, Hesse, Nassau, Bavaria, Saxe-Meiningen, Saxe-Weimer).

The main difficulty in determining the exact proportionality of features is that nearly half of the respondents listed "Prussia" as their place of origin, which could indicate any of the provinces of the Prussian Empire (the light grey area in Map 2). Jordan notes that this includes immigrants from Westphalia, Holstein, parts of Hannover, Hesse, and Nassau, and in particular the Wetzlar-Braunfels area of Hesse that supplied many of the immigrants to the western settlements. The ʃ-variant was likely the dominant form in the donor dialects listed as "Unspecified Prussia," and in the dialect mixture as a whole. Therefore, Table 2 gives us a lower bound: a minimum of 23% of the immigrant population came from areas with the s-variant for *Wurst*, with a somewhat higher percentage in the eastern settlements.

Trudgill's model predicts that the ʃ-variant for *Wurst* will eventually replace the s-variant. Previous research in New Braunfels indicates the presence of variation, but with some evidence of focusing: Gilbert (1972) found that 87% of the speakers used the ʃ-variant, and Boas (2009) found a higher proportion of 94% for his respondents three decades later.

For the other words, *Haarbürste* and *Donnerstag*, Gilbert found the s-variant to be more prevalent in New Braunfels, and in fact it was in the majority for *Donnerstag*. In both cases Boas found an increase in the usage of the ʃ-variant. These findings are summarized in Table 3, which is culled from Boas (2009:153). For *Donnerstag*, this goes against the expectation that the majority form will overtake the minority form.

Boas writes that "the change could be explained in terms of leveling, where in the mid-20th century *Donnerstag* was one of the few words for which the majority of informants still preferred [s] over [ʃ] in this consonant cluster. By the beginning of the 21st century, the almost exclusive use of [ʃ] in this context was extended from other words to *Donnerstag*" (Boas, 2009:154). Boas also notes that in open interviews some speakers use an ʃ-variant for *erst* ('first'), which did not originally vary.

In the current analysis, I examine the distribution of *Wurst*, *Haarbürste*, and *Donnerstag* variants from all forty-four different Texas counties for which we have data.

**Table 3.** Proportion of ʃ- to s-variants in New Braunfels

| | Wurst | Haarbürste | Donnerstag |
|---|---|---|---|
| Gilbert 1972 (n = 15) | 87% | 53% | 13% |
| Boas 2009 (n = 52) | 94% | 96% | 77% |

## 3. Study

The data for this analysis come from recorded questionnaires which were collected between 2004 and 2016 by the TGDP. A typical interview session consists of a brief survey on the background and language attitudes of the speaker, an open interview in German on any topic, and the elicitation of the Gilbert and/or Eikel questionnaires. For these questionnaires, the speaker is prompted with a word or sentence in English and asked to provide a translation in Texas German. Each survey consists of approximately 200 questions. Boas et al (2010) is a comprehensive description of the workflow of the TGDP. I compiled the responses to those questions in the Gilbert questionnaire for s/ʃ-alternations in the words *Haarbürste* 'hairbrush', *Donnerstag* 'Thursday', and *Wurst* 'sausage'. I examined the same questions listed in Boas (2009). For this questionnaire, questions (1) and (16) are identical, as are (15) and (113). The relevant words are in bold:

1: A **hairbrush**

14: **Thursday**

15: He took the most **sausage**.

16: A **hairbrush**

113: He took the most **sausage**.

These tokens came from untranscribed audio recordings from the TGDP database.[4] The determination of whether a particular token instantiated the s-variant or the ʃ-variant was made by aural inspection. I used a self-written Python computer program which played every recorded response for a particular question without any identifying information and in a randomized order. I excluded responses which contained neither variant, but I included compounds and variant pronunciations of vowels. For example, for *Haarbürste* ('hairbrush'), I made the determination based entirely on the quality of the sibilant and included variations like *Birste*, *Haarberst*, *Kaambürst* ('comb-brush'), etc. I checked my judgments with a non-German speaker, who was presented with 10% of the tokens for each question. Our judgments were in agreement for 97.1% (100/103) of these tokens.

From the biographical questionnaire forms filled out by each informant I recorded gender, year of birth, and county of birth. The collected data come from 486 informants (193 female, 282 male, and 10 other/undetermined). There were 130 uniquely listed birthplaces in 45 different Texas counties. These speakers were born between 1908 and 1979 with a median birth year of 1933. Most informants were born in rural communities, and informants typically described place of birth with reference to either the name of a small town or settlement (e.g., "Converse," "Warda"), or county ("Lee County," "Bexar County"), or orientation with respect to a town or city ("13 miles north of Fredericksburg"). I converted this information into geographic coordinates using the online geocoding tool at http://LatLong.net, a search function that provides geographic coordinates for cities, towns, and counties. There were 1068 tokens in total, of which 95 were excluded because a confident determination of the variant could

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

not be made (454 for *Wurst*, 255 for *Donnerstag*, and 264 for *Haarbürste*).

### 3.1 Nested model comparison

In order to determine which predictors may characterize *rst*-alternation, I considered the following variables for each token:

Dependent Variable:

Value (i.e., [s]/ [ʃ])

Predictors:

1. Word (*Haarbürste, Donnerstag, Wurst*)
2. Gender of Speaker (Male, Female, none)
3. Speaker's Year of Birth
4. Speaker's Place of Birth
5. Speaker ID Number

Any of these four predictors could plausibly correlate with the observed feature variation. However, if the feature variation is evenly distributed and stable in the population, then Place of Birth and Gender will not be significant predictors. Year of Birth (YOB) may be a significant predictor. This would indicate phonological change over the course of the generations represented in the survey. The dependent variant Value is coded as a binary variable (s = 0, ʃ = 1). Word, Gender, and Speaker ID are categorical variables. I set "Wurst" as the reference for the Word variable and "F" as the reference for the Gender variable. Year of Birth (YOB) is coded as a continuous variable, and Place of Birth is coded with two continuous variables Latitude and Longitude. I scaled and z-normalized the three continuous variables in order to compare the effect sizes of each variable. I used the 800 tokens that contained no missing data (some respondents did not record their age or birthplace).

I determined which predictors are significant by comparing logistic regression models pairwise to determine which model best fit the data. I used the Akaike Information Criterion to compare model fitness (AIC; Akaike, 1974). A lower AIC, taken to be a reduction of at least 2, means that the added complexity of the model is warranted given the goodness of fit. First, I ran a generalized logistic regression in R (R Core Team, 2017). I used the predictor variables Word, Gender, YOB, Latitude, and Longitude. I compared this with a mixed-effect logistic regression with the same predictors and with Speaker ID as a random effect variable. I ran this using the glmer package in R (Bates et al, 2014).

The purpose of this comparison is to see whether the model improves significantly when individual speaker variation is taken into account. Table 4 is a comparison of the results of these two models. The mixed-effect model performs decidedly better (AIC = 412.9, compared to 462.9), which indicates that the speakers are not a homogenous group. The mixed-effect model is a better fit. The only significant predictor at this level of complexity comes from a Word variable.

In order to determine which variables are significant, I performed a nested model comparison of mixed-effects models. I excluded each variable in turn and compared the AIC to the original model. Excluding Gender gave the greatest positive change in AIC (8.7), and so I conclude that a model which excludes Gender is a better fit. I then excluded each variable in turn from this new model, but the AIC improvement in each case was less than two. Therefore, the best fit was a model which removed Gender from consideration, leaving Word, Latitude, Longitude and YOB. This indicates that Word, Latitude, and Gender are allpotentially significant variables. Table 5 is a summary of this model.

**Table 4.** Model Comparison of Generalized Linear Model with Mixed Effect Model for all variables

| Generalized Linear Model | | | | |
|---|---|---|---|---|
| Coefficients: | Estimate | Std. Error | z value | Pr(>\|z\|) |
| (Intercept) | 2.6311 | 0.2533 | 10.389 | <2e–16 *** |
| Word-Hairbrush | 0.3146 | 0.4096 | 0.768 | 0.443 |
| Word-Thursday | −1.5988 | 0.2830 | −5.649 | 1.96e–08 *** |
| Gender-unspec | −0.1693 | 1.1529 | −0.147 | 0.883 |
| Gender-M | 0.4843 | 0.2585 | 1.873 | 0.061 |
| YOB.z | 0.2158 | 0.1319 | 1.635 | 0.102 |
| Long.z | −0.1399 | 0.1386 | −1.009 | 0.313 |
| Lat.z | 0.1761 | 0.1202 | 1.465 | 0.143 |
| AIC: 462.9 | | | | |
| Generalized Linear Mixed-Effect Model | | | | |
| Coefficients: | Estimate | Std. Error | z value | Pr(>\|z\|) |
| (Intercept) | 5.6260112 | 0.0007137 | 7882.9 | <2e–16*** |
| Word-Hairbrush | 1.1374962 | 0.0007137 | 1593.9 | <2e–16*** |
| Word-Thursday | −2.6047295 | 0.0007137 | −3649.8 | <2e–16*** |
| Gender-unspec | 0.6621558 | 0.0007137 | 927.8 | <2e–16*** |
| Gender-M | 0.8512719 | 0.0007137 | 1192.8 | <2e–16*** |
| YOB.z | 0.6175153 | 0.0007137 | 865.3 | <2e–16*** |
| Long.z | −0.3556039 | 0.0007137 | −482.5 | <2e–16*** |
| Lat.z | 0.4296658 | 0.0007137 | 582.9 | <2e–16*** |
| AIC: 412.9 | | | | |

Word is the only variable with a significant correlation. In order to evaluate the robustness of these correlations, I also conducted a bootstrapping resampling on the correlation coefficients, drawing randomly with replacement for 1000 iterations. The result of this analysis is in Table 6. If the coefficient of a variable ranges between positive and negative values, this indicates that the variable is not significant. From this I conclude that the only variable that significantly correlates with the pronunciation of sibilants in *rst*-clusters is the lexical item elicited (Word). I did not find a correlation with Gender, Year of Birth, or Birthplace.

**Table 5.** Mixed Effects Model with Word, YOB, Longitude, and Latitude

| Fixed Effects: | Estimate | Std. Error | Z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | 10.40003 | 1.60598 | 6.476 | 9.43e–11 *** |
| Word-Hairbrush | 0.99697 | 0.61419 | 1.623 | 0.105 |
| Word-Thursday | −4.34782 | 0.83999 | −5.176 | 2.27e–07 *** |
| YOB.z | 0.43592 | 0.439 | 0.993 | 0.321 |
| Long.z | −0.09205 | 0.42917 | −0.214 | 0.83 |
| Lat.z | 0.23536 | 0.44855 | 0.525 | 0.6 |
| AIC: 404.2 | | | | |

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

**Table 6.** Coefficient Correlation Confidence Intervals (Bootstrap Resampling of 1000 iterations

| | 2.5% | 97.5% |
|---|---|---|
| (Intercept) | 14.7652199 | 26.100204 |
| WordHairbrush | −1.275258 | 8.317408 |
| WordThursday | −14.37048 | −6.553844 |
| YOB.z | −0.7918391 | 1.044521 |
| Long.z | −1.9747366 | 0.767494 |
| Lat.z | −0.9443762 | 1.183687 |

### 3.2 Variation and word, gender, and year of birth

The nested model comparison tells us that the only significant predictor of the sibilant feature is the word in which the feature is used. In particular, the word *Donnerstag* ('Thursday') was significant. This is an expected result, if we consider that Boas (2009) found *Wurst* (the reference factor) and *Haarbürste* to have a similar proportionality in contrast to *Donnerstag*. Each of the words have their own distribution, and it is necessary to consider each word separately in the analysis to follow. Strikingly, the overall proportionality of each word across the entire dataset is practically identical to that described in Boas (2009) for the community of New Braunfels, as can be seen in Table 7.

While it is true that New Braunfels as a single community is well-represented in the dataset (13.7% of the tokens come from speakers who were born in New Braunfels, and 20.0% of the tokens come from speakers who were born elsewhere in Comal County), the close correspondence suggests that New Braunfels is a good proxy for Texas German as a whole, at least with respect to this particular feature. It would be interesting to see whether this is the case for the other features described in Boas (2009).

The statistical analysis did not find gender to be a significant factor. In the study sample, 44% of the speakers were female and 52% were male. Table 8 shows the results.

We expect a general trend toward the leveling out of the s-variant over time. However, the statistical analysis did not find Year of Birth to be a significant predictor, which would be expected if this leveling out is the result of an ongoing phonological trend. Gilbert (1972) and Eikel (1954) indicate that the relative proportionality of the ʃ-variant was lower in previous generations than we find in the modern data. It may be the case that this phonological trend towards the leveling out of the ʃ-variant which occurred in previous generations stabilized before the final generation. In other words, the variation in the current generation is stable variation which has continued from the oldest speakers to the youngest speakers in the current dataset. On the one hand, this fits with the idea from Boas (2009) that Texas German was following a path towards the complete leveling out of all variants, and during the third generation the massive social changes that led to the endangerment of Texas German caused it to "freeze" in its development. However, this is separate from the question of whether these variants have become sufficiently distributed geographically so as to constitute a new dialect, because stable variation may also occur in a relatively homogenous language.

**Table 7.** Overall Percentage of ʃ-variant by Lexical Item

| | Wurst | Donnerstag | Haarbürste | Total |
|---|---|---|---|---|
| Gilbert (1972) | 87% (n = 15) | 13% (n = 15) | 53% (n = 15) | 51% (n = 45) |
| Boas (2009) | 94% (n = 49) | 77% (n = 48) | 96% (n = 25) | 88% (n = 122) |
| Current Study | 94% (n = 454) | 77% (n = 255) | 94% (n = 264) | 90% (n = 973) |

**Table 8.** Overall Percentage of ʃ-variant by Gender

| | Total |
|---|---|
| Female | 87.3% |
| Male | 91.9% |

### 3.3 Individual speaker variation in the open interviews

The relative intra- and interspeaker variation is important for a proper understanding of the mechanisms of dialect formation. In particular, it would be useful to know how much variation can be attributed to variation within a single speaker. If most of the variation in variant choice is attributable to the speaker, and the individual differences between speakers are relatively small, then this would be evidence for focusing. Texas German has focused into a unary dialect with a probabilistic distribution of feature variants.

Gilbert (1972) recorded that one of his fifteen informants produced *Wurst* variably with either [s] or [ʃ], and that one of his fifteen informants produced *Donnerstag* variably. It is unclear if this is the same individual in both cases, but in any case there is reason to believe that some Texas German speakers produce these lexical items variably.

The survey results are not ideal for studying feature variation, because the questionnaires have only a few examples of multiple tokens of the same type from the same informant (there were two questions each for *Haarbürste* and *Wurst*). The items with multiple representations are those for which the ʃ-variant is nearly 20 times as common as the s-variant. Considering the low number of tokens per type, it is almost certain that the level of intraspeaker variation will be underestimated by the survey.

To get a better understanding of individual speaker variation, I analyzed the distribution of *rst* clusters from the open interviews separately. The open interviews are longer recordings of conversations between the interviewer and the respondent on any topic. The majority of respondents describe growing up on a farm in the Texas Hill Country, speaking German at home and learning English for the first time in school. A substantial proportion of these interviews have been transcribed and annotated. I used the concordancer tool at the online Texas German Dialect Archive to search through the transcribed open interviews (Speaker IDs 1 - 395). I collected every word which was transcribed with an "rst" cluster (1332 tokens) or an "rscht" cluster (300 tokens).

By far, the most common *rst*-word in the open interviews is *erst* ('first'). Including all derived forms and the adverb *zuerst*, this word appears 758 times in the corpus. The word is common in narratives about growing up in the Texas German belt.

**Table 9.** Individual speaker variation in the ʃ and s-variants of *Wurst*

| s-variant only | ʃ-variant only | Mixed s and ʃ |
|---|---|---|
| 5% (n = 2) | 88% (n = 37) | 7% (n = 3) |

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

**Table 10.** Individual speaker variation in the ʃ and s-variants of *erst*

| s-variant only | ʃ-variant only | Mixed s and ʃ |
|---|---|---|
| 78% (n = 70) | 5% (n = 3) | 17% (n = 15) |

**Table 11.** Percentage of ʃ-variant by word in four major counties

| County | Wurst | Donnerstag | Haarbürste |
|---|---|---|---|
| Comal | 88% (n = 89) | 75% (n = 46) | 100% (n = 48) |
| Gillespie | 97% (n = 133) | 76% (n = 69) | 92% (n = 83) |
| Fayette | 100% (n = 39) | 85% (n = 19) | 100% (n = 14) |
| Lee | 100% (n = 35) | 100% (n = 19) | 100% (n = 23) |

**Table 12.** Variation in the lexical item "Pumpkin"

| Bungis | Pumpkin | Kurbis | Galawas | Other |
|---|---|---|---|---|
| 52% (n = 70) | 31% (n = 41) | 6% (n = 8) | 5% (n = 7) | 6% (n = 8) |

The most common noun with an *rst* cluster is *Wurst* ('sausage'). Including types of sausage (*Blutwurst*, *Fleischwurst*, *Knackwurst*, etc.) and other compounds (*Wurstfest*, *Wurstkäse*, *Wurststube*), the word appears 291 times in the corpus. Fredericksburg in particular is known for its yearly German cultural festival, *Wurstfest*. Discussions of tourism and development in modern-day Fredericksburg often mention the impact of this holiday. A great many speakers also discussed the process of preparing and smoking sausage at home.

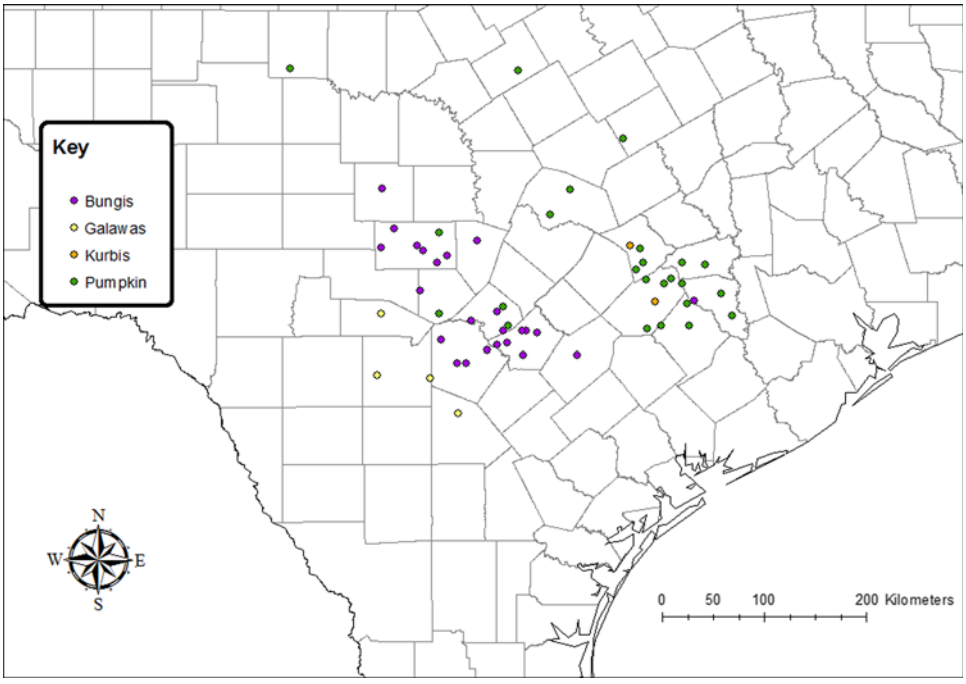
Among the words analyzed in the surveys, *Donnerstag* appears only 13 times and a variant of *Haarbürste* appears once. The remaining words in the open interviews are (1) always pronounced with [ʃ] according to German phonology (*verstehen*, *erstaunlich*, *Bürgersteig*); (2) second-person verb forms (*warst*, *erinnerst*); or (3) English loanwords or code-switching in English (*airstrip*, *understand*, *first*).

I listened to every interview clip that contained an annotation "Wurst" or "Wurscht" and made a determination of whether the sibilant was an [s] or [ʃ]. Out of the 51 speakers, 42 of them used the word *Wurst* more than once for a total of 208 tokens (2–20 tokens per speaker, with a median of 3 and mean of 4.9). The results are presented in Table 9.

The percentage of mixed speakers is probably still underestimated. If the average speaker uses the s-variant only 5% of the time, they are unlikely to use it once in five tokens. The data is consistent with a pattern in which every speaker is a mixed speaker and uses both variants. It is also consistent with a pattern in which there are exclusively ʃ-variant speakers and a minority of mixed speakers.

Boas (2009) mentions that the ʃ-variant pronunciation has apparently spread to the word *erst* for some speakers, although the s-variant is still in the majority. I also searched the transcribed open interviews for the annotations "erst" and "erscht." While I was not able to audibly verify the annotations in most cases, the transcriptions were generally perceived to be correct. If a bias exists, it will be towards the Standard German spelling "erst." Out of 100 speakers, 89 used the word *erst* more than once, for a total of 774 tokens (2–38 per speaker with a median of 7 and a mean of 8.7). The results are in Table 10.

This is consistent with a pattern in which the majority of speakers are exclusive s-variant users and about one-fifth are mixed speakers. This means that some focusing has occurred, but speakers are not a completely homogenous group—some speakers behave differently from others. I now turn to the question of whether these differences manifest geographically, meaning that different communities express different patterns, or if this variation is distributed homogenously throughout the Texas German belt.




**Map 3.** Distribution of variants for "Pumpkin"

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

**Table 13.** Moran's *I* Analysis for *Bungis/Pumpkin*

| Cutoff (km) | Moran's *I* | z-score | p-value | Result |
|---|---|---|---|---|
| 20 | 0.209271 | 8.032487 | <0.000001 | highly clustered |
| 40 | 0.218946 | 11.149148 | <0.000001 | highly clustered |
| 60 | 0.226641 | 13.125318 | <0.000001 | highly clustered |
| 80 | 0.242965 | 16.107117 | <0.000001 | highly clustered |
| 100 | 0.207658 | 21.200864 | <0.000001 | highly clustered |
| 120 | 0.186511 | 26.457181 | <0.000001 | highly clustered |
| 140 | 0.155158 | 25.892242 | <0.000001 | highly clustered |
| 160 | 0.126054 | 23.967244 | <0.000001 | highly clustered |
| 180 | 0.098621 | 20.688795 | <0.000001 | highly clustered |
| 200 | 0.039101 | 12.580116 | <0.000001 | highly clustered |

### 3.5 Variation and geography

Birthplace was not a significant predictor in the statistical model, but this will not tell us whether there are markedly different patterns in different communities. Table 11 shows the distribution for these variants in four representative counties: Gillespie and Comal in the west and Fayette and Lee in the east.

Overall the percentages are similar in each county, but there are some basic differences: *Donnerstag* appears to be nearly leveled out in the east, and the s-variant for *Wurst* is unusually common in Comal County. But we cannot tell from the data if these are statistically reliable patterns. In order to see whether there are reliable patterns, we need to look at measures of spatial autocorrelation.

### 3.6 A Demonstration of spatial autocorrelation

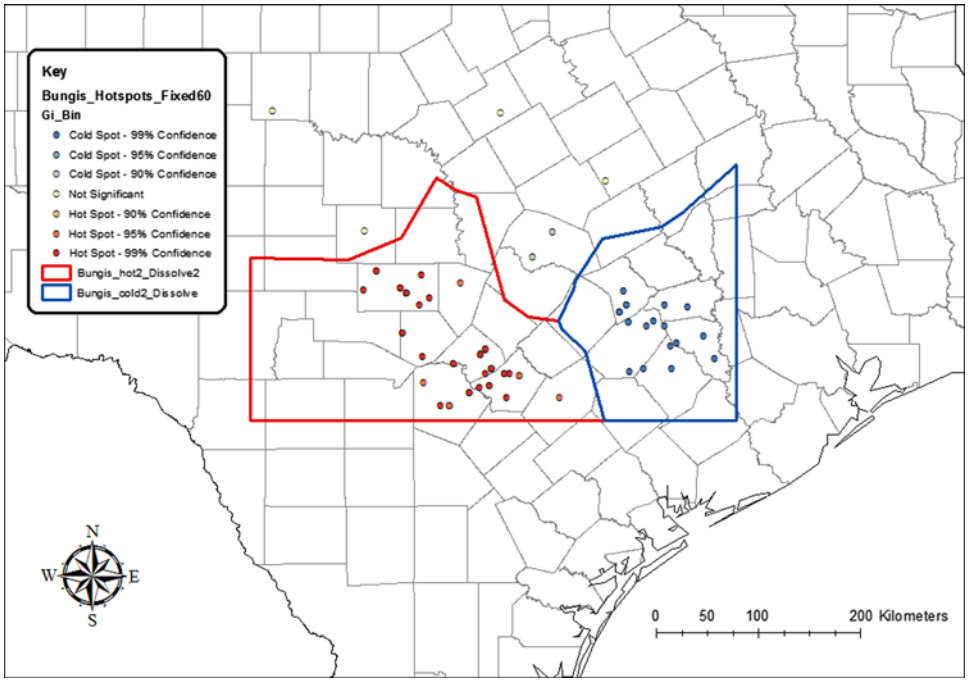
To examine the question of whether sibilant variation is diffused geographically, I first consider an unrelated feature of Texas German that does exhibit a clear clustering pattern of regional variation. In contrast to sibilant variation, there are certain lexical items which are clearly associated with certain regions. The feature presented here is the lexical item for "pumpkin," *Kurbis* in Standard German. I analyze this feature as a general demonstration of spatial autocorrelation statistics before moving on to sibilant variation. The data come from another question elicited by the Gilbert survey:

125. A **pumpkin** (the fruit of Cucurbita pepo)

Table 12 depicts the most common variants for this lexical item. The most common term is *Bungis* (including variants *Pungis*, *Bunkis*). The next most common term is an English loanword *Pumpkin*. The term *Galawas* (*Galavasa*, *Galawa*) is associated particularly with the Alsatian Texas German of Castroville and Medina County. Because many speakers report the same birthplace, some points in Map 3 represent multiple tokens.

*Bungis* appears to be the most common variant in the western settlements (excluding the influence of Alsatian Texas German in the southwest), while *Pumpkin* dominates in the eastern settlements. We can test the statistical reliability of these clusters with measures of global and local spatial autocorrelation using the Moran's *I* and Getis-Ord G* statistics (Grieve et al, 2011; Zanuttini et al, 2018; Tamminga, 2013).

The Global Moran's *I* statistic is useful for determining whether there is an overall pattern of regional clustering of a feature variant. Using ArcMap's Moran's *I* tool, I tested the comparative regional distribution of the two major variants *Bungis* and *Pumpkin* (coding them as 1 and 0 respectively). I used a fixed band with a specified cut-off distance (Grieve et al, 2011). This cutoff distance corresponds to the resolution of the clusters: a low cutoff is better for identifying small clusters, and a higher cutoff is better for identifying large clusters. Table 13 shows the Moran's *I* analysis for *Bungis/Pumpkin* at ten different cutoff distances from 20 kilometers to 200 kilometers (60 kilometers is roughly the size of a county, and 200 kilometers is about half the overall length of the German belt).




**Map 4.** Hotspot analysis for *Bungis/Pumpkin*

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

The Moran's *I* statistic ranges from −1 to 1. A score close to 0 indicates that the feature is scattered randomly throughout the region. 1 indicates the maximum level of regional clustering, and −1 indicates that the variable is maximally dispersed. For every tested cutoff distance, there is a highly significant result indicative of a pattern of regional clustering.

For local autocorrelation, I used the Getis-Ord Gi* statistic with ArcMap's hotspot analysis tool (Ord & Getis, 1995, 2001). For each data point, this statistic determines the degree to which it is surrounded by similar values. Regional clusters show up as "hot spots" or "cold spots." I ran the statistic on several fixed-band cutoff distances and found 60 kilometers to be representative of the general pattern. In Map 4, I created a boundary around points with a confidence interval of 95% or greater (I used the Thiessen polygon tool to partition the map into regions defined by distance to the closest data point). These can be thought of as isogloss boundaries for *Bungis* (hot spot) and *Pumpkin* (cold spot).

### 3.7 Spatial clustering of rst-variation

Returning to the *rst* variants, I performed the same analysis for each of the three words *Wurst*, *Donnerstag*, and *Haarbürste*. The global autocorrelation analysis for each word is presented in Table 14.

One of the three words, *Donnerstag*, shows a pattern consistent with a random distribution at every cutoff distance. We can conclude that the distribution of sibilant variants in *Donnerstag* is homogenous in the Texas German belt.

The other two words show a pattern consistent with randomness (or even dispersal) at higher resolutions, but they both show significant results for clustering below 60 kilometers. This means that there is no widescale pattern, but we may find significant clustering at the level of particular counties or settlements. I ran a hotspot analysis on *Wurst* and *Haarbürste* with a fixed band cutoff corresponding to the highest Moran's *I* p-value (20 kilometers for *Wurst* and 40 kilometers for *Haarbürste*) to maximize the potential to pick out hot spots.

The only clustered areas are in the west. The cold spot in the south is somewhat outside the German belt and corresponds to single judgments, so it may be considered noise. The cold spot in the north is centered near the settlement of Fredericksburg, and the hot spot is centered near the settlement of New Braunfels. These correspond to relatively low and high percentages of the ʃ-variant in *Haarbürste* in these communities. We find the reverse situation for *Wurst*.

Again, there are cold spots corresponding to single judgements from speakers born outside the traditional area of the German belt. These can be excluded from analysis. We find a cold spot centered around New Braunfels and a hot spot centered around Fredericksburg, corresponding to relatively low and high percentages of the ʃ-variant for *Wurst* in these communities.

The two words for which we do have some evidence for spatial clustering are both words in which the overall average percentage for the ʃ-variant is about 95%. In Fredericksburg, the relative cold spot for *Haarbürste* corresponds to a proportion of 91%, while in New Braunfels, the relative cold spot for *Wurst* corresponds with a proportion of 83%. This low result is somewhat surprising considering that Boas (2009) found a proportion of 94%. In any case, these communities do not appear to have substantially different patterns of *rst* variation.

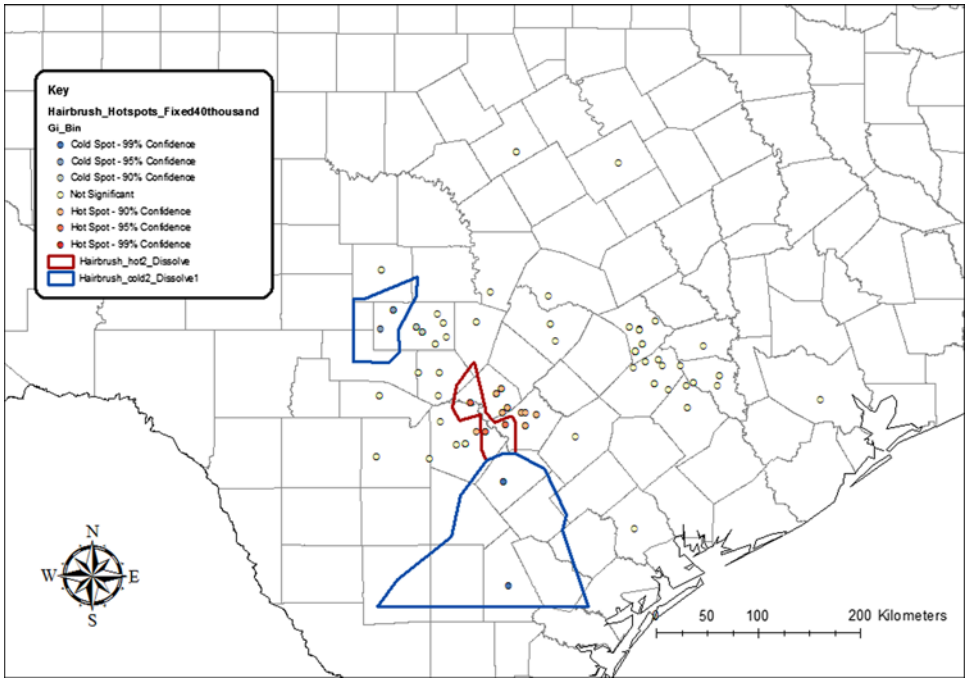
Considering the geographic distribution of German donor dialects to central Texas, we would expect broad differences between east and west, as was clearly evident in the variant lexical items for "pumpkin." The eastern settlements should have a higher percentage of the s-variant because more settlers came from northern Germany to these areas, while the ʃ-variant should predominate in the western settlements. We might also expect the southeast to have a different pattern due to the influence of Alsatian German. We do not find this, but rather a much more homogenous pattern, particularly in the east, with marginal differences between New Braunfels and Fredericksburg in the west.

I conclude that for *rst* variation, the distribution of variation is reasonably uncorrelated with geographic region. This variation looks more like stable variation in a homogenous Texas German variety than variation between subvarieties.

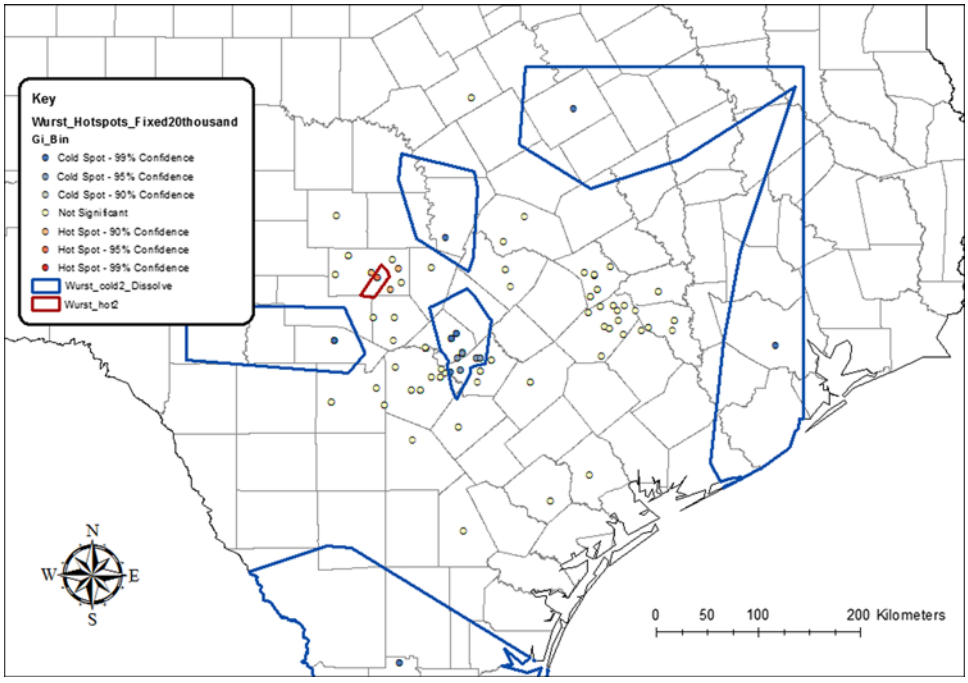
**Table 14.** Moran's *I* Analysis for *Wurst, Donnerstag, and Haarbürste*

| Word | Cutoff (km) | Moran's *I* | z-score | p-value | Result |
|---|---|---|---|---|---|
| Wurst | 20 | 0.024888 | 3.375601 | 0.000737 | highly clustered |
| Wurst | 40 | 0.007624 | 1.637967 | 0.101429 | random |
| Wurst | 60 | 0.010968 | 2.462766 | 0.013787 | clustered |
| Wurst | 80 | −0.002704 | 0.001356 | 0.998918 | random |
| Wurst | 100 | −0.007685 | −1.663147 | 0.096283 | slightly dispersed |
| Wurst | 120 | −0.004355 | −0.00271 | 0.417146 | random |
| Wurst | 140 | −0.005356 | −1.511036 | 0.130779 | random |
| Wurst | 160 | −0.004537 | −1.170381 | 0.241848 | random |
| Wurst | 180 | −0.00554 | −1.971882 | 0.048623 | dispersed |
| Wurst | 200 | −0.003101 | −0.336099 | 0.736796 | random |
| Donnerstag | 20 | −0.010989 | −0.383959 | 0.701009 | random |
| Donnerstag | 40 | −0.014479 | −0.789144 | 0.430028 | random |
| Donnerstag | 60 | −0.015409 | −0.994469 | 0.319994 | random |
| Donnerstag | 80 | −0.012347 | −0.846987 | 0.397002 | random |
| Donnerstag | 100 | −0.00779 | −0.508061 | 0.61141 | random |
| Donnerstag | 120 | −0.006698 | −0.469308 | 0.638849 | random |
| Donnerstag | 140 | −0.004648 | 0.061823 | 0.950704 | random |
| Donnerstag | 160 | −0.005101 | 0.083419 | 0.933519 | random |
| Donnerstag | 180 | −0.007736 | −1.078405 | 0.280853 | random |
| Donnerstag | 200 | −0.00665 | −0.857155 | 0.391359 | random |
| Haarbürste | 20 | 0.019814 | 1.902257 | 0.057138 | slightly clustered |
| Haarbürste | 40 | 0.026318 | 3.049959 | 0.002289 | highly clustered |
| Haarbürste | 60 | 0.018005 | 2.543024 | 0.01099 | clustered |
| Haarbürste | 80 | 0.004375 | 1.194792 | 0.232168 | random |
| Haarbürste | 100 | −0.001207 | 0.666598 | 0.505029 | random |
| Haarbürste | 120 | −0.004391 | 0.032663 | 0.973943 | random |
| Haarbürste | 140 | −0.006171 | −0.547831 | 0.583808 | random |
| Haarbürste | 160 | −0.001924 | 0.947839 | 0.343211 | random |
| Haarbürste | 180 | −0.000887 | 1.471142 | 0.141253 | random |
| Haarbürste | 200 | −0.00665 | −0.857155 | 0.391359 | random |
| Haarbürste | 140 | −0.006171 | −0.547831 | 0.583808 | random |

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4




**Map 5.** Hotspot analysis for *Haarbürste*




**Map 6.** Hotspot analysis for *Wurst*

## Conclusion

The presence of variation in a particular linguistic variety may be correlated with many different factors. For the variable pronunciation of *rst*-clusters in Texas German, I have examined the possibility that it may be correlated with particular lexical items or with some speaker feature such as age, gender, or place of birth. Texas German is an endangered dialect, and some variation should be expected due to the chaotic processes of language attrition. For this particular feature, however, and for other phonological features examined by Boas (2009), the diachronic tendency has been in the direction of uniformity. Boas emphasizes the striking lack of increased variability that might be expected in a situation of language death (Sasse, 1992): most of the trends of phonological variation described by Gilbert (1972) are found in Boas (2009), and in fact leveling processes have continued in the same direction, but are not yet completed. The ʃ-variant has virtually replaced

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

the s-variant for some lexical items, continuing a trend noted in Gilbert (1972).

It is clear that different lexical items are associated with different levels of variation among the three lexical items *Haarbürste*, *Donnerstag*, and *Wurst*. This pattern has been extended to other words with this cluster, as in *erst*. An examination of intraspeaker variation in the open interviews indicates that the speakers are not a homogenous group: in regular speech some speakers appear to use the majority form exclusively, while other speakers use the markers variably. This heterogeneity is reminiscent of East Sutherland Gaelic, in which there is stable local variation.

I found no statistical correlation between gender or age and the expression of particular variants. Taken together, these findings suggest a lack of diachronic change for these variants in the recent history of Texas German. If this is broadly the case, then the variation present in modern-day Texas German is relatively stable. This can be taken either as an argument that the dialect formation process has completed and converged on a variable pronunciation, or else that the process was interrupted before a final invariant pattern could emerge.

I have argued that the question of whether a group of dialects have cohered into a new dialect cannot be answered by the mere presence of variation, but should rather be answered by the geographic distribution of variant proportions. This is because variation of many types can exist in a homogenous language community. The factors that allow for a minority variant to survive as a stable variant in the new dialect remain to be explored, but the important issue is that the end state should be measured by the sufficient diffusion of variant proportions across the dialect region.

The geographic analysis of variation in *rst*-clusters show a pattern characteristic of a randomly scattered feature, with no regions with statistically higher or lower proportionality. An expected pattern for non-convergence would be multiple hotspots corresponding to particular settlement patterns, and by and large this is not what I found with this feature. However, the current study has explored a single variant in depth, and it would be necessary to look at other variant features in order to determine whether Texas German appears to be more or less homogenous. Certain lexical items such as the word for "pumpkin" are clearly associated with particular regions of the German belt. It would also be helpful to compare this analysis with other linguistic enclaves to see what a thoroughly heterogenous dialect community would look like. The analysis of this particular feature demonstrates how these techniques may be applied to other variant features in Texas German in order to provide an overall picture of the uniformity or heterogeneity of the dialect region.

## Endnotes

**1** http://www.arcgis.com

**2** The maps are available online at http://regionalsprache.de.

**3** In the original Wenker survey forms upon which the maps are based, one of the questions does contain the word *Bürste* ('brush', cf. *Haarbürste* 'hairbrush'), so an isogloss could be constructed from the original survey data. My impression is that the isogloss would be similar to *Wurst* but extend further south to encompass more of the Hessen-Nassau Region. There is evidence from later extension maps of southern Germany and Austria that there is substantial variation between s and ʃ-variants of *Donnerstag* in the south. These brief observations are in line with the difference in proportions for *Wurst*, *Donnerstag*, and *Haarbürste* found in Table 3.

**4** Downloaded from the archives and available at http://speechislands.org.

## REFERENCES


Akaike, Hirotugo. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6). 716–723.

Bates, Douglas, *et al.* 2014. "Package 'lme4'." *R foundation for statistical computing*, Vienna 12.

Baxter, Gareth J., Richard A. Blythe, William Croft & Alan J. McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change 21*(2). 257–296.

Biesele, Rudolph Leopold. 1930. *The history of the German settlements in Texas, 1831–1861*. Austin, TX: Press of Von Boeckmann-Jones Company.

Boas, Hans C. 2009. *The life and death of Texas German (American Speech: A Quarterly of Linguistic Usage)*. Durham: Duke University Press.

Boas, Hans C. 2002. The Texas German Dialect Archive as a tool for analyzing sound change. In Peter Austin, Helen Aristar-Dry & Peter Wittenburg (eds.), *Proceedings of the International Workshop on Resources and Tools in Field Linguistics*, 28.1–28.4.

Boas, Hans C., Karen Ewing, Cheryl Moran & Jana Thompson. 2004. Towards determining the influence of internal and external factors on recent developments in Texas German phonology. *University of Pennsylvania Working Papers in Linguistics 10*(1). 47–59.

Boas, Hans C., Marc Pierce, Hunter Weilbacher,Karen Roesch, and Guido Halder. 2010. The Texas German Dialect Archive: A multimedia resource for research, teaching, and outreach. *Journal of German Linguistics 22*(3). 277–296.

Clardy, Catherine. 1954. *A Description and Analysis of German Language Spoken in New Braunfels, Texas*. Austin, TX: University of Texas at Austin PhD thesis.

Dorian, Nancy C. 1973. Grammatical change in a dying dialect. *Language 49*. 413–438.

Dorian, Nancy C. 1983. Natural second language acquisition from the perspective of the study of language death. In R.W. Anderson (ed.) *Pidginization and creolization as language acquisition*, 158–167. Rowley, MA: Newbury House.

Dorian, Nancy C. 1994. Varieties of variation in a very small place: Social homogeneity, prestige norms, and linguistic variation. *Language 70*(4). 631–696.

Eikel, Fred. 1949. The use of cases in New Braunfels German. *American Speech 24*(4). 278–281.

Eikel, Fred. 1954. *The New Braunfels German Dialect*. Baltimore, MD: Johns Hopkins University PhD thesis.

Eikel, Fred. 1966a. New Braunfels German: Part I. *American Speech 41*(1). 5–16.

Eikel, Fred. 1966b. New Braunfels German: Part II. *American Speech 41*(4). 254–260.

Eikel, Fred. 1967. New Braunfels German: Part III. *American Speech 32*(2). 83–104.

Gilbert, Glenn G. 1963. *The German Dialect Spoken in Kendall and Gillespie Counties*. Austin, TX: University of Texas Press.

Gilbert, Glenn G. 1965. English loanwords in the German of Fredericksburg, Texas. *American Speech 40*(2). 102–112.

Gilbert, Glenn G. 1972. *Linguistic Atlas of Texas German*. Austin, TX: University of Texas Press.

Grieve, Jack, Dirk Speelman & Dirk Geeraerts. 2011. A statistical method for the identification and aggregation of regional linguistic variation. *Language Variation and Change 23*(2). 193–221.

Guion, Susan. 1996. The death of Texas German in Gillespie County. In P. Sture Ureland & Iain Clarkson (eds.), *Language contact across the North Atlantic*, 443–463. Tübingen: Max Niemeyer Verlag.

Johannessen, Janne B. & Joseph C. Salmons. 2015. *Germanic Heritage Languages in North America: Acquisition, Attrition and Change (Studies in Language Variation, 18)*. Amsterdam: John Benjamins Publishing Company.

Jordan, Terry G. 1977. The German element in Texas: An overview. *Rice Institute Pamphlet-Rice University Studies 63*(3).

Jordan, Terry G. 1966. *German seed in Texas soil: Immigrant farmers in nineteenth-century Texas*. Austin, TX: University of Texas Press.

Kerswill, Paul. 2002. Koinéization and accommodation. In J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The Handbook of Variation and Change*, 669–701. Malden, MA: Blackwell.


Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4

Kerswill, Paul & Peter Trudgill. 2005. The birth of new dialects. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect change: Convergence and divergence in European languages*, 196–220. Cambridge: Cambridge University Press.

Nützel, Daniel C. 1998. *Language death and morphological decay: The case of Haysville East Franconian*. West Lafayette, IN: Purdue University Phd thesis.

Nützel, Daniel C. 2009. *The East Franconian dialect of Haysville, Indiana: A study in language death*. Regensburg: Edition Vulpes.

Nützel, Daniel C. & Joseph C. Salmons. 2011. Language contact and new dialect formation: evidence from German in North America. *Language and Linguistics Compass 5*(10). 705–717.

Ord, J. Keith & Arthur Getis. 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis 27*(4). 286–306.

Ord, J. Keith & Arthur Getis. 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science 41*(3). 411–432.

Pierce, Marc, Hans C. Boas & Karen Roesch. 2015. The history of front rounded vowels in New Braunfels German. In Janne Bondi Johannessen & Joseph C. Salmons (eds.), *Germanic heritage languages in North America*, 117–132. Amsterdam: John Benjamins Publishing Company.

R Core Team. 2017. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL https://www.R-project.org/.

Roesch, Karen A. 2012. *Language maintenance and language death: The decline of Texas Alsatian*. Amsterdam: John Benjamins Publishing.

Salmons, Joseph C. 1983. Issues in Texas German language maintenance and shift. *Monatshefte 75*. 187–196.

Salmons, Joseph C. & Felecia A. Lucht. 2006. Standard German in Texas. In Linda Thornburg-Panther & Janet M. Fuller (eds.), *Studies in contact linguistics: Essays in honor of Glenn G. Gilbert*, 165–86. New York: Peter Lang Inc., International Academic Publishers

Sasse, Hans-Jürgen. 1992. Theory of language death. *Contributions to the Sociology of Language 64*. 7–7.

Schmidt, Jürgen Erich, Joachim Herrgen & Roland Kehrein (eds.). 2008. *Regionalsprache.de (REDE). Forschungsplattform zu den modernen Regionalsprachen des Deutschen. [Research platform for the modern German regional languages]*. Prepared by Dennis Bock, Brigitte Ganswindt, Heiko Girnth, Simon Kasper, Roland Kehrein, Alfred Lameli, Slawomir Messner, Christoph Purschke & Anna Wolańska. Marburg: Forschungszentrum Deutscher Sprachatlas.

Siegel, Jeff. 2001. Koine formation and creole genesis. *Creole Language Library 23*. 175–198.

Tamminga, Meredith. 2013. Phonology and morphology in Dutch indefinite determiner syncretism: Spatial and quantitative perspectives. *Journal of Linguistic Geography 1*(2). 115–124.

Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. New York: Oxford University Press.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. New York: Oxford University Press.

Wenker, Georg. 1888–1923. Sprachatlas des Deutschen Reichs. Handgezeichnetes Original von Emil Maurmann, Georg Wenker und Ferdinand Wrede [Linguistic atlas of the German Empire. Hand-drawn original by Emil Maurmann, Georg Wekner and Ferdinand Wrede]. Marburg. Publiziert als Digitaler Wenker-Atlas (DiWA) [Published as the Digital Wenker-Atlas (DiWA)]. www.regionalsprache.de.

Wells, Christopher. J. 1985. *German, a linguistic history to 1945*. New York: Oxford University Press.

Wilson, Joseph B. 1960. The Texas German of Lee and Fayette counties. *Rice Institute Pamphlet 47*. 83–98.

Wilson, Joseph B. 1977. The German language in central Texas today. In Joseph Wilson (ed.), *Texas and Germany. Crosscurrents 63*(3), 47–58. Houston, TX: Rice University Studies.

Wood, Jim, Laurence Horn, Raffaella Zanuttini & Luke Lindemann. 2015. The Southern Dative Presentative meets Mechanical Turk. *American Speech 90*(3). 291–320.

Zanuttini, Raffaella & Laurence Horn. 2014. *Micro-syntactic variation in North American English*. New York: Oxford University Press.

Zanuttini, Raffaella, Jim Wood, Jason Zentz & Laurence Horn. 2018. The Yale Grammatical Diversity Project: Morphosyntactic variation in North American English. *Linguistics Vanguard 4*(1). DOI: https://doi.org/10.1515/lingvan-2016-0070

Downloaded from https://www.cambridge.org/core. University of Texas Libraries, on 29 Oct 2021 at 14:18:33, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jlg.2019.4