Listener Factors in Accent Recognition: A Perceptual-dialectology Study of Frisian

Cesko Voeten, Anne-France Pinget, Martijn Kingma, Nika Stefan, and Hans Van de Velde*

1 Introduction

Within the field of perceptual dialectology, a significant amount of research on dialect recognition has focused on effects related to the *speaker* (see Montgomery 2012 and McKenzie et al. 2019 for a review). The current study illustrates a growing tendency to take *listener*-related factors into consideration to broaden our understanding of dialect recognition and its more general, crosslinguistic nature. We investigate how listeners' social profiles impact dialect recognition by looking at a small and rather cohesive language situation, i.e., Frisian as spoken in the province of Fryslân in the Netherlands, using the map-based accent recognition task (Pinget and Voeten 2023, Voeten and Pinget submitted).

Based on prior research reviewed in section 1.2., we know that listeners' representations of regional language variation are shaped in a complex way by the listeners' familiarity with the different regional varieties. Moreover, there is some evidence that listeners' recognition aptitudes are further modulated by their individual characteristics, for instance their age, level of education, and geographical knowledge. This study aims to refine our understanding of the role of listeners' social factors in dialect recognition. Before discussing these factors, the Frisian language situation is introduced.

1.1 The Language Situation in Fryslân: West Frisian

Western Frisian (ISO 639-2: *fry*) refers to a group of West Frisian dialects spoken in the north of the Netherlands, and is commonly called *Frisian* (we also use this term in the current study). Frisian is spoken by about 430.000 people (Centraal Bureau voor de Statistiek 2023, Provincie Fryslân 2020) in the province of Fryslân (5,749 km^{2}; see Figure 1).

The linguistic situation in Fryslân is complex. Traditionally, three major regional varieties of Frisian are distinguished (Bloemhoff et al. 2014): *Klaaifrysk* (Clay Frisian), *Wâldfrysk* (Wood Frisian) and *Súdwesthoeksk* (South-west Corner Frisian), which differ mostly at the lexical and phonetic level. Some Frisian linguists further distinguish two subvarieties within *Klaaifrysk*: Main Clay Frisian and *Noardhoeksk* (Northern Corner Frisian; e.g., Duijff 2008). In addition to the main Frisian dialects, there are some smaller, archaic Frisian dialects (e.g., in Hylpen and on the islands Skylge and Schiermûntseach), a number of Frisian-Dutch contact varieties (e.g., Town Frisian) and Low-Saxon dialects spoken in Fryslân (Bloemhoff et al. 2014). The geographical distribution of these language varieties in the province of Fryslân is shown in Figure 1. The current study focuses on the main (mainland) Frisian varieties (in green). This means that archaic dialects, such as that of Hylpen (Versloot 2020) and contact varieties are excluded.

Apart from being complex, the Frisian language situation is also characterized by strong local varieties (i.e. varieties more actively used as compared to the rest of the Netherlands or parts of Europe), because there is little pressure from a spoken standard, nor from the written standard (as the vast majority does no read nor write Frisian; Stefan 2022). Despite a moderate body of work on Frisian dialectology (e.g., Boelens and van der Woude 1955, Hilton et al. 2020, Hof 1933, van der Veen 2001), there is to the best of our knowledge no available insight about how variation in Frisian is perceived and used by language users. In this study, we therefore

*We gratefully acknowledge Hielke Vriesendorp, who provided helpful comments and critiques on an earlier version of the manuscript, and Johan Feikens who created the maps. We thank the colleagues from the Digital Humanities Lab at Utrecht University, especially Ben Bonfil who helped with setting up the experiment. Moreover, we are thankful to our colleague Femke van Seijen for assisting in setting up this experiment and Mettje de Vries and the public broadcast station *Omrop Fryslân* for their help with advertising the experiment.

The Faculty Ethics Assessment Committee of the Faculty of Humanities at Utrecht University has reviewed the research project 'Accentherkenning - Friese taal' and approved the study (FEtC-H reference number: 21-161-03, 29 November 2021).

investigate the accuracy with which Frisian listeners are able to infer regional origins in speech perception, and whether this accuracy differs for different types of listeners. In this study, the focus lies on phonetic variation, which is the largest component of regional variation in Fryslân (Duijff 2008). Frisian dialects of course further display lexical and morpho-syntactic variation, but our study excludes these types of variation from consideration. Consequently, we will further use the term 'accents' to refer to the Frisian varieties used in this study.

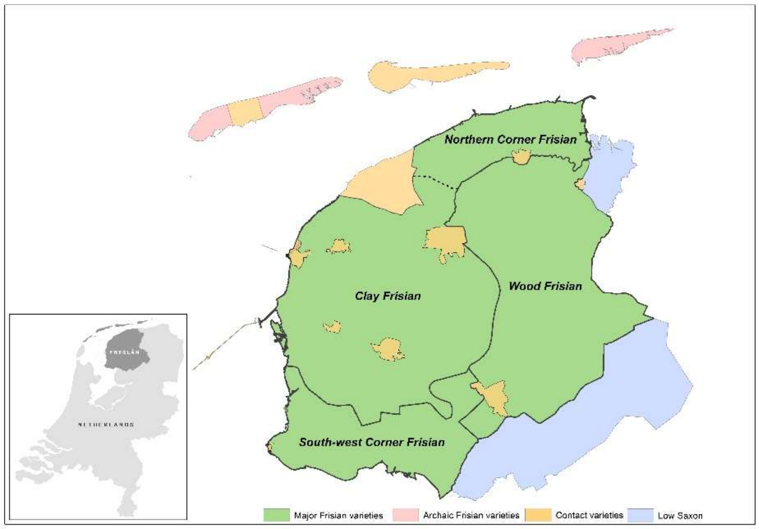

Figure 1: The geographical distribution of the language varieties in the province of Fryslân. Colors indicate the different types of language varieties; labels within the green area distinguish the four main varieties within Frisian.

1.2 Listener Factors in Accent Recognition

The most important listener-related predictor of accent recognition is arguably the listener's familiarity with the target variety. While it is quite straightforward to observe that the degree of familiarity directly facilitates recognition (as it has been done in many studies, e.g., McKenzie 2015, McKenzie et al. 2019, Montgomery 2012, Pinget and Voeten 2023, Ryan 1983), investigating which type of familiarity matters is complicated, since the target variety can be familiar to listeners for many different reasons (Kerswill and Williams 2002). Pinget and Voeten 2023 distinguished four types of familiarity: (1) being a speaker of the variety; (2) being in extended contact with speakers of the variety (i.e., geographical mobility); (3) knowing an individual speaker of the variety; (4) being confronted with the variety through mass or social media. Studies have reported effects for each specific type of familiarity (e.g., Williams et al. 1999, Diercks 2002, Purschke 2012), but considering the more rural nature of Fryslân (compared to the rest of the country), the second type of familiarity could be the most important one in this study. To start with, it is necessary to include a measure of the distance between the speaker and the listener as an estimate of the effect of proximity. On top of that, previous studies have demonstrated relatively high levels of accurate recognition among listeners with greater levels of experience with dialects within the country and/or geographical mobility than less experienced/mobile listeners (Baker et al. 2009, Clopper and Pisoni 2004, McKenzie et al. 2019). Hence, it is hypothesized that the geographical distance between speaker and listener origin directly affects accent recognition because of the way speakers live, travel, and keep in touch with each other. Furthermore, in a region like Fryslân, geographical distance is a good proxy for linguistic distance (cf. Wieling, Nerbonne, and Baayen 2011 for Dutch in the whole country).

Besides one's familiarity with a variety, accent recognition can be modulated by a range of factors, including the listener's age, gender and education. A few studies have looked at the effect of listeners' age on their abilities to identify speakers' places of origin (Diercks 2002, Floccia et al. 2009, Girard et al. 2008, Jones et al. 2017, Pinget and Voeten 2023, Williams et al. 1999). Jones et al. 2017 presented samples of American English varieties to child and adult listeners and asked them to group the talkers by region in a free-classification task. While they showed that some of the youngest listeners perceived dialect variation, they did not observe adult-like accuracy before the age of sixteen or seventeen years. Diercks 2002 demonstrated that German teenagers and young adults (15–29 years old) displayed a lower ability to recognize German dialects than older adults (30–64 years old). Pinget and Voeten 2023 found a curvilinear effect of age on dialect recognition, such that middle-aged listeners showed more accurate recognition patterns than young adults and older listeners, and interpreted this pattern as a reflection of age-related changes in geographical and social mobility.

Accent recognition can also be influenced by the listener's gender and education level. Kerswill and Williams 2002 systematically varied judges by social class (working versus middle class) and social network (local network versus nonlocal network, thus geographically mobile) in the English context. They showed that it was network, rather than social class, that was the decisive factor in recognition. Alternatively, and especially in the Frisian situation, we expect the listener's education level to be a crucial factor in the amount of language input one is confronted with outside one's own place of origin. In the Netherlands, sociological research has shown a strong correlation between education level and social and geographical mobility (Hensen et al. 2009; van de Werfhorst 2002). Concerning gender, women are traditionally seen as poor recognizers, because they tend to be less mobile both socially and geographically than men. In his German dialect recognition study, Diercks 2002 included gender as a factor and found – against expectations – that female listeners scored better. However, because there was a confound between gender and mobility, his results were difficult to interpret. In this study, we therefore include education level, gender and age and try to disentangle their effects.

The principal aim of this study is to investigate to what extent these factors play a role in the recognition of Frisian accents. In order to answer this research question, we use the same methodology as in Pinget and Voeten 2023, i.e., a map-based accent-recognition task. As explained in section 1.1, the Frisian linguistic case is quite complex and the geographical and social community can be said to be 'tight-knit'. We hypothesize that this might result in better accent recognition. Moreover, we expect the above-mentioned listeners' factors to play a role in recognition patterns.

2 Method

2.1 Stimuli

Forty speakers were selected from 20 locations spread over the Frisian language area. In each location, we recorded a young adult (18–30 years old) and an older adult (55–80 years old), born and raised, and still living in the location at the time of recording.

The sites were geographically well-spread and reflected variation within the Frisian dialects. The selection of sites can be found in Figure 2. Seven sites represent Wood Frisian, four sites represent the Southwestern variety, and nine sites represent Clay Frisian, of which six lie in the main Clay Frisian area and three in the Northern Corner.

A spontaneous conversation was conducted with all speakers, either online (winter 2021–2022; via Microsoft Teams because of COVID restrictions) or face-to-face (summer 2022). There was no apparent difference in the sociolinguistic situation and type of speech recorded in these two settings. Stimuli consisted of an extract of around 20 seconds (the usual length of stimuli in this type of study, see also Watson and Clark 2015). The content of the stimuli was carefully selected so that it contained no information that could be linked to a geographical area or any specific place, no indication about the age, social class and regional origin of the speakers, no noticeable hesitations or slips of the tongue and no disturbing background noise. The extracted sound files were levelled to 75 dB SPL.

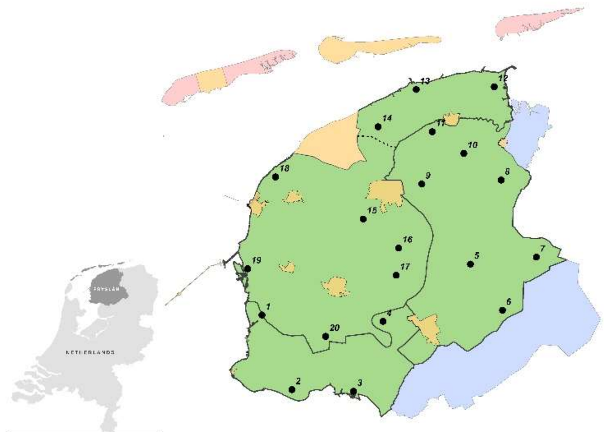

Figure 2: The geographical distribution of the twenty sites in the province of Fryslân. Colors indicate the different language varieties in the province (adapted from Bloemhoff et al, 2014: p.721). Place names in Frisian (Dutch name between brackets): [1] Warkum (Workum), [2] Aldemardum (Oudemirdum), [3] De Lemmer (Lemmer), [4], De Jouwer (Joure), [5] Nij Beets (Nij Beets), [6] Jobbegea (Jubbega), [7] Bakkefean (Bakkeveen), [8] De Harkema (Harkema), [9] Tytsjerk (Tietjerk), [10] De Westereen (Zwaagwesteinde), [11] Rinsumageast (Rinsumageest), [12] Eanjum (Anjum), [13] Holwert (Holwerd), [14] Hallum (Hallum), [15] Weidum (Weidum), [16] Grou (Grouw), [17] Akkrum (Akkrum), [18] Seisbierrum (Sexbierum), [19] Makkum (Makkum), [20] Wâldsein (Woudsend).

2.2 Listeners

In Spring 2023, data were collected from 1,610 listeners (766 men, 824 women, 5 other, and 15 preferred not to say) with Frisian as one of the native languages[1] and living in the province of Fryslân. The age range was between 18 and 78 years ($M = 46$, $SD = 18$, based on the participants' self-provided birth years). Participants filled in the highest level of education they completed: either primary school ($N = 9$), preparatory vocational secondary education (Dutch *vmbo*; $N = 85$), senior general secondary education (Dutch *havo*; $N = 119$), university preparatory secondary education (Dutch *vwo*; $N = 44$), vocational education and training (Dutch *mbo*; ($N = 500$), higher professional education (Dutch *hbo*; $N = 660$), or research-oriented university education (Dutch *wo*, $N = 193$). The origin of the listeners was determined by asking them for the four-digit postcode of their place of residence. The task was disseminated through social media and the regional broadcaster to reach as many people as possible. Listeners originated from all over the province, with a slight bias towards the eastern area (Wood area). Listeners also provided their self-estimated geographical knowledge (on a scale from 1 to 10).

2.3 Procedure

Participants did an online map-based accent-recognition task using Qualtrics, as designed by Pinget and Voeten 2023 and developed by Research Software Lab (Centre for Digital Humanities, Utrecht University 2024).[2] They were instructed to assign each speech stimulus to the site where they thought the speaker came from by clicking on a map provided by Google Maps.

As explained by Pinget and Voeten (2023), the Google Maps interface being by far the most popular mapping application in the Netherlands has several advantages. First, we can assume that participants have at least some experience with the application design and functionalities (e.g., the zoom-in and zoom-out functions). Second, Google Maps provides a lot of historical-political and geographical information (e.g., province lines and city locations)

[1]Speakers of Bildts, Town Frisian, and Stellingswerfs were excluded from the study, except if they also stated one of the varieties of main Frisian as one of their native languages.

[2]All materials for the map-based accent-recognition experiment, including the Qualtrics code, the sound stimuli, and the practice trials are available at https://zenodo.org/records/10890009.

familiar to listeners, which they can use as support when localizing the accent that they heard. Participants could not click on a location situated outside the province of Fryslân, nor in the water. The zoom level used by participants while clicking on a location was recorded during the experiment. This was recorded as a number ranging from 1 (fully zoomed out) to 19 (zoomed in down to the street level), with a zoom level of 9 being the default zoom setting.

Eight semi-randomized lists, each containing ten speakers balanced for gender and locality, were created. Each listener was asked to identify the origins of all ten speakers from one of the lists, presented in random order. Participants could choose to take more than one list (around 10% of them did so), but for each participant we analyzed only the first list.

Feedback was given to participants about the accuracy of their response after each item: responses within a range of 10 km were marked as correct and for all other responses the distance in kilometers between the given answer and the correct speaker origin was given. Each participant received a final score (out of 10), which made the experiment similar to a game or quiz. The gamification of the experimental design probably contributed to its attractiveness to a large pool of participants.

2.4 Data Analysis

The data were analyzed[3] as in Voeten and Pinget (2023); we refer to that publication for more details. We hence fitted a location-scale gamma generalized additive model with a logarithmic link, using the gam function from R package mgcv (Wood 2017) with the Extended Fellner-Schall optimizer (Wood and Fasiolo 2017). We based our analysis on all listener clicks that were inside mainland Fryslân; due to limitations in the statistical model we present here, accent placements on the Wadden islands were removed (24 observations out of 16,100). The dependent variable was the 'accent recognition error', defined as the distance in kilometers between the locations clicked by the listeners and the corresponding true speaker origins.

The predictors were an intercept, fixed effects for listener gender (sum-coded), fixed effects for listener education level (Helmert-coded[4]), as well as smooth terms for listener geographical knowledge, zoom level, age, and for the distance between the listener's self-reported geographical origin and the true speaker origin. These smooths were defined as thin-plate regression splines with ten basis functions. In addition, we included a spline on the sphere for the <latitude, longitude> coordinates of the true speaker origins, again including ten basis functions which were verified to be sufficient. We also included random effects by speakers: a random intercept, random slopes for the aforementioned fixed effects, and random smooths (with the customary first-derivative penalties) for all of the aforementioned thin-plate regression splines (but not the spline on the sphere for speaker origin, since that denotes a speaker-specific property, which are already handled by the random intercepts). Of the aforementioned predictors, all were used to model the dependent variable of 'accent recognition error'; furthermore, the subset of these predictors that correspond to geographical locations (i.e. the spline on the sphere for speaker origin, the thin-plate regression spline for the distance between the speaker and the listener, and the by-speaker random intercept and random smooth for the distance between the speaker and the listener) was included as predictors of the scale parameter of the model.[5]

For the parametric terms, inference was based on the significance of the *z*-score. For the smooth terms, we based our inference on the fitted splines. For the 1-D thin-plate regression splines, we predicted each spline onto a grid of 100 points; for the 2-D spline on the sphere, we predicted onto a 100×100-point grid spanning the entire mainland area of Fryslân, pruning away points outside its boundaries to yield a grid of 3,890 points. We then computed the model's predictions for these points as well as their 95% credible interval, based on the procedure by Wood (2017:293-294). Differences are considered significant where the 95% credible interval excludes zero.

---

[3]The data and analysis code are available at https://figshare.com/s/e303e946444173569e2d.

[4]The order of the categories for the factor 'education level' was primary school < vmbo < mbo < havo < vwo < hbo < wo.

[5]We refer to Voeten & Pinget's 2023 discussion of 'edge effects' for the background and necessity of this approach to dealing with heteroscedasticity, which is inherent to this type of data.

3 Results

The raw data of the accent recognition task (N = 16,074) are presented in Figure 3; the data have been binned along a 10×10-point grid to facilitate visualization. Each facet in Figure 3 plots the data for a particular set of two speakers from the same place; names of these places are in the facet labels, and they are indicated on the map by triangles. Observe that there is a strong tendency for listeners to place accents in the areas where they belong.

We discuss the effects of our predictors on the mean accent-recognition error in two parts: first, we discuss the parametric terms, i.e. the intercept and the factor predictors. Next, we discuss the smooth terms separately. In our interpretation of these effects, we rely on the 'accent-recognition-error ratio', henceforth 'ARE ratio'. This ratio is obtained by exponentiating the log-scale linear predictors. Its interpretation is similar to the odds ratio known from logistic regression: an ARE ratio of 1.01 means that a one-unit change in the predictor incurs a multiplication of the accent-recognition error in kilometers by 1.01, or a 1% relative increase. Conversely, an ARE ratio of 0.99 implies that the listener incurs 'only' 99% of the usual accent-recognition error, for every one-unit increase.

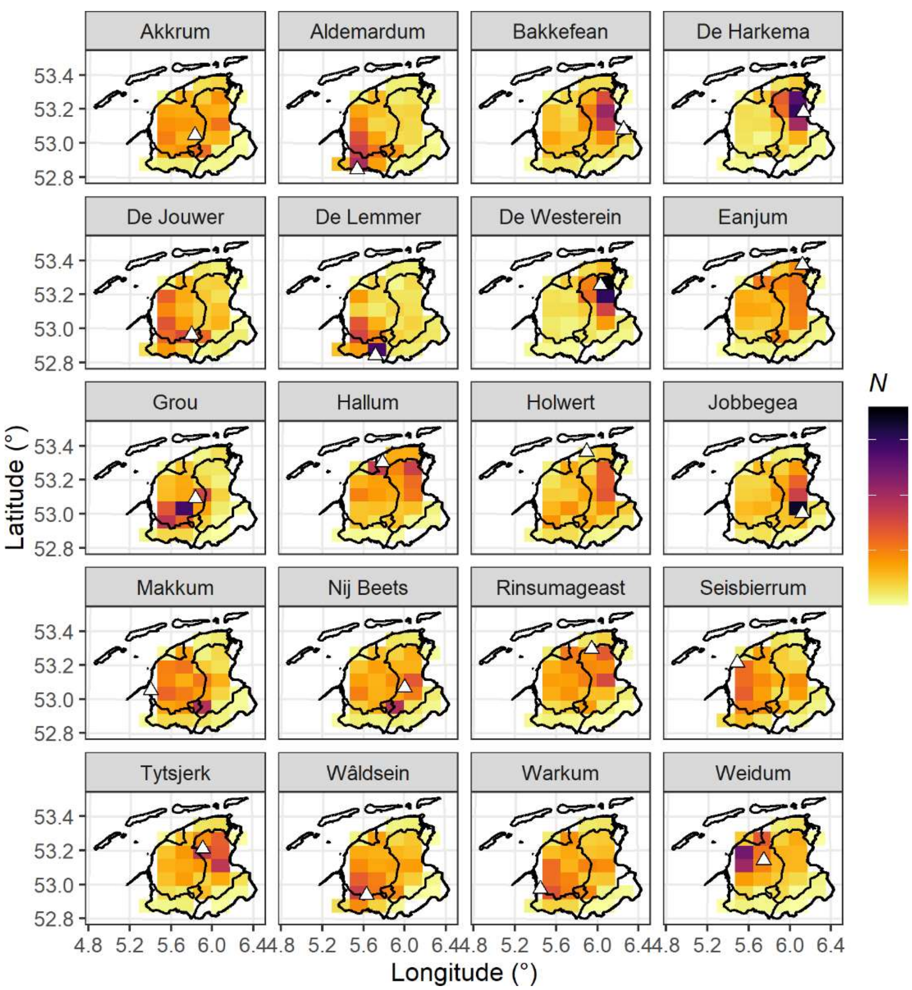

Figure 3: Absolute number of recognition clicks, split up by origin of the speakers. The site of origin itself is indicated with a white triangle.

3.1 The Parametric Terms

The parametric terms are summarized in Table 1. The intercept is placed at 3.07 log km, which corresponds to a baseline accent-recognition error of 21.46 kilometers. This is, however, the only parametric term that reaches significance: no other effects even come close.

Predictor	Estimate (SE)	ARE Ratio	*z*-value	*p*-value	Sig.
Intercept	3.07 (0.04)	21.46	70.49	<.001	***
Education = vmbo	0.01 (0.04)	1.01	0.30	0.77	
Education = mbo	0.003 (0.01)	1.00	0.27	0.79	
Education = havo	0.003 (0.01)	1.00	0.48	0.63	
Education = vwo	0.002 (0.01)	1.00	0.28	0.78	
Education = hbo	−0.002 (0.003)	1.00	−0.58	0.56	
Education = university	−0.0003 (0.003)	1.00	−0.11	0.92	
Gender = other	−0.04 (0.06)	0.96	−0.60	0.55	
Gender = prefer not to say	0.002 (0.04)	1.01	0.04	0.97	
Gender = male	0.01 (0.02)	1.01	0.60	0.55	

Table 1: Parametric coefficients.

3.2 The Smooth Terms

For reasons of space, we only discuss smooth terms that achieved significant effects along their domain; this turned out to be the case for all except for *geographical knowledge*. These significant effects are plotted in Figure 4. We discuss these effects in turn.

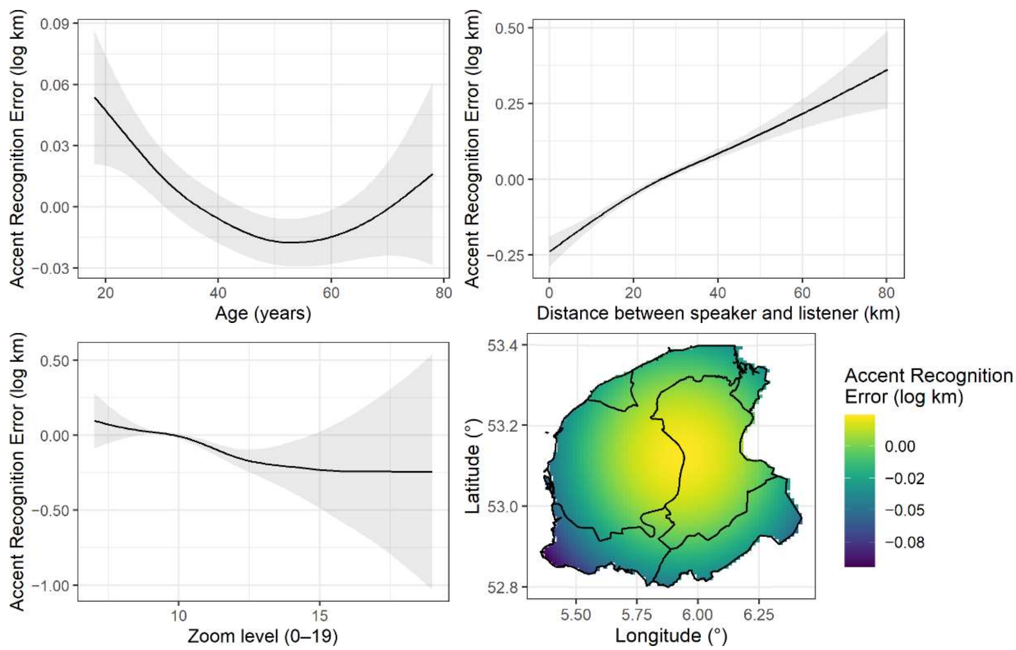

Figure 4: The smooth terms which, for some part of their span, reached significant effects. For the top and bottom-left plots, the 95% credible interval is indicated using a gray band along the trajectory; differences are significant where this band does not include zero.

First, the effect of *age* shows a curvilinear pattern (see Figure 4, top-left panel). Listeners younger than 30.5 years of age have significantly higher accent-recognition errors, with the highest value of +0.05 log km reserved for the youngest listeners of 18 years old. This corresponds to an ARE ratio of 1.06, which means that, overall, an 18-year-old listener has a 6% higher error in kilometers than average. Conversely, middle-aged listeners, here specifically those between 44 and 61.5 years old, have significantly lower accent-recognition errors, reaching a minimum of −0.018 log km at 53 years old. This corresponds to an ARE ratio of 0.98, or a 2% lower error than average.

Second, the *distance between speakers and listeners* has a very large and robustly significant effect, which is visualised in the top-right panel of Figure 4. When speakers and listeners come from the same site (insofar as the granularity of our postcode data allow us to say), the effect reaches its minimum of −0.24 log km, or an ARE ratio of 0.78. Conversely, the largest distance between a speaker and a listener in our data of 80 km leads to an effect of +0.36 log km, which corresponds to an ARE ratio of 1.44. This means that listeners' accent recognitions get anywhere from 22% smaller to 44% larger than average, depending on how close they are to the speaker's geographical origin. The progression of this effect between these extremes of nearly zero to 80 kilometers appears nearly linear on the log-km scale, which, at these values, results in similar near-linearity at the ARE-ratio scale. So, the closer the distance between the speaker and the listener, the better the accent is recognized.

Third, the effect of *zoom level* is significant in two areas in Figure 4 (bottom-left panel). Listeners who zoomed from −0.5 to +0.5 from the default value of 9 had significantly higher accent-recognition errors from +0.033 to 0.013 log km, which translates to ARE ratios from 1.034 to 1.013. In addition, listeners who zoomed in to zoom levels 10.4 to 14.6 incurred significantly smaller accent-recognition errors, an effect which ranged from −0.027 log km (ARE ratio = 0.97, 2.7% lower than average) to −0.223 log km (ARE ratio = 0.800, 20% lower than average).

Finally, the bottom-right plot in Figure 4 pictures the spline on the sphere coding for the *listener origins*. Overall, we observe that the highest accent-recognition errors emanate from a large area in the center of the province. By contrast, listeners whose regional origins are along the edges of the province achieve lower accent-recognition errors than average. The biggest difference, at the southwesternmost tip of the province, is one of −0.094 log km, which is an ARE ratio of 0.91, or a 9% improvement in accent recognition.

4 Discussion and Conclusion

We have conducted a large-scale online accent recognition experiment that in the Frisian language area. The aim of this study was to investigate how listener-related factors impact accent recognition in a small and socially cohesive language area.

First, we observed that Frisian listeners were on average relatively good at recognizing native accents, since the mean of the baseline accent-recognition error was 21 km. Combining this quantitative insight with the interpretation of the patterns in Figure 3, it seems that listeners at least have some awareness of the major varieties (Clay/Wood/Southwest-Corner Frisian), and in some cases they even can be more precise than that, indicating the exact village or a village nearby where the speaker actually came from. For a small and cohesive language area such as Fryslân, it therefore seems that the cognitive representation of accent variation is relatively precise. This evidence contrasts with for instance Bent and Holt 2017, who summarized the evidence available from several studies in the United States and the United Kingdom by saying that the cognitive representation of variation is relatively broad, showing only rough distinction along a single geographical axis (e.g., north versus south). However, it seems logical that listeners of a small and socially cohesive language area have – throughout their lifespan – obtained more experience with speakers of different varieties than listeners in a very large area (for instance, the United States). Consequently, their recognition patterns are more refined.

In order to further differentiate between 'better' and 'worse' listeners, we looked specifically at how listener's social profiles in terms of age, gender, level of education and place of origin impacted their abilities to map speakers to places. In contrast to Pinget and Voeten 2023 and Kerswill and Williams 2002, we did not find an effect of education, but we did find effects of listener age, place of origin, and distance to the speakers.

We found that age showed a curvilinear relationship with accent-recognition errors. Middle-aged listeners showed more accurate recognition patterns than young adults and older listeners. Even though this pattern only achieved significance in the middle of the age continuum (44–61.5 years), it seems quite stable and parallels the findings by Pinget and Voeten 2023. Both younger and older adults thus showed less sensitivity to variation than middle-aged adults, a pattern that can be interpreted as a reflection of age-related changes in geographical and social mobility (see Pinget and Voeten 2023 for an extensive discussion). In Fryslân, very similarly to the situation in the rest of the Netherlands, middle-aged adults are often in the middle of their working life. They behave themselves in large and solidified social networks, are exposed to more variation and hence show the highest awareness of language variation.

Furthermore, it was shown that the larger the distance between the listener's own place of origin and the speakers' place of origin, the more difficult accent recognition was. Listeners recognized the accent of their own place of origin relatively easily. This advantage in recognition decreases in a (nearly) linear manner as the distance to the speaker's origin increases. This effect compares to the effect of familiarity found in a larger language area in

Pinget and Voeten 2023 with the same methodology, and in studies like Baker et al. 2009 and Williams et al. 1999 with forced-choice categorization designs. These different types of evidence point at the fact that geographical distance between speaker and listener origins seems to be a very reliable predictor of accent recognition. At this stage, it would however be interesting to test whether this effect holds in much larger geographic spaces, such as the United States or Australia. Moreover, our perception of language variation is definitely affected by the salience of features, and specifically the relative ability of a linguistic variant to evoke regional meaning. In a future study, we would like to look at phonetic correlates, and examine how the relative salience of variables impacts listeners' recognition aptitudes.

Listener origin also turned out as a significant predictor of accent recognition: listeners from the central area of Fryslân turned out to be slightly worse at accent recognition. This result is in line with evidence provided by Pinget and Voeten 2023, where listeners from the central area of the Netherlands turned out to be slightly worse at accent recognition than listeners from the periphery of the country. In the case of Fryslân, the central area is considered to be a kind of standard: especially speakers from Grou have the reputation to be the 'closest to the standard' and the area is often used as a reference in dialectological study (e.g., Heeringa 2005). It seems unlikely, though, that speakers from this part of the province are exposed to less linguistic variability and are therefore less familiar with other accents. A way to further investigate this tendency could be by introducing a measure of mobility as a listener factor to control for how much input from other varieties a speaker/listener has experience with.

In conclusion, this study investigated how social factors impact accent recognition by looking at the Frisian language situation, which is small, complex, and rather socially cohesive. Results largely confirm patterns found in larger language areas. Clear effects of the distance between the speaker's and listener's origins were shown, and listeners recognized their own accents more easily than others. Moreover, listeners' recognition aptitudes were shown to be significantly determined by listeners' age and place of origin. All these listener-related factors modulate familiarity effects with the speakers in complex ways. Hence, listener-related factors seem to be crucial in shaping sensitivity to language variation and in fostering the capacity to map variation onto geographical representations.

References

Baker, Wendy, David Eddington, and Lyndsey Nay. 2009. Dialect identification: the effects of region of origin and amount of experience. *American Speech*, 84, 48–71.

Bent, Tessa and Rachael F. Holt. 2017. Representation of speech variability. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8, e1434.

Bloemhoff, Henk, Germen J. de Haan, and Arjen P. Versloot. 2014. Language varieties in the province of Fryslân. In: F. Hinskens and J. Taeldeman (eds). *Language and space: an international handbook of linguistic variation – Vol 3: Dutch* (pp. 721–738). De Gruyter Mouton.

Boelens, Krine, and van der Woude, Goasse (in collaboration with Klaas Fokkema and Edgard Blancquaert). 1955. *Dialect-atlas van Friesland*. De Sikkel.

Centraal Bureau voor de Statistiek. 2023. *StatLine – Bevolking op 1 januari en gemiddeld; geslacht, leeftijd en regio*. Retrieved on April 25, 2024 at https://opendata.cbs.nl/statline/#/CBS/nl/dataset/03759ned/table?fromstatweb.

Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111–140.

Diercks, Willy. 2002. Mental maps. In: D. Long, and D. Preston (eds). *Handbook of Perceptual Dialectology*. Volume 2 (pp. 51–70). John Benjamins.

Duijff, Pieter. 2008. Towards Standard Frisian in the Friesch Woordenboek. In: M. Mooijaart and M. van der Wal (eds). *Yesterday's words: contemporary, current and future lexicography* (pp. 53–66). Cambridge Scholars Publishing.

Floccia, Caroline, Joseph Butler, Frédérique Girard, and Jeremy Goslin. 2009. Categorization of regional and foreign accent in 5- to 7-year-old British children. *International Journal of Behavioral Development*, 33, 366–375.

Girard, Frédérique, Caroline Floccia, and Jeremy Goslin. 2008. Perception and awareness of accents in young children. *British Journal of Developmental Psychology*, 26, 409–433.

Heeringa, Wilbert. 2005. Dialect variation in and around Frisia; classification and relationships. *Us Wurk*, 54, 125–167.

Hensen, Maud M., Robert De Vries, and Frank Cörvers. 2009. The role of geographic mobility in reducing education-job mismatches in The Netherlands. *Papers in Regional Science*, 88, 667–682.

Hilton, Nanna H., Goffe T. Jensma, and Willem Visser. 2020. Frisian. In *Linguistic minorities in Europe online: A born-digital, multimodal, peer-reviewed online reference resource*. De Gruyter.

Hof, Jan J. 1933. *Friesche Dialectgeographie*. Martinus Nijhoff.

Jones, Zack, Qingyang Yan, Laura Wagner, and Cynthia G. Clopper. 2017. The development of dialect classification across the lifespan. *Journal of Phonetics*, 60, 20–37.

Kerswill, Paul, and Ann Williams. 2002. Dialect recognition and speech community focusing in new and old towns in England. In D. Long and D Preston (Eds), *Handbook of Perceptual Dialectology. Volume 2.* (pp. 173–204). Amsterdam/Philadelphia: John Benjamins.

McKenzie, Robert M. 2015. The sociolinguistics of variety identification and categorisation: Free classification of varieties of spoken English amongst non-linguist listeners. *Language Awareness,* 24, 150–168.

McKenzie, Robert M., Mimi Huang, Theng Theng Ong, and Navaporn Snodin. 2019. Socio-psychological salience and categorisation accuracy of speaker place of origin. *Lingua*, 228, 102705.

Montgomery, Chris. 2012. The effect of proximity in perceptual dialectology. *Journal of Sociolinguistics*, 16, 638–668.

Pinget, Anne-France and Cesko C. Voeten. 2023. Social factors in accent recognition: a large-scale study in perceptual dialectology. *Journal of Linguistic Geography*, 11, 78–90.

Provincie Fryslân 2020. *De Fryske Taalatlas 2020*. Retrieved on April 25, 2024 from https://www.fryslan.frl/flysystem/media/Fryske%20Taalatlas%202020%20%28Nederlands%29.pdf

Purschke, Christoph. 2012. *Regionalsprache und Hörerurteil: Grundzüge einer perzeptiven Variationslinguistik.* Franz Steiner.

Research Software Lab, Centre for Digital Humanities, Utrecht University. 2024. *Accent recognition: a perceptual-dialectological study of Frisian* (0.0.6). Zenodo. https://doi.org/10.5281/zenodo.10890009.

Ryan, Ellen B. 1983. Social psychological mechanisms underlying native speaker evaluations of non-native speech. *Studies in Second Language Acquisition*, 5, 148–159.

Stefan, Nika. 2022. *Spoken Frisian*. LOT.

van de Werfhorst, Herman G. 2002. A detailed examination of the role of education in intergenerational social-class mobility. *Social Science Information*, 41, 407–438.

van der Veen, Klaas F. 2001. West Frisian dialectology and dialects. In H. Munske (Ed), *Handbuch des Friesischen / Handbook of Frisian Studies* (pp. 98–115). Max Niemeyer.

Versloot, Arjen. 2020. Historical dialectology: West Frisian in seven centuries. In: S. D. Brunn & R. Kehrein (eds.). Handbook of the changing world map (pp. 405–442). Springer.

Voeten, Cesko C., and Anne-France Pinget. Submitted. *The online map-based accent-recognition task: a new pathway for perceptual dialectology*.

Watson, Kevin and Lynn Clark. 2015. Exploring listeners' real-time reactions to regional accents. *Language Awareness*, 24, 38–59.

Wieling, Martijn, John Nerbonne, and R. Harald Baayen. 2011. Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PloS One*, 6, e23613.

Williams, Angie, Peter Garrett, and Nikolas Coupland. 1999. Dialect recognition. In D. Preston (Ed.), *Handbook of perceptual dialectology. Volume 1.* (pp. 369–383). John Benjamins.

Wood, Simon N. 2017. *Generalized additive models: an introduction with R*. 2nd edn. Chapman & Hall/CRC.

Wood, Simon N., and Matteo Fasiolo (2017). A generalized Fellner-Schall method for smoothing parameter optimization with application to Tweedie location, scale and shape models. *Biometrics* 73, 1071–1081.

University of Pennsylvania
Department of Linguistics / Department of Biology
Philadelphia, PA, USA
cvoeten@sas.upenn.edu

University of Amsterdam
Amsterdam Center for Language and Communciation
Amsterdam, the Netherlands
c.c.voeten@uva.nl, m.m.kingma@uva.nl

Fryske Akademy
Ljouwert/Leeuwarden, the Netherlands
cvoeten@fryske-akademy.nl
apinget@fryske-akademy.nl
mkingma@fryske-akademy.nl
nstefan@fryske-akademy.nl
hvandevelde@fryske-akademy.nl

Utrecht University
Institute of Language Sciences,
Utrecht, the Netherlands
a.c.h.Pinget@uu.nl, h.vandevelde@uu.nl