MA Thesis

Study on self-evaluation of French regional accents

Chloé Vincent (190727659)

Queen Mary University of London

MA Linguistics - 2020

Supervisor: Devyani Sharma

Word count: 10 662

Study on self-evaluation of French regional accents

Abstract:

This thesis explores the attitudes towards regional accents in France. Many studies have been interested in the subject of accent evaluation in France, e.g. Kuiper (1999, 2005), Paltridge and Giles (1984), Pustka et al. (2019). However, none have focused on evaluations of French speakers' own accents.

I focus on self-evaluations of accents, in the sense of an accent well known to the participant, whether because they are themselves speaker of an accented variety or because they have lived in an area where they consider an accent can be heard.

Following recent attitude studies on metropolitan varieties of French, e.g. Avanzi and de Mareüil (2019), I conducted an online survey with Gorilla Experiment Builder (Anwyl-Irvine et al. 2019) and distributed it among native speakers of French through personal and social networks.

Mixed-effects models enables me to uncover what drive the evaluation of accent, previously grouped by emotional and status judgment thanks to a Principal Component Analysis. The main factors that are found significant are the region where the accent is spoken, the amount of remarks the participant received in their life: the more frequent the remarks (positive or negative) the higher the evaluation, and social class measures such as the socio-professional category of the participant: the higher the social-class, the lower the evaluation.

This study shows that the traditional hierarchies of accent are challenged: the Parisian standard is not the only accent highly rated on the status scale, while the non-standard varieties benefit from

mostly high emotional judgment. Moreover, it marks the importance of discrimination and power dynamics in the evaluation of an accent.

The data collected could potentially be used for later studies given that part of the data can be further explored, and many potential tracks were uncovered.

Introduction

Not two people in the world speak the same way. They differentiate by their language, but even within speakers of a same language variations are observed. Not only physical characteristics, such as the pitch, but also the way to pronounce some words or phrases, are a true identification of an individual. The way people speak may indicate where they grew up, what social class or social groups they belong to (Eckert 2012).

In this dissertation, I focus on the case of French regional accents and how they are perceived by French native speakers familiar with those accents.

France has a long tradition of prescriptivism around the unity of the nation based partly on the language unity (Encrevé and Braudeau 2007). Following the French Revolution, the influent Abbé Grégoire recommended the erasure of local languages spoken in France in favour of the French language as it is spoken in Paris. The French "republican ideal" requires the uniformity in language. Children were long forbidden to speak their own regional languages at school. This is still the case today concerning immigration languages especially (Cheshire 2018). Nowadays, accents are still frowned upon in higher instances of the state, on national media channels or in some branches of the industry: where the power lies (Bourdieu 1977).

Recent debates in France illustrated what Blanchet (2016) calls glottophobia: the discrimination because of an accent. Jean-Luc Mélanchon, a *député* (MP) and presidential candidate, made fun of a Toulousan journalist who asked a question he did not like, with a Southern accent. Marie-Arlette Carlotti, also *députée*, was also mocked when she spoke on television, twice on the same day, once with a standard accent to comment on the national news as a MP, and the other with a strong accent from Marseille as a candidate for the Mayor of Marseille. The very recent

appointment of Jean Castex as Prime Minister of France had many people talking: he is one of the very few high-level politicians to be speaking with a non-standard accent. Jean Castex is indeed originally mayor of a town in Gers, a South-West department of France, and speaks with a strong accent he is proud of. This rare situation shows that the reactions are still deeply intricated with ideologies towards the accent. For instance, his predecessor Edouard Philippe introduced him by saying *"il a un accent énorme,* mais *il est très compétent"* ["he has a huge accent, *but* he is very competent"][1] (emphasis is mine). Others said hearing him reminded them of holidays, or rugby, both being associated with the South-West. However, it also shows that there might be some change in the attitudes toward accents. In Castex case, having an accent might be taken as a seal of authenticity after the critics of elitism from the Macron administration — especially after the yellow vest movement.

The present study is interested in the evaluation French people make of "accents". It relies on folk definition of the accent as a way to speak that differs from the standard and/or differs from one's own speech. The accent can be regional or social, and it can be performative as in "accent de banlieue": the accent from suburbs, implicitly deprived estate, which regroup many stereotypes (Candea 2017).

I specifically focus on self-evaluations: the evaluation people make on their own accent or on the accent they are in contact with daily. That is to say, the accent is not necessarily their own, neither completely foreign. In several studies participants had to comment on their accent as well as on French accent from other regions, nationally and internationally, e.g. Toulouse speakers

---

[1] All translations from French quotes are mine.

commenting on regional French from Paris, Alsace and Provence as well as Quebec and African French (Pustka et al. 2019). However, no study focuses specifically on self-evaluation.

Studies have shown that people rate their own accent higher. In France "Parisian French" is considered the standard, and benefit from status prestige. Some regional accents benefit from solidarity prestige, e.g. Southern accent, while others have no prestige at all, e.g. Northern or Eastern accent. Moreover, there is a levelling in all Northern part of France.

I look at whether out-group evaluations are replicated in in-group evaluations. My study consisted in analysing a survey distributed online to native French speakers. Figure 1[2] shows a map of France where the points indicate where the participants of my study evaluated an accent. Each coloured area corresponds to an administrative region of France.

---

[2] All plots in this report are created with the R package ggplot (Wickham 2016). The plots that display points use the function jitter() to make all points visible: there is an error margin of 0.1 square: points might be off 0.1.
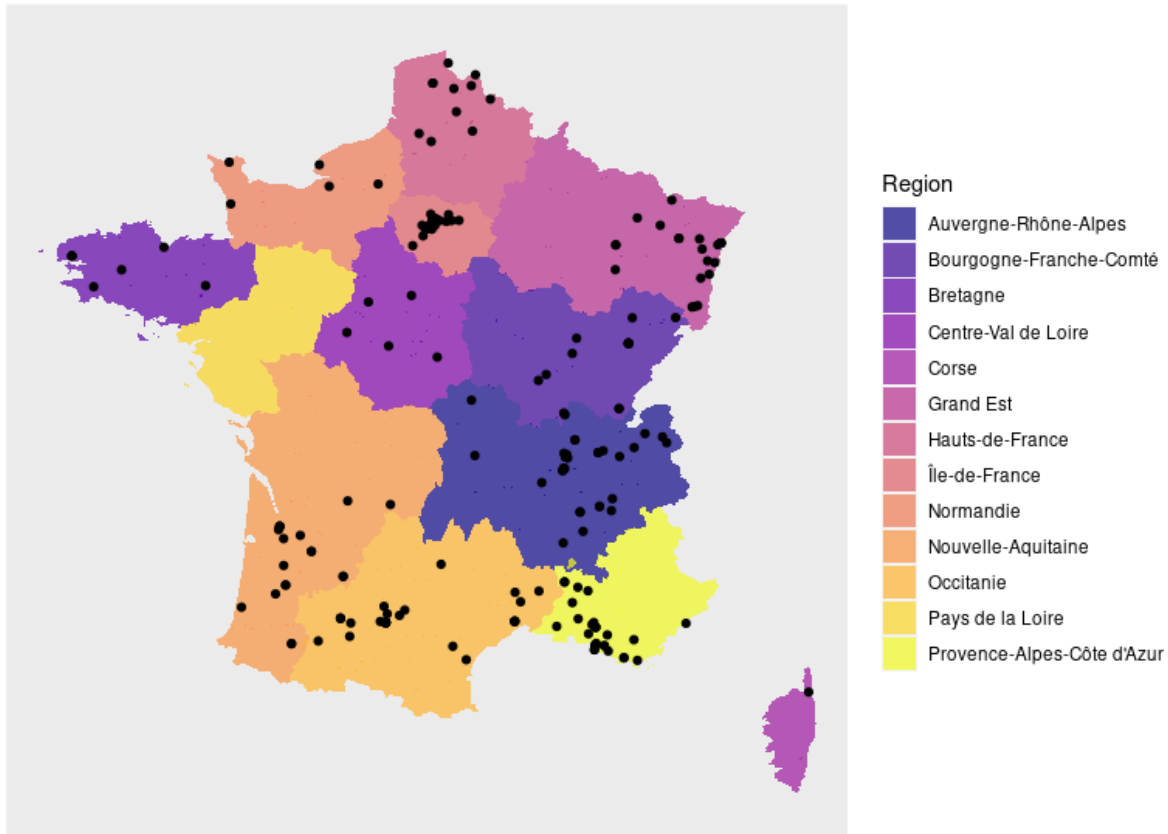
Figure 1: Map of the participants residences

In this report, I first review the existing literature on accents and language. I focus on research on the perception of accents, as well as the attitude towards accent, with a special interest in studies on French accents. I then present the methodology that I followed in order to collect the data I analyse. I describe the choices I made to build the online survey. I also explain how the survey was distributed, and who are my participants. In order to conduct the statistical analysis, an initial data processing was needed. This pre-processing is presented in the section Initial Data Processing. Finally, I show the Main Statistical Results in a section followed by the Discussion in which I comment those results in addition to other findings from the survey.

My research questions are as follow:

- How do people evaluate accents they really know?

- Are the categories based on location still relevant?

- Is social class (wages, education, occupation) a factor of evaluation?

- Is personal experience a factor? (own accent, region evaluation, personality)

Literature Review

In this section I present a literature review related to my dissertation topic. I first define the object of my work: accents, and review relevant studies on accent production. I then move on to accent perception and attitude studies. Finally, I review the specific case of France, and attitude to French accents.

Note that I focus here on the topic related literature. Further methodological papers are cited later to clarify some choices I made in building my survey.

Accent

In the present dissertation, I discuss accent as defined by Lippi-Green (1997, p. 142) as a "loose bundles of prosodic and segmented features distributed over geographic and/or social space" (cited in Candea et al., 2019). From a linguist point of view there is no neutral accent or non-accent (Lippi-Green 2012). However, from a folk perspective, an accent corresponds to a pronunciation that differs from the standard. It also includes pronunciation that differs from one's own pronunciation (Boyer 2015). In France, *dialecte* refers to local languages (Hall 2018). Beyond variations of French, Breton (from Brittany) and Basque (from the South-West) are also considered as *dialectes* even though they are not related to French. In the present study, I refer to accent as the "contemporary regional or social varieties of French" (Blanchet and Armstrong 2006).

Studies of accent have traditionally been focused on NORMs (Nonmobile, Older, Rural, Males) in dialectology (Chambers and Trudgill 1998). This branch of linguistics was attached to describing the accents or dialects of ideal speakers representing territorial areas, e.g. the Linguistic Atlas of Gilliéron and Edmont (1902-1910) (cited in Chambers and Trudgill, 1998). Sociolinguistics as developed by Labov (1966), Trudgill (1974) and Weinreich et al. (1968),

demonstrated the importance of variation and its link with social evaluation. I am interested in the reason for this variation. I focus on the attitudes toward the accent: how the accents are evaluated, as a potential explanation for having an accent.

Many personal characteristics are found to have an influence of the way people speak, e.g. social class: lower social class use more vernacular features (Labov 1966), gender: women use more standard forms (Trudgill 1983; Wolfram 1969), age: adolescent lead change (Tagliamonte and D'Arcy 2007), personal experiences and life stories (Sankoff 2004), attachment to region: L2 speakers have less accent when more attached to the country and the culture in addition to other factors such as duration of stay (Drummond 2013; Moyer 2007), or type of study (Baran 2014). I use many of those characteristics as potential factors for the evaluation of the accent. In return, regional accents can also be used as a way to negotiate one's identity and can be negotiated over and over again.

Accent identity and accent evaluation are closely linked through indirect indexicality (Ochs 1992; Silverstein 2003): for instance, Southern French accents are often associated with the sun and holidays, and having an accent might in turn be interpreted with the indexical meaning of friendliness, or laziness, through the process of iconisation (Eckert 2012; Irvine and Gal 2000; Peirce 1932). It is associated with "fractal recursivity" (Irvine and Gal 2000) that creates hierarchies between the standard and local or social varieties, as well as within those varieties. In my survey I look at what indexical meaning is associated with the regional accents, and whether this is linked other factors to create hierarchies between accents.

The ideologies (Woolard 1998) that maintain those hierarchies have been described in term of prestige. The standard recognized by the norm benefits from overt prestige, in comparison to vernaculars and local varieties which may benefit from covert prestige (Trudgill 1974, 1983).

Accents and languages can be evaluated on different scales depending on how the speakers sound (Giles 1970): status scale corresponds to indexical meaning of sounding smart, hard-working or wealthy, while solidarity scale corresponds to sounding honest, generous, friendly or humble. Prestige might also be locally determined, defined within community of practices (Eckert 1989).

Perception and attitude

Whatever the accent production, not only the perception of accents but also the attitude to them vary with respect to the ideologies and indexical meaning behind the accents. Matched guise studies are a way to elicit participants attitudes towards speech varieties by asking to rate voice samples of bilingual speakers on multiple scale, e.g. smart, nice (Lambert et al. 1960). I use similar scales in my survey. The main approach of those studies is to fix all parameters except the accent, such that the participants are only judging the accent as all other parameters are unchanged. They allow to get a better impression of participants attitude to real voice samples. However, those studies have been criticized given that the voice sample were not authentic (Garrett et al. 1999). Other studies were based on perceptual mapping: participants are asked to draw areas where they recognize an accent (Niedzielski and Preston 2010). Finally, many studies offer lists with options to rank or evaluate characteristics related to accents (Coupland and Bishop 2007; Levon et al. 2020). Multiple factors have been found significant across studies: people favour their own accent, women evaluate more positively, and young people attribute less prestige to standard variety, e.g. Coupland and Bishop (2007). I built my survey following those studies as developed in the Methodology section.

French accent description

France is known to be a very centralized country and with a very strong prescriptivist tradition concerning French language (Cheshire 2019; Encrevé and Braudeau 2007; Secova et al. 2018).

This led to the very worrying situation of minority languages of France, many of which are close to disappearance (Blanchet and Armstrong 2006). Accents are also subject to levelling (Armstrong and Pooley 2010; Boughton 2006). Nevertheless, some variations persist (Candea 2017; Gadet 2007; Gasquet-Cyrus 2013; Hall 2013; Pustka 2009).

The PFC project (Durand et al. 2009) aims at giving a clear protocol for phonological studies in the French speaking world in order to construct a corpus of comparable sociolinguistic data. It gathers more than fifty surveys and four hundred speakers, in France and around French speaking regions of the world. For instance, Sobotta (2006) and Ostby (2016) are two PhD studies that were conducted in this project on Aveyron and Guadeloupe, and on Parisian elite.

Recent studies in the UK and in France have used online crowdsourced surveys that allow to access a larger sample and get statistically significant results despite skewed samples (Avanzi 2018; Leemann et al. 2018).

French accent perception

In France, Kuiper (1999, 2005) conducted a study on Parisian and Provençal (South-East) attitudes towards regional accents. He presents maps of areas that were previously delimited by participants as well as attitude toward the regional varieties. The results show that Parisian French has the highest status, while regional varieties benefit from covert prestige. South regions are found to be especially friendly and beautiful. Only Alsace and North region accents do not benefit from any prestige for Parisians and Provence participants. Similar results are obtained using matched guise techniques to study attitudes towards accented French for Alsace, Paris, Provence and Bretagne (Paltridge and Giles 1984). Alsace accent is poorly rated, even by in-group participants. My survey

aims at comparing the results of those initial perception studies on French accents to the evaluation of people living among speakers of the accented variety of French.

Boughton (2006) shows that although Kuiper Parisian have strong feelings toward self-described regional accents, in reality the accents might not be distinguishable in terms of regions but only in terms of social differences (rural/urban), at least in the Northern part of France, between Rennes and Nancy. Moreover, Avanzi and de Mareüil (2019) demonstrates that while locals are able to recognize the accent of their region more than half of the time, there are confusions on other regions accents. My project draws on these conclusions and allows participants to comment on the accent they actually know and recognize.

Pustka et al. (2019) discuss the changes regarding prestige and attitudes towards French regional and international accents. Their hypothesis is that stigma comes from Paris but that, in the regions, the hierarchies are not equivalent and are shifting. They show that not only Paris benefit from overt prestige: "peripheries" are highly rated in in-group evaluations. Marseille is the only exception, probably influenced by out-group representation that are very salient (Gasquet-Cyrus 2012; Pustka 2011). My study could also help document this change in prestige by focussing on in-group evaluation of the accent. To do so, participants evaluate only the accent they are in contact with daily, but do not necessarily speak.

In this dissertation, I focus on accents. Keeping in mind the sociolinguistic variations that can be found, I focus on the attitude towards regional accents. The indexical meaning they might have results in, or comes from, hierarchies of accents between standard and local or social varieties. Various studies have highlighted some hierarchies of accent perception, where Paris is placed at the top of the status scale while regional varieties benefit from covert prestige. Studies on French mostly focus on out-group attitudes.

Methodology

Following many other sociolinguistics studies, I conducted a survey to elicit the evaluation of regional accent in France (Coupland and Bishop 2007; Kuiper 2005; Levon et al. 2020; Oakes and Warren 2007; Pustka et al. 2019).

I first describe how the survey was built. I then explain how I distributed the survey in order to get a total of one hundred and forty-two participants. Finally, some more information is provided on the variables I used for the models presented in the section Initial Data Processing and Main Results.

Building the survey

I constructed an online survey in order to get data on French native speakers' attitudes toward French regional accents. The survey was built using the Gorilla software (www.gorilla.sc). It is composed of three main parts, which are details thereafter. The first part asks participants to describe accents of places where they have lived and to rate those accents on multiple scales: friendly, beautiful, correct, serious. It also asks about attitude to the place itself. A second optional part provides the possibility for participants to record themselves reading a newspaper-like article which was designed by phoneticians for the research project Phonologie du Français Contemporain (PFC: Phonology of Contemporary French). This enables us to get access to the real accent of participants. Finally, in the third part, demographics details are collected to account for the variability in the data, e.g. gender, age, highest degree. Screenshots of the survey are provided in Annex A, including the consent form at the start of the survey. Details on the building of the survey and on the variables chosen and provided thereafter.

The survey starts with a consent form that must be agreed by participant as a prerequisite before accessing the main parts. The consent form was inspired from a survey by Gabriel Thiberge and Heather Burnett, yet to be published. It explains who is conducting the survey, what are the goal of the project, and how the survey unfolds. It also informs the participants about the anonymity of their response.

The first step of the survey records "life trajectories" information on the participants. The goal was to get information on where the participants spent most of their lives, and whether they often moved between regions. It also permitted to centre the attention to the fact that participants had to have lived in the regions they commented on, they were not supposed to comment on regions they had not experienced on a daily life level for at least three years. For the second step, the participants had to choose among the thirteen regions of metropolitan France (as they were redrawn in 2016) and the overseas regions.

Twelve out of the thirteen metropolitan regions were commented. This classification was problematic because the administrative regions of France are not properly matching linguistics areas. However, it was chosen for practical reasons: the frontiers are clear and official contrary to linguistic areas which are not as strict, and they are well known to any potential participant. In order to recentre the attention of participants to the area where they lived, I asked the exact place of residence in the region, suggesting giving details on the type of area it was, e.g. countryside, working class neighbourhood. I then asked details on the accent they were commenting, such as the area corresponding to the accent and the accent itself, which also allows to get how precisely they defined the accent.

Figure 1, in Introduction, shows a map of where participants had lived in the region they commented on. The map is inferred from the life trajectories: the postcodes are mapped with

coordinates from the postal service, available online[3]. We notice that most participants commented on parts of regions where the accent is stronger: that is the case in Nouvelle-Aquitaine (the most South-Western region) where most replies come from the lower half of the region, and in the North-East Grand-Est, where most replies come from the Eastern part of the region. Note that the mapping accuracy is slightly disturbed by the few (less than ten) participants who commented multiple accents from the same region, the middle of the two locations is used to represent such replies on the map.

The next questions were related to the relation of the participant to the accent: whether they had themselves an accent, and the proportion of friends and family who talk with this accent. I asked about judgment towards the accent via five Likert scales (Dőrnyei 2003; Schleef 2014) ranking this way of speaking as e.g. *"incorrect-correct"* ["incorrect-correct"] with the following scale: *"très incorrect", "plutôt incorrect", "ni correct ni incorrect", "plutôt correct", "très correct"* ["very incorrect", "rather incorrect", "neither correct nor incorrect", "rather correct" and "very correct"], and similarly with *"ridicule-sérieux"* ["ridiculous-serious"], *"antipathique-sympathique"* ["unfriendly-friendly"], *"laide-belle"* ["ugly-beautiful"] and *"marquée socialement"* ["socially marked"]. The ratings were inspired from previous survey on the topics e.g. Kuiper (2005) and Pustka et al. (2019). Two additional numeric scales (1 to 8) asked whether the participant *"aime cette façon de parler"* ["like this way of speaking"] and whether they though French people in general liked this way of speaking. I then collected judgments toward the region commented on itself (Moyer 2007). It consisted in rating participants agreement towards six

---

[3] https://www.data.gouv.fr/fr/datasets/base-officielle-des-codes-postaux/

statements concerning the experience in the region, see Annex B for more details on the questions and the options given.

The third optional step consisted in getting participants to read a passage out loud. The passage is titled *"Le Premier Ministre ira-t-il à Beaulieu ?"* ["Will the Prime Minister go to Beaulieu?"] and is provided among the screenshots of the survey in Annex A. It was created for the PFC project (Durand et al. 2009) with the explicit intention to cover all sounds of French. It has a special focus on schwa and liaisons, providing multiple contexts in which they occur. It also ensures the accessibility to a large range of readers, avoiding complex vocabulary and phrasing. This optional part motivated the choice of the Gorilla software for this survey, as it was the only one found to offer a way to get audio recordings via a web interface.

Finally, the fourth step gathered general information about the participants. It started with personality type questions, such as hobbies or the importance to be locally integrated. It followed with general questions related to their accent, such as remarks they got or if they liked hearing accents: what Coupland and Bishop (2007) calls "sociolinguistic diversity". A few classical demographics questions, inspired from Levon et al. (2020), were asked, such as gender or other language fluently spoken. Multiple questions were asked to elicit social class. Participants had to provide their socio-professional category from INSEE (Institut National de la Statistique et des Études Économiques), which is a classic way to get socio-economic status in French surveys (Bouchet-Valat and Jayet 2019; Chauvel 2001; Payne 2013; Penissat 2012; Savage et al. 2013)—student and stay at home parent categories were added as they recover different realities that "long term unemployed" in term of social class. I used a list of twenty-four occupations that were collapsed to six categories according to the pyramid of occupation defined by Chauvel (2001). The socio-professional category was collected both for the participants and their parents. The

household income was also collected. The options were chosen looking at the minimum salary and the medium salary given by INSEE (Bouchet-Valat et al. 2016). For the analysis, the two highest categories are merged. Finally, the highest degree obtained (levels are inspired from INSEE categorisations) together with the number of other languages spoken can be a sign of social status (Bouchet-Valat 2014; Vallot 2019).

The second step could be retaken up to three times to leave the option to participants to comment on multiple regions they had lived in. Seven participants commented on three regions, forty-five commented two and ninety commented one. This resulted in a total of two hundred and one comments of accents. One was discarded as it seems that, by the end of the survey, the participant got bored and did not answer anything, even answering *"flemme"* ["can't be bothered"] to a question on extra languages spoken. Another apologised in the comments saying they clicked on one region by mistake and could not go back to correct. I removed those two instances from the final data frame.

Gorilla Experiment Builder (www.gorilla.sc) was chosen because it provided the option to get an audio recording (Anwyl-Irvine et al. 2019). This option made it unfortunately impossible to offer a survey available on smartphones and tablets: it was only available on computers (laptop or desktop). This was pointed out by would-be participants who could not take the survey since they did not have access to a computer. This was especially true for people from lower socio-economic status.

Distributing the survey

Before distributing the survey on a larger scale, I had asked eight friends and family members to take the survey. It gave me the opportunity to improve the survey by clarifying a few questions and rephrasing explanations that were misunderstood.

The survey was first distributed through my personal network: friends and family, who themselves shared it in their personal networks (the first eight are not included in the results). I used the tool RISC[4] (Relai d'Information sur les Sciences de la Cognition) from the CNRS (Centre National de la Recherche Scientifique). It offers to distribute experiences to volunteers and give access to participants from an older demographics.

I then shared it by tweeting the link of the survey. A few French linguists with many followers on twitter gave a very large audience to my tweet by retweeting it, among them Maria Candea (@MarCandea – 5839 followers), Mathieu Avanzi (@MathieuAvanzi – 14.1K), Mederic Gasquet-Cyrus (@MedericGC – 1250) and Laelia Véron (@Laelia_Ve – 64.7K). A lot of followers were probably put off by the fact that they could not access the survey on their phone. More than eight hundred and fifty people clicked on the link provided in the tweet, only four hundred and seventy people actually started the survey, most of those who did not complete the survey stopped even before the consent form. I got big jumps in the completion of the survey after the retweets of the previously cited linguists.

The final channel to distribute the survey was Facebook groups. Overall, I contacted a hundred and thirty-eight local groups, yet the success rate of recruiting was very low. I first tried to target groups of local food producers to access potentially more rural population. I did not get accepted

---

[4] https://www.risc.cnrs.fr/

to many of those group: given their focus, my survey was considered as spam. I then tried and use exchange, mutual aid, swapping and selling websites. Most accepted only locals but I could share my survey on some. And finally, I shared it on the "Wanted community" Facebook groups which are mutual aid groups, in which I got the most engagement (likes and comments) but did not get many more participants.

The combination of the distribution through my personal network and via academic twitter accounts resulted in a skewed sample towards younger, educated participants, from privileged social class as Figure 2, 3 and 4 show. Moreover, the survey had many free text questions that could discourage uneducated participants. Finally, as mentioned before, allowing to answer the survey on smartphones and tablets could have expanded the participants profiles.
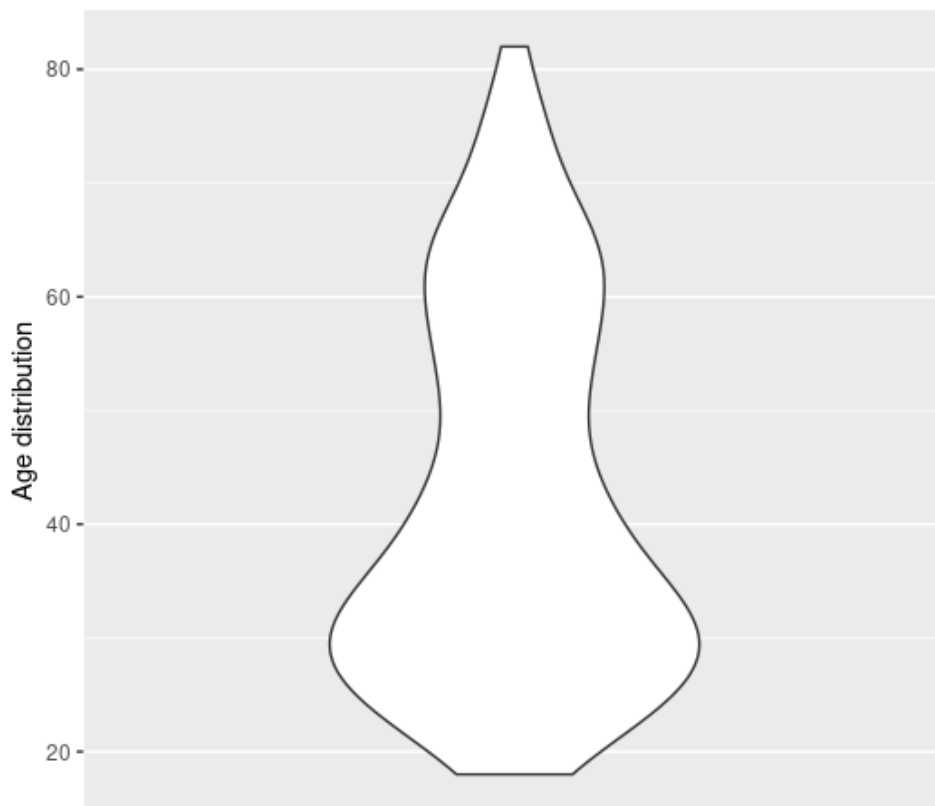
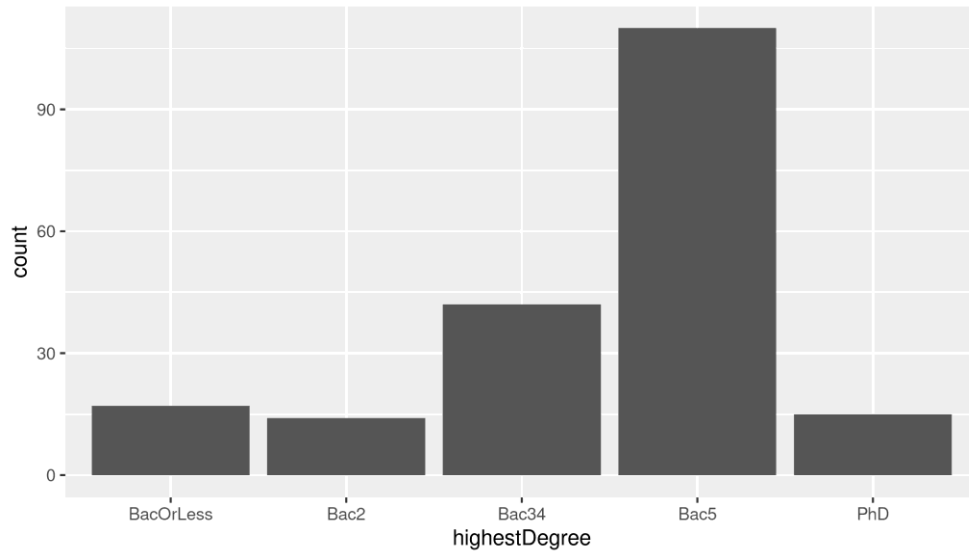

Figure 2: Distribution of participant age

Figure 3: Distribution of participant highest degree (Bac is for the French high school degree *Baccalauréat*)
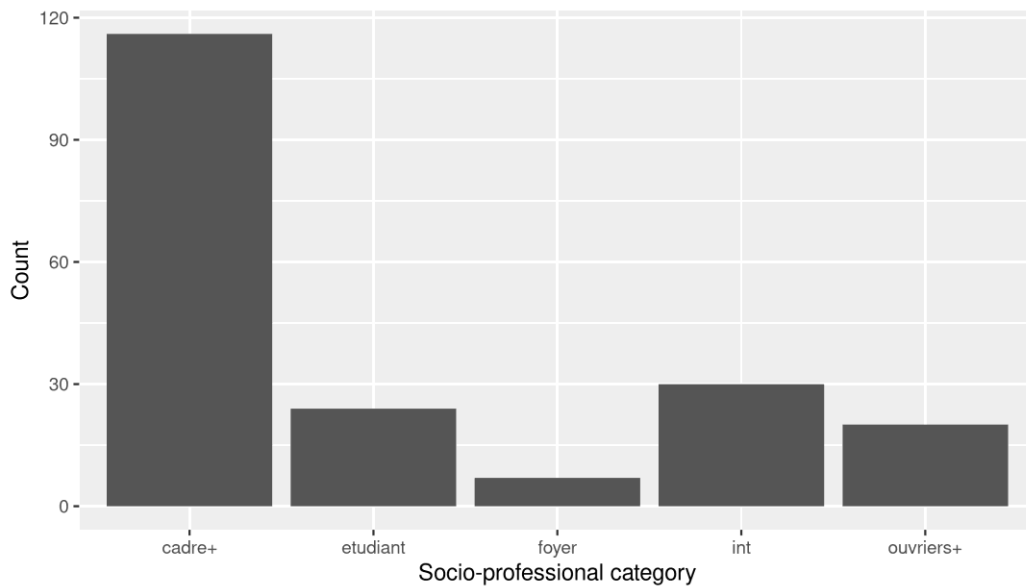

Figure 4: Distribution of participant occupation (*cadres* are higher social class, *etudiant* is for student, *foyer* is for stay-at-home parent, *int* is for middle class, and *ouvriers* is for lower social class)

Extracting variables from the survey response

The survey responses are used to build the dependent and independent variables necessary to the statistical analysis presented in section Initial Data Processing and section Main Results. Table 1 and Table 2 show the dependent and independent variables names, including the variables that are created after the Initial Data Processing. The column "Question" indicates the question from the survey screenshots provided in Annex A. Annex B provides details on the variables levels.

I detail here the variables that needed to be processed from the responses to the survey questions.

Variable name	Meaning	Question
correct	correct	B
serieux	serious	C
sympa	friendly	D
beau	beautiful	E
marque	socially-marked	F
jAime	I like the accent	G
emotionalJudgment	combination of sympa, beau and jAime	D, E, G
statusJudgment	combination of correct and serieux	B, C

Table 1: List of **dependent** variables built from the survey, for the statistical analysis

The dependent variables are built from the responses to questions of the survey eliciting the attitude to the accent.

The following independent variables were modified from the survey responses. The accent area describes the size of the geographical area where the accent is spoken, based on the description the participants made. The levels are as follow:

loc: very localised description of accent (a city, a valley)

dep: scale of a department size-wise (even if not corresponding to an official administrative department)

reg: scale of multiple department - might be a former region, or a list of departments

area: a broad area, such as "South-West", "all cities"

I added a variable (isCountryside) to indicate whether the participant lived in a rural or urban environment, based on location details. Another variable (isSameAsOrigin) indicated whether the region commented is the same as where the participant grew up. The origin region was deduced from the participants life trajectories given with postcodes and periods.

The types of degree obtained were classified as follow:

- linguistics and literature related studies (lit)
- other humanities (hum)
- economy (eco)
- politics and law (pol)
- care, doctors, nurses (ser)
- engineers, scientists (eng)

The comments about the remarks received by the participants were classified as positive, neutral or negative. Given that I specifically asked to comment on whether the remarks were positive or negative, the classification posed no issues, e.g. positive: *"je donne l'impression d'être en vacances"* ["I give the impression to be on holidays"], neutral: *"Ce n'est pas négatif ou positif"* ["neither positive or negative"], or negative: *"les gens se sont moqués de mon accent"* ["people made fun of my accent"], *"Généralement des remarques négatives, liées au caractère vieillot et campagnard de mon accent"* ["Generally negative remarks, tied to the old and countryside character of my accent"].

Finally, based on the audio recordings, I rated the accents on a scale from 1 to 4, as follow — 20% of the recordings were blind checked by another researcher, Celia Richy, only one recording was rated differently:

1. cannot hear any accent

2. very subtle, on a few words

3. audible accent

4. strong accent

Finally, some questions were not used to build variables for the statistical model. The variable relating to the judgment of the accent (positive/neutral/negative) in question A was dropped given that it was too uncertain. Some comments were clearly negative, e.g. *"disgracieux"* ["graceless"], *"aggressive"* ["aggressive"], *"pas sexy"* ["not sexy"], or positive, e.g. *"agréable"* ["pleasant"], *"amical"* ["friendly"], or neutral, e.g. description of vowel pronunciation: *"'et' prononcés 'é' (ex : dire 'chalé' plutôt que 'chalè' pour 'chalet')"* ["'et' pronounced 'é' (ex : say 'chalé' instead of 'chalè' for 'chalet')"]. However, others were difficult to judge, such as *"qui a des opinions"* ["has opinions"] or *"rigolo"* ["funny"], which without more context are difficult to estimate. Finally,

description like *"chantant"* ["singing"] can be seen as quite positive, but participants commented that they do not always appreciate the positive comments on the surface as *"c'est souvent associé à un fond de condescendance"* ["it is often associated with a background of condescension"].

Variable name	Meaning	Question
accentArea	Geographical area where the accent is spoken	6
accentStrength.audio	Rating of accent based on the audio recordings	16
accentStrength.region	People from the area have an accent	5
accentStrength.self	The participant think they have an accent	7
age	Age	27
csp	Socio-professional category from INSEE	25
csp.dad	See csp, of the father	29
csp.mum	See csp, of the mother	30
diplomeType	Type of degree obtained	24
frenchLikeAccent	"French people in general like this way of speaking"	9
frenchLikeRegion	French people in general like this region	15
friendsFamilyAccentDiversity	Proportion of friends and family who speak with the accent in this region	8
gender	Gender	22
hasAccentChangedInLife	"Do you think you have changed your accent during your life?"	18
highestDegree	Highest obtained degree	23
ID	Unique ID of a participant, provided by Gorilla	n/a

isCountryside	Based on the location details of where participants lived, indicates whether it was in a rural (TRUE) or urban (FALSE) environment	4
isSameAsOrigin	Region where each participant had grown up, indicates whether the origin region is the same as the one commented	3
likeHearingAccents	"I like hearing different accents"	21
localIntegrationImportance	It is important to be integrated in the local life	17
otherLanguageSpoken	Number of languages the participants listed as fluent	28
region	Administrative region of France	2
regionBelonging	Belonging to the region	10,11
regionEmotionalLink	Emotional link to the region	12,13
remarksFrequency	How often participants get remarks on the way they speak	19
remarksType	Detail on the remarks they get	20
wage	Income of participant household	26

Table 2: List of **independent** variables built from the survey, for the statistical analysis

Initial Data-Processing: Principal Component Analysis

In this section I present the three Principal Component Analysis (PCA) I conducted (Cavallaro and Chin 2009). The first PCA let us reduce the number of dependent variables relating to accent evaluation in order to have fewer variables to explore while still explaining most of the data. The second and third aimed to reduce the number of independent variables according to two groupings: evaluation of the region in which the accent is spoken, and social class.

Reducing the number of dependent variables

In the survey, seven variables measured evaluative judgments of accents. Six of them were rated on Likert scales: serious, correct, friendly, beautiful, socially marked and "I like this accent". A seventh variable consisted in a rating of the words chosen to describe the accent. This variable was dropped as it proved too uncertain too judge, see Methodology section.

I conducted a PCA in order to reduce the scales to two dependent variables. I used the function `prcomp()` from base R. Table 3 provides the component matrix and Figure 5 shows the resulting plot for the first two components. The cumulative proportion shows that three components explain 81% of the variation. PC1 and PC3 indicate that "socially marked" variable does not load with the rest of the variables and given that no significant factor was found to explain the variation in the data, this variable was dropped. The presence of the option "socially marked" might have had an influence on the correctness and seriousness ratings by suggesting that those could be separate scales. PC2 allows to discriminate two groups of variables that load together: correctness and seriousness on the one hand, and friendliness, beauty and "I like this accent" on the other hand. I thus created two dependent variables from the average (rescaled) of those ratings: status and

emotional judgments. We note that friendliness could be separated from beauty and liking to create a fourth dependent variable, but for now I keep it together with beauty and liking.

	PC1	PC2	PC3	PC4	PC5	PC6
jAime	0.508	-0.276	0.055	0.162	-0.264	0.753
sympa	0.38	-0.508	0.24	-0.637	0.295	-0.219
beau	0.503	-0.19	0.005	0.591	-0.139	-0.585
correct	0.411	0.414	-0.304	0.094	0.735	0.135
serieux	0.386	0.454	-0.357	-0.459	-0.532	-0.156
marque	0.166	0.502	0.848	0.007	-0.026	-0.001
Standard deviation	1.666	1.114	0.921	0.669	0.655	0.507
Proportion of Variance	0.463	0.207	0.141	0.075	0.072	0.043
Cumulative Proportion	0.463	0.67	0.811	0.886	0.957	1

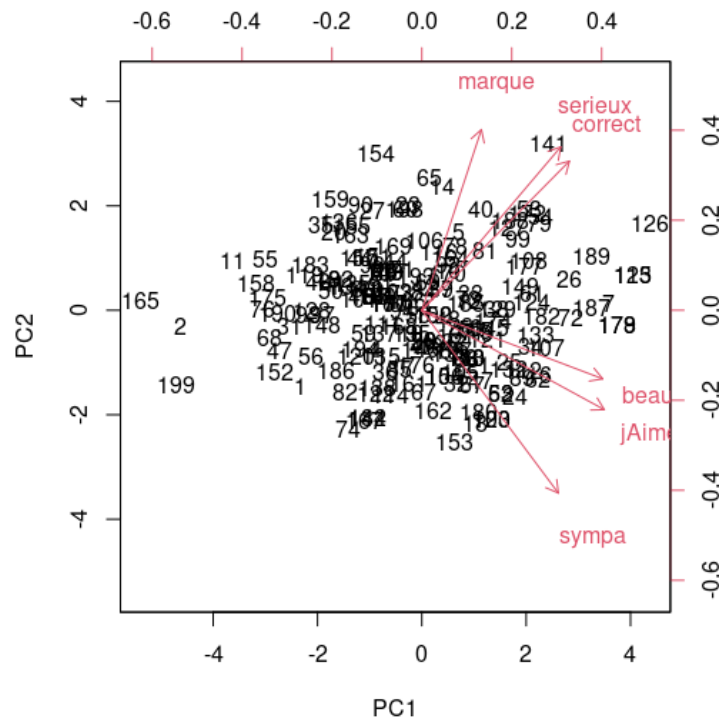Table 3: Component matrix of the PCA for the dependent variables

Figure 5: Biplot resulting from the PCA analysis to build the dependent variables: serieux (serious) and correct load together; beau (beautiful), sympa (friendly) and jAime (I like) load together

Reducing the number of independent variables

I also conducted a PCA on a set of independent variables relating to the participant judgment towards the region. The goal was to see if some loaded together and could be grouped. Table 4 and Figure 6 show the results for the different variables under consideration: "I like the region", "I like the people", "I have good souvenirs", "attachment", and "being a part of". Looking at PC2, the first three load together under the sense of emotional link, while the last two variables load together under the general sense of belonging. Two components are enough to explain 81% of the

data. I expect that the emotional link should be correlated with the emotional judgment of the accent.

	PC1	PC2	PC3	PC4	PC5
jaimeLaRegion	0.466	-0.318	0.51	0.053	-0.647
jaimeLesGensRegion	0.446	-0.385	-0.761	0.264	-0.067
bonSouvenirsRegion	0.469	-0.316	0.202	-0.516	0.611
attachesRegion	0.401	0.679	-0.251	-0.484	-0.283
fairePartieRegion	0.45	0.435	0.239	0.653	0.352
Standard deviation	1.819	0.872	0.596	0.558	0.514
Proportion of Variance	0.662	0.152	0.071	0.062	0.053
Cumulative Proportion	0.662	0.814	0.885	0.947	1

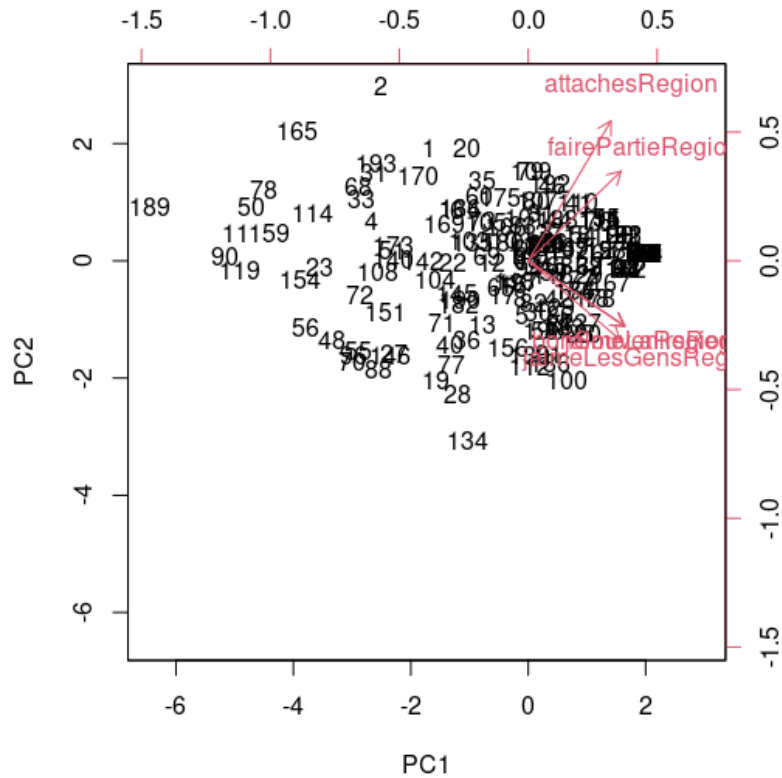Table 4: Component matrix of the PCA for the region variables

Figure 6: Biplot resulting from the PCA analysis to build the region judgment: attaches-Region (attachment to the region) and fairePartieRegion (to be part of the region) load together, jAimeLaRegion (I like the region), jAimeLesGensRegion (I like the people), bonSouvenirsRegion (good memories in the region) load together

Finally, I conducted a PCA on social class variables: wage, socio-professional category (csp) of participants and their parents and highest degree, see Table 5. Only the socio-professional category of the parents loaded together. I chose to only keep socio-professional category of the father in the model given that this is the traditional way to look at social class, with the father's occupation defining the family social class (Armstrong and Unsworth 1999). I kept all the other independent variables separately.

	PC1	PC2	PC3	PC4	PC5
demo_revenue.quantised	0.403	0.458	-0.771	0.164	0.08
csp	0.473	0.38	0.597	0.452	0.266
csp.dad	0.401	-0.54	-0.093	-0.248	0.69
csp.mum	0.329	-0.589	-0.107	0.565	-0.463
demo_diplome.quantised	0.588	0.078	0.171	-0.623	-0.481
Standard deviation	1.321	1.184	0.823	0.787	0.747
Proportion of Variance	0.349	0.28	0.135	0.124	0.112
Cumulative Proportion	0.349	0.629	0.765	0.889	1

Table 5: Component matrix of the PCA for the social class variables

Main Results: Mixed-effect models

In this section I present the analysis results of the survey to uncover which factors drive the evaluation of accents spoken in places well known to the participants. The factors under study correspond to the independent variables presented in Methodology section, see Table 2.

I detail how I built two separate models for emotional and status judgment, before presenting the results for each. Following common practice in sociolinguistics (Avanzi and de Mareüil 2019; Baayen 2013; Tagliamonte and Baayen 2012), I used mixed-effect models with participant as random effect. I conducted the analysis in R, with the function lmer() of package lme4 (Banta et al. 2010) and lmerTest (Bates et al. 2014; Kuznetsova et al. 2017). It allows to account for the fact that participants had the opportunity to comment multiple accents if they wished too, and that some participants might have rated very low or very high for uncontrolled factors. I built both models starting by including all independent variables. I first removed only those necessary for the model to converge, e.g. variables with too many missing values, and to avoid singular fits. Singular fits indicate that the covariance is close to zero which might by a sign that some variables are correlated with others. To check this, I looked at the correlation matrices that I obtained thanks to the function vcov(), which outputs the variance-covariance matrix from the mixed-effect model, and the function cov2cor(), which outputs the correlation matrix from the covariance matrix.

The handling of missing values by lmer() can lead to significant reductions in the data set: all entries missing one variable are ignored (see Annex C for a complete list of the missing values). Two third of the participants did not complete the audio task, which consisted in reading the PFC text presented in Methodology, and therefore do not have an audio-based accent strength measure. This variable was thus excluded from the model to avoid a significant reduction in usable responses

for the analysis. I rely instead on self-assessed accent strength for the present analysis and reserve audio-based accent strength data for a later study. Figure 7 shows that the two are correlated, which reassures us that the self-assessments are a reliable measure. Note that I focus on cases where the self-assessed accent strength corresponds to the region where the participants grew up: given that participants could comment multiple accents, they might have rated very low the accent strength of one region while they speak with the accent of their region of origin.
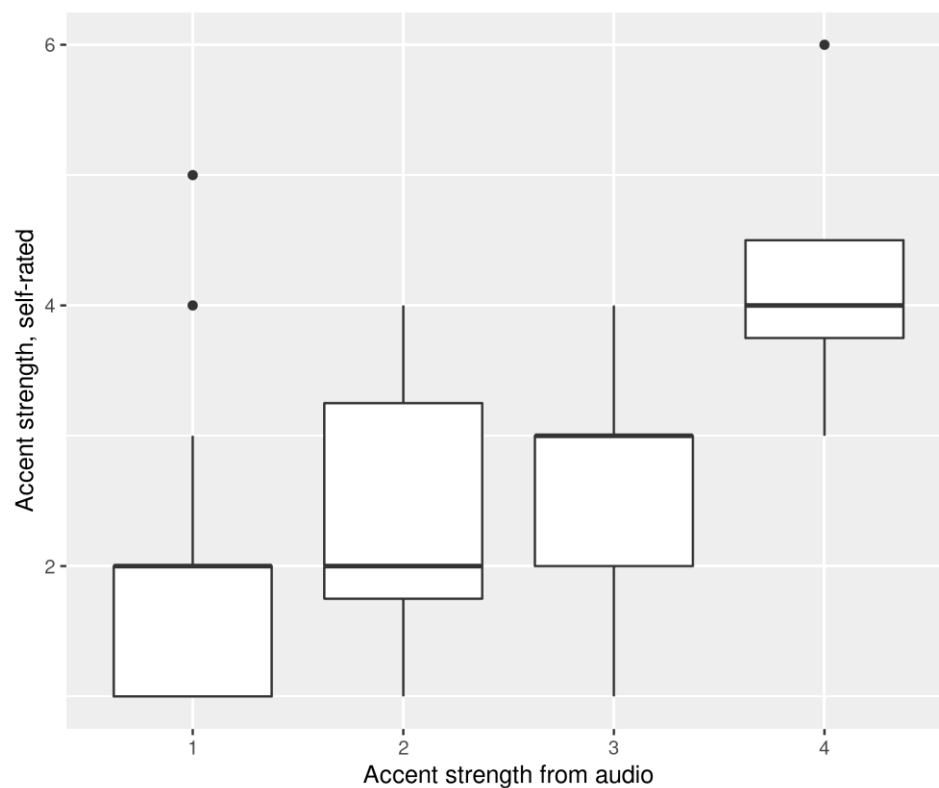

Figure 7: The accent strength of participants from audio is correlated with their self-rating of their accent

Results for Emotional Judgment of Accents

In this section I look at how the emotional evaluation of accents is driven by the independent variables presented in Methodology section, see Table 2. Before I present the results I clarify the exclusion of the variables "emotional link to the region" and "type of remarks received by the participants.

The high correlation of emotional link to the region with emotional judgment evaluation ($p = 3.1e-6$) is cautionary. I argue that being emotionally attached to a geographical area includes being emotionally attached to the accent spoken there. Figure 8 shows that only one participant rated very low for the emotional link to the region whereas high on the emotional evaluation of the accent. Looking more closely at their responses, it is a bit odd that this person is the only one who rated 1 for all the region questions: it could be the case that they inverted the scales. They did not write any comment, so it is difficult to understand their position. In any case, this person is an exception: while it seems possible to be attached to the region but not like the accent, liking the accent involves liking the region. In other words, if someone dislike a place, they will devalue the accent spoken there. Therefore, the two variables cannot be treated as independent because of the risk of endogeneity (Sharma 2017).
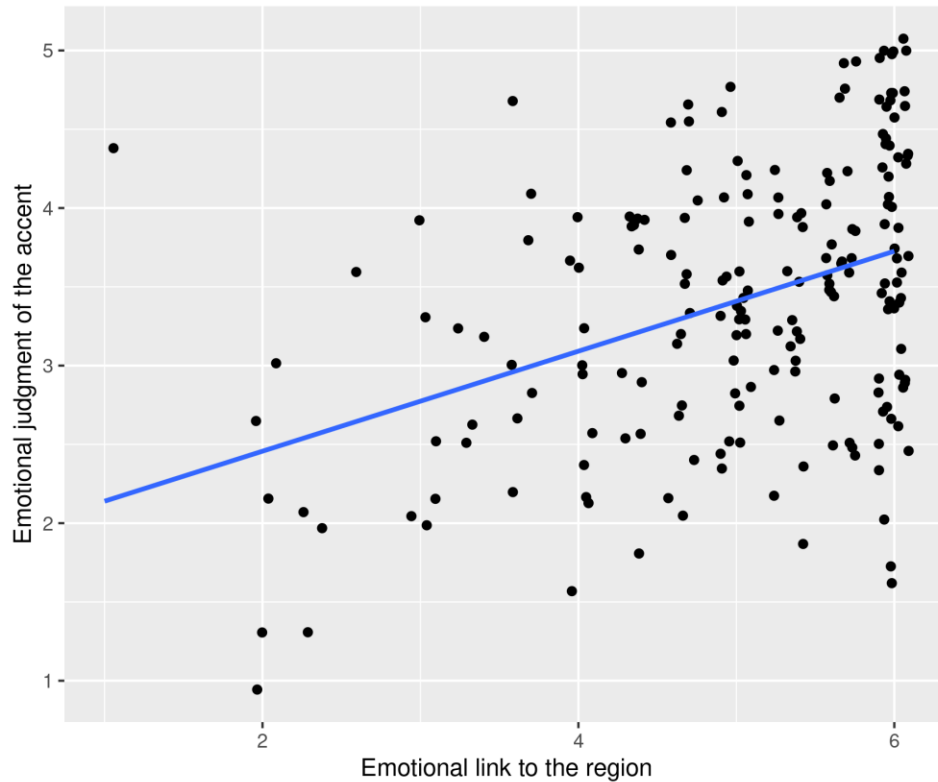
Figure 8: Emotional Judgment is significantly correlated with emotional link to the region

Given the handling of missing values by lmer(), keeping some variables in the model results in dropping many participants from the results. For this reason, I also removed wage and type of remarks from the independent variables since they are not significant in the global model and have many missing values. Other variables missing values were either significant, e.g. type of degree, or removing them did not change the main effects of the model, e.g. socio-professional category or gender. I kept the model maximal.

Figure 9 presents the ANOVA results of the resulting model. It gives the significance of independent variables, instead of the classic summary which gives it by levels. Given the number of levels for many variables, the summary results constitute a very large table. It also does not

always present relevant results since it shows the results in comparison to one reference level

instead of all combination possible, which would be an even bigger table.

```
> m.E.noRT.noWage <- lmer(emotionalJudgment~ region+isSameAsOrigin+isCountr
yside+accentStrength.region+accentArea+accentStrength.self+friendsFamilyAcc
entDiversity+frenchLikeAccent+regionBelonging+frenchLikeRegion+localIntegra
tionImportance+hasAccentChangedInLife+remarksFrequency+likeHearingAccents+g
ender+highestDegree+diplomeType+csp+age+otherLanguageSpoken+csp.dad+(1|ID),
 data=df.m.reg)
> anova(m.E.noRT.noWage)
Type III Analysis of Variance Table with Satterthwaite's method
                              Sum Sq Mean Sq NumDF  DenDF F value  Pr(>F)
region                        7.8079 0.86754     9 82.138  2.6135 0.01047 *
isSameAsOrigin                0.4826 0.48259     1 70.596  1.4538 0.23194
isCountryside                 0.0153 0.01530     1 85.564  0.0461 0.83053
accentStrength.region         1.1332 0.22664     5 78.801  0.6828 0.63784
accentArea                    0.1477 0.04922     3 81.293  0.1483 0.93051
accentStrength.self           1.4932 0.37330     4 84.950  1.1246 0.35041
friendsFamilyAccentDiversity  0.5022 0.12555     4 80.298  0.3782 0.82357
frenchLikeAccent              5.9925 0.85607     7 82.113  2.5789 0.01872 *
regionBelonging               2.0252 2.02524     1 72.864  6.1010 0.01585 *
frenchLikeRegion              1.4741 0.36853     4 79.893  1.1102 0.35760
localIntegrationImportance    1.2742 0.31854     4 61.160  0.9596 0.43619
hasAccentChangedInLife        0.4126 0.41257     1 65.708  1.2429 0.26898
remarksFrequency              1.9366 0.64554     3 57.246  1.9447 0.13259
likeHearingAccents            0.2857 0.09523     3 58.017  0.2869 0.83468
gender                        0.4713 0.47127     1 54.007  1.4197 0.23866
highestDegree                 2.7860 0.69650     4 61.657  2.0982 0.09180 .
diplomeType                   4.8387 0.96775     5 54.334  2.9153 0.02107 *
csp                           3.7151 0.92878     4 57.720  2.7979 0.03417 *
age                           0.0346 0.03463     1 54.046  0.1043 0.74794
otherLanguageSpoken           2.9294 0.97646     3 59.523  2.9416 0.04026 *
csp.dad                       1.9611 0.65370     3 52.367  1.9693 0.12993
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> |
```

Figure 9: Emotional judgment, the following variables are significant: region, frenchLikeAccent, regionBelonging, highestDegree,diplomeType, csp and otherLanguageSpoken

The ANOVA results Figure 9 show that for French native speakers, the emotional judgment of

accent is driven by the following significant factors: the region where the accent can be heard

(p=0.010), whether the participant thinks French people like the accent (p=0.019), the feeling of

belonging to the region where the accent is spoken (p=0.016), the type of degree (p=0.021), the

socio-professional category (p=0.034), and the number of other languages spoken (p=0.040). The

highest degree obtained is close to significant (p=0.092). Contrary to what previous studies had

found (Coupland and Bishop 2007; Paltridge and Giles 1984), in my results, the evaluation is not

significantly driven by age, gender, sociolinguistic diversity (like hearing accents) and whether the accent is their own.

I discuss those results in more details in the Discussion section.

Finally, I added an interaction term between region and "French like accent", however adding the interaction to the model had the consequence of making the model rank deficient. Even when collapsing levels of "French like accent", given the many regions under study, seven columns had to be dropped: many cells were empty. An ANOVA to compare the two models with and without interaction shows that they are significantly different and that with interaction is only slightly a better fit (lower AIC: 371.64 vs 371.12). Given that the interaction is not significant I chose to keep the interaction out. Note that the ANOVA comparison between models was not possible when removing previous variables, because since some values were missing, the dataset was reduced and could not be compared with the larger dataset.

Results for Status Judgment of Accents

I followed the same process to build the model of the status judgment. Similarly, the accent strength variable collected from the audio task had to be dropped for the model to converge. I also excluded socio-professional category of the dad and wage to avoid singular fit, since those two variables were not significant.

Given that remark frequency is a significant factor, I added the interaction with remarks type. An ANOVA to compare both models with and without interaction shows that the model is better with the interaction (AIC = 253.28 < 261.01). The results are presented in Figure 10. It shows that for French native speakers, the status judgment of accent is driven by the following significant factors: the frequency of remarks the participant received concerning their accent (p=0.003), the age

(p=0.020), whether the participant thinks French people like the accent (p=0.029), the region where

the accent can be heard (p=0.037). The three following factors are close to significant: whether the

participants lives in the countryside (p=0.053), the highest degree obtained (p=0.082) and the type

of degree (p=0.098). Contrary to previous studies, gender, sociolinguistic diversity (Coupland and

Bishop 2007), and the emotional link and belonging to the region (Moyer 2007) are not significant.

In the next section, I present the results in more detail with a full discussion.

```
> m.S.int.noCSPdad.noWage <- lmer(statusJudgment~ region+isSameAsOrigin+isCount
ryside+accentStrength.region+accentArea+accentStrength.self+friendsFamilyAccent
Diversity+frenchLikeAccent+regionEmotionalLink+ regionBelonging+frenchLikeRegio
n+localIntegrationImportance+hasAccentChangedInLife+remarksFrequency*remarksTyp
e+likeHearingAccents+gender+highestDegree+diplomeType+csp+age+otherLanguageSpok
en+(1|ID), data=df.model)
> anova(m.S.int.noCSPdad.noWage)
Type III Analysis of Variance Table with Satterthwaite's method
                                Sum Sq Mean Sq NumDF  DenDF F value   Pr(>F)
region                         6.5933 0.73259     9 43.730  2.2402 0.037078 *
isSameAsOrigin                 0.6546 0.65464     1 43.317  2.0018 0.164259
isCountryside                  1.2995 1.29952     1 42.681  3.9738 0.052631 .
accentStrength.region          1.0762 0.21523     5 43.229  0.6582 0.656987
accentArea                     0.6694 0.22314     3 44.553  0.6823 0.567538
accentStrength.self            2.5473 0.63682     4 44.405  1.9473 0.119232
friendsFamilyAccentDiversity   0.3394 0.08484     4 42.744  0.2594 0.902280
frenchLikeAccent               5.7583 0.82261     7 44.039  2.5155 0.028862 *
regionEmotionalLink            0.0963 0.09632     1 44.600  0.2945 0.590028
regionBelonging                0.1793 0.17931     1 44.612  0.5483 0.462893
frenchLikeRegion               1.6363 0.40906     4 44.650  1.2509 0.303405
localIntegrationImportance     1.5482 0.38705     4 33.651  1.1835 0.335757
hasAccentChangedInLife         0.0166 0.01659     1 34.574  0.0507 0.823118
remarksFrequency               4.6787 2.33935     2 29.243  7.1535 0.002958 **
remarksType                    0.3087 0.15436     2 33.855  0.4720 0.627781
likeHearingAccents             0.8764 0.29213     3 33.217  0.8933 0.454811
gender                         0.0308 0.03084     1 37.171  0.0943 0.760493
highestDegree                  2.9698 0.74244     4 33.888  2.2703 0.082004 .
diplomeType                    3.3672 0.67344     5 30.505  2.0593 0.098063 .
csp                            2.1419 0.53546     4 27.505  1.6374 0.193181
age                            1.9903 1.99033     1 27.180  6.0862 0.020210 *
otherLanguageSpoken            0.1416 0.04719     3 30.472  0.1443 0.932542
remarksFrequency:remarksType   2.1412 0.53529     4 31.467  1.6369 0.189560
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 10: Status judgment: remarksFrequency, region, frenchLikeAccent, age, isCoun-

tryside, highestDegree, diplomeType are significant

Note that the missing values for the type of remarks are not randomly distributed. Keeping the variable in the model excludes all participants who never received remarks about their way of speaking. When removing remarks type from the model, only a subset of the previously mentioned remain significant: the frequency of remarks, the socio-professional category and whether French people like the accent.

Discussion

In this section I discuss the variables that were found to be significant from the ANOVA of the mixed-effects models covered in the previous section. I first focus on the relation of language and place. I then look at how participants use their speech to make a stance about their regional and social belonging.

Changing regional accent dynamics

The results show that region is significant in participants evaluation of the local accent.

Not only is accent part of what makes the region, it also indexes belonging to the place (Dubois and Horvath 1998; Irvine and Gal 2000). Participants mention it as part of their identity: *"c'est l'accent de mon enfance"* ["it is the accent of my childhood"], *"c'est mon accent"* ["it is my accent"]. Some regret "losing it": *"quand j'entends des gens me dire que je n'ai pas d'accent alors que je viens du Tarn, ça me rend un peu triste"* ["when I hear people telling me that I don't have an accent although I come from Tarn, it makes me a bit sad"]. Other participants are proud to have adopted the accent of a region where they move to, as a way to show their authenticity: "Ils disent que je deviens vraiment marseillaise" ["They say I'm really becoming Marseillaise/from Marseille"].

Figure 11 shows the evaluations towards the regions, detailed in Figure 12. As mentioned earlier, using administrative regions as a subdivision of France may not be the most relevant geographical splitting with respect to linguistics, yet it is an established and well-known categorisation which allows to observe patterns and trends in our survey results. The evaluations are overall positive: this confirms the observation that in attitude studies, participants rate their own region higher, e.g. Coupland and Bishop (2007). Considering that whether the participants had grown up in the region

or not was not a significant factor, "own region" could be seen not only as where participants grew up, but also wherever they have lived for at least three years in their life.
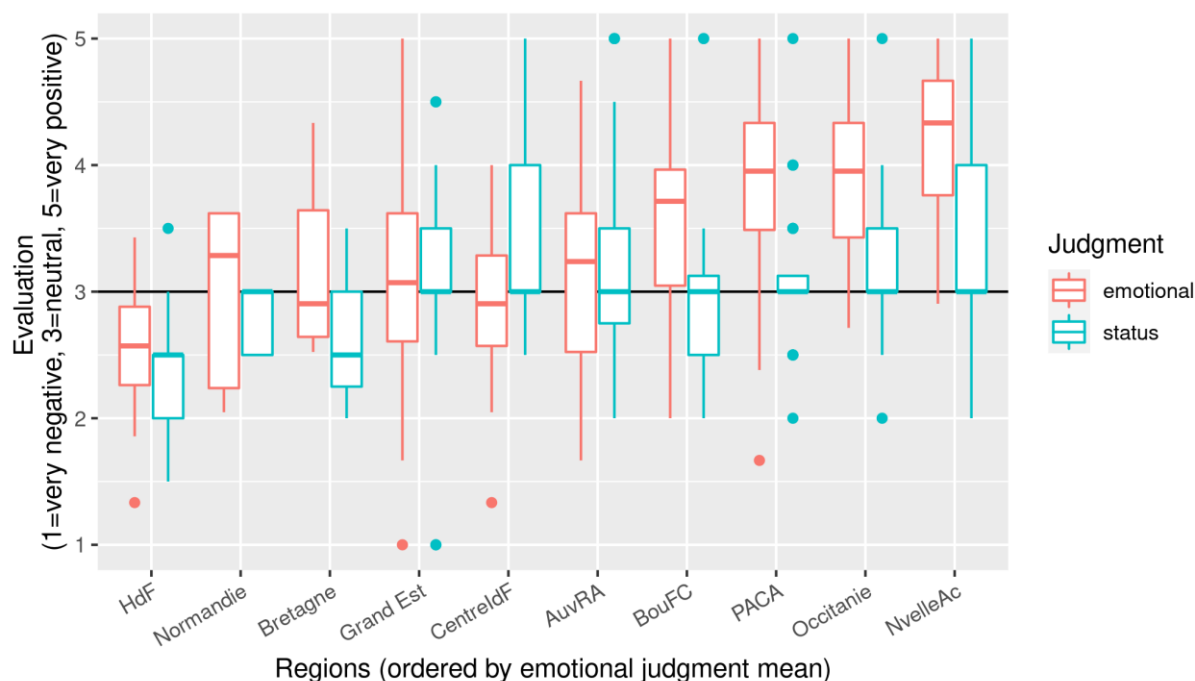

Figure 11: Distribution of emotional and status judgement in the regions

The ranking of regions is similar to what has been found in previous studies, mostly based on outsiders evaluations (Kuiper 1999, 2005; Paltridge and Giles 1984): South regions are highest in emotional evaluations, Ile-de-France and Centre are highest in status, while North (HdF) is lowest for both. We also observe an increase in correctness from East to West in the South, from PACA, to Occitanie, to Nouvelle-Aquitaine. The accent of Marseille (main city of PACA) is associated with vulgarity and lower social class, higher social class having a softer accent (Gasquet-Cyrus 2013). On the other hand, status in Nouvelle-Aquitaine is even surprisingly high, up to the level of Ile-de-France and Centre. The South-West accent is considered by some as more correct than

standard French given that it is argued to be closer to the written norm (Durand et al. 1987; Pustka et al. 2019). Bordeaux (main city of Nouvelle-Aquitaine) is closer to Paris and is associated with products of the higher-class culture: wine, oysters and bourgeoisie.

Looking in more details at the ungrouped variables in Figure 12, it is interesting to note that globally the accents are rated a bit higher on friendliness than on beauty. The only exception is Ile-de-France (Paris) which is considered neutral on the beauty scale but not friendly. The accent is described by some as *"prétentieux"* ["pretentious"], *"condescendant"* ["condescending"], or *"pas mélodieux"* ["not melodious"]. Previous studies have found that Parisian French still benefits from emotional prestige, in comparison to other regions: first ranked is own region, Paris comes in the top positions and other poorly rated regions follow (Kuiper 1999, 2005; Paltridge and Giles 1984). Here, on an absolute scale, Parisian French is the lowest rated emotionally, comparing to everyone self-evaluations. CentreIdf represents the standard accent (Lippi-Green 1997) that holds the overt prestige (Trudgill 1974), e.g. *"au travail"* ["at work"], *"en contextes sérieux"* ["in serious contexts"]. All other accents are considered friendly and hold local covert prestige (Eckert 1989).
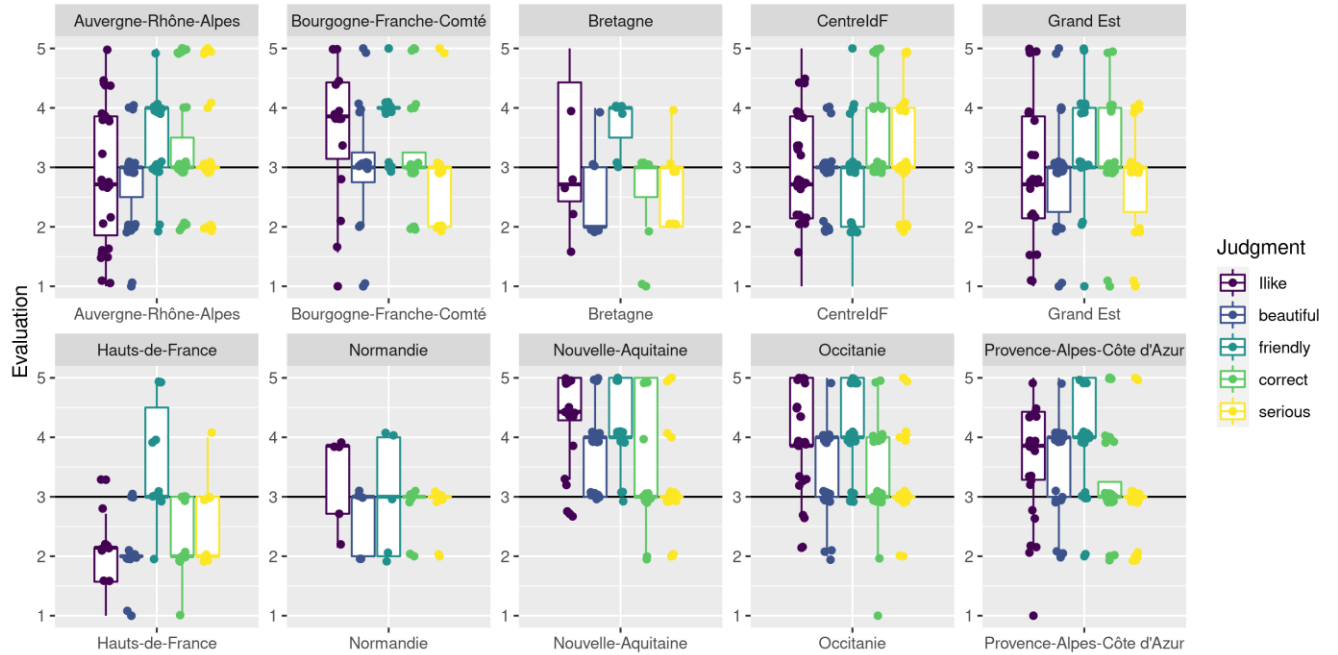
Figure 12: All evaluation variables for each region

Surprisingly evaluations of Grand-Est accent are not as low as expected from previous attitude studies. The status of Eastern accents may have changed over the years, as Figure 13 shows according to the apparent time hypothesis (Eckert 1997). For instance, Strasbourg (main city in Alsace) is now a very attractive city compared to the North region which still suffers from bad stereotypes similar to those of mining regions in the UK. Moreover, Alsace low ratings on a national level are linked to the proximity with Germany and the prejudices dating back to the several wars between Germany and France. On the contrary, people from Alsace, or people who have lived there long enough are potentially not affected as much by such negative stereotype, while still describing the accent as *"germanique"* ["Germanic"].

``
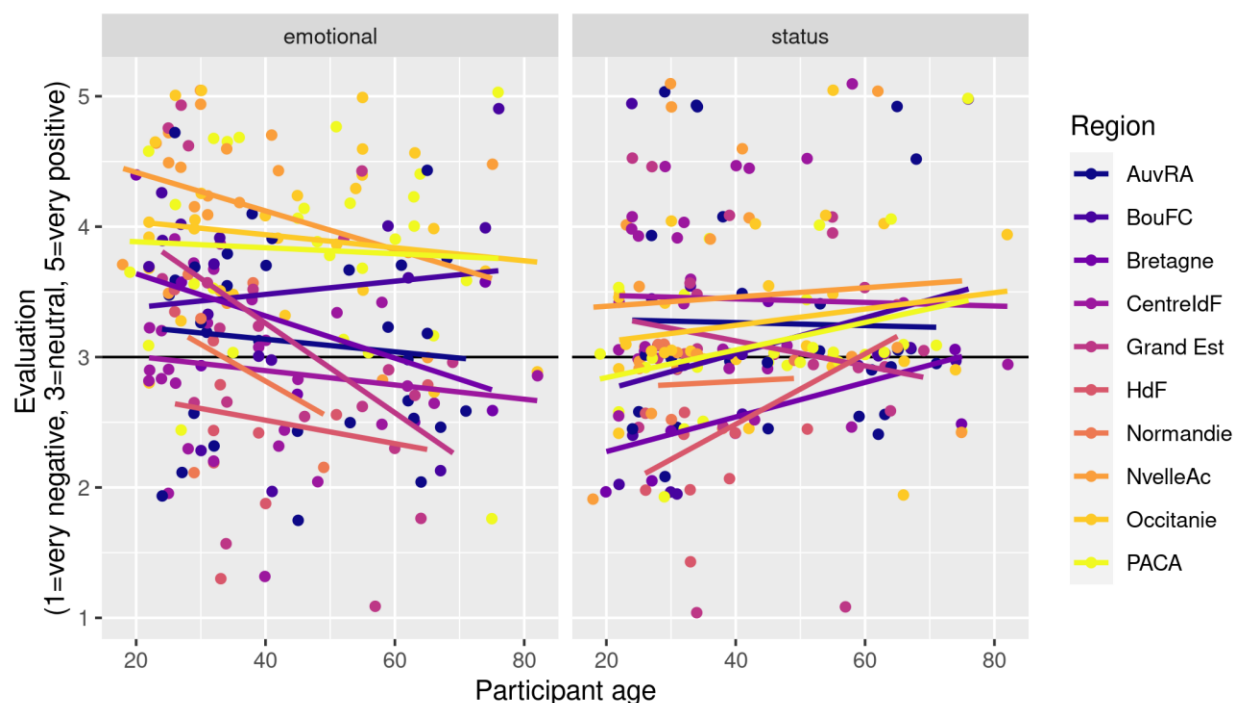
Figure 13: Evolution of judgment depending on participant age and region commented: evaluation towards Grand-Est accent increases with time, with younger participants giving higher ratings than older ones.

The levelling in northern French accents results in a social class divide in place of the regional divide (Armstrong and Pooley 2010; Boughton 2006; Gadet 2007). This is illustrated in Figure 14: the accent in Normandy, Brittany and North region are rated as especially socially marked. Many participants who rated an accent as very socially marked mention that they needed to hide their accent in some contexts: e.g. *"je le 'cache' lorsque je quitte ma région, sinon on vous catalogue de beauf"* ["I 'hide' it when I leave my region, otherwise you're categorized as hick"], *"J'ai un proche (qui essaye de le gommer) qui prétend que c'est un accent de 'pequenaud'"* ["I have a friend (who tries to erase it) who pretends it's a 'hick' accent"], *"C'est donc un accent 'de pauvre' mais aussi de 'campagnard' maintefois stigmatisé dans les médias ou les espaces culturels"* ["it is a 'poor people' accent, and also 'countryside', many times stigmatised in the media or cultural spaces"], *"I do everything to correct my ch'ti accent"* ["je fais tout pour corriger mon accent

ch'ti"]. They therefore use different varieties of French in order to make a stance and disalign from the ideology (they think is) associated with the accent (Du Bois 2007). I also note that many are perfectly aware of the stereotypes associated with the accent, some regret playing along the rules they do not approve of.
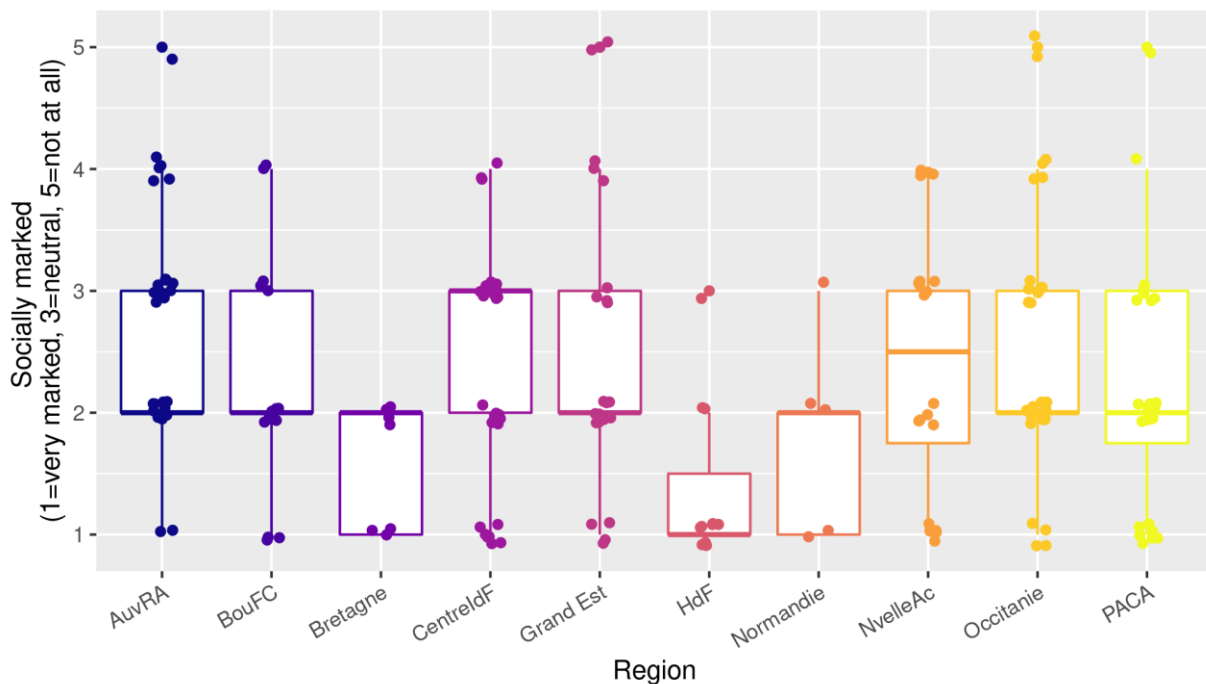

Figure 14: Accent in Normandy, Brittany and North regions are rated as socially marked

Figure 15 shows that there is a clear correlation between participants' own evaluations of regional accent and how participants see the national ideology toward these accents. It is not surprising that the national ideology concerning accent, whether French people like the accent, is significant. Indeed, participants are not independent from the society they live in, so we can expect that the participants evaluation reflects the national ideology. The question remains whether the

participants are influenced by an ideology previously in place or whether they claim the ideology reflects their views somehow.
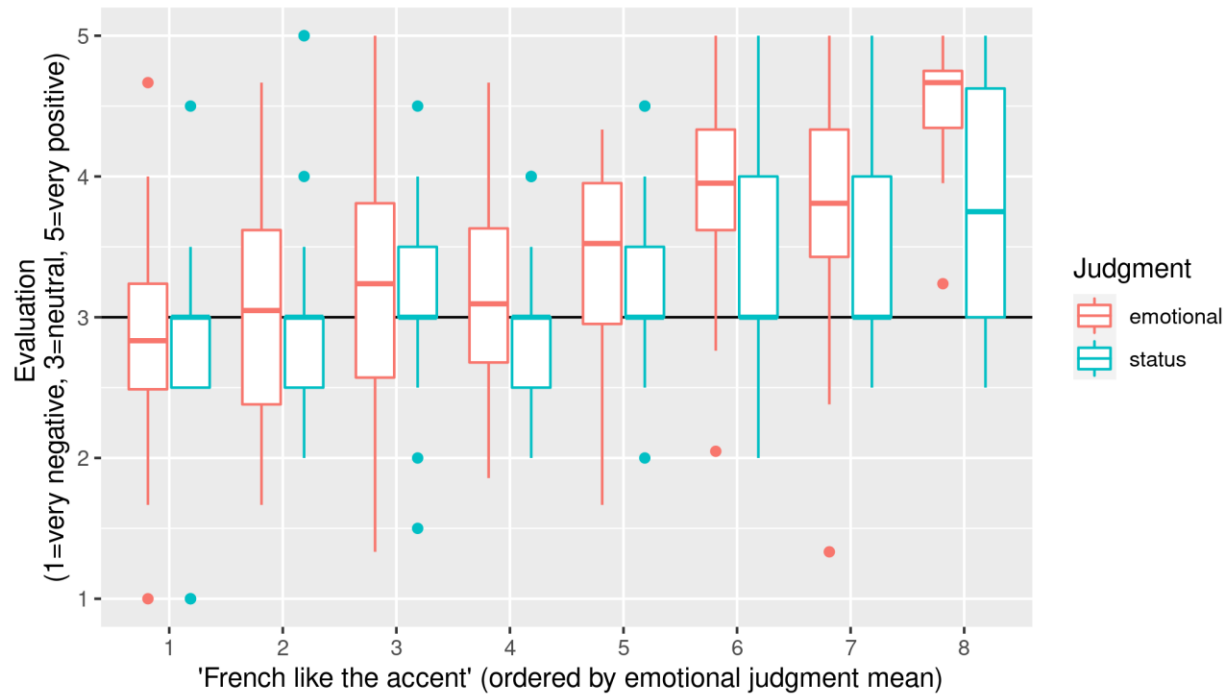

Figure 15: Distribution of emotional and status judgement depending on whether French people like the accent

Finally, Figure 16 shows that the emotional evaluation increases with participants sense of belonging to the region of the accent commented. There is a light tendency of emotional evaluation increasing with participants sense of belonging to the region, although we could have expected a starker tendency here. Belonging could be taken as a measure in the affirmation that participants rate their "own region" higher (Coupland and Bishop 2007), and thus we would expect a stronger correlation between belonging to the region and the evaluation of accent. Status evaluation is not driven by the sense of belonging to the region.

Figure 16: Evolution of judgment in relation to participant belonging to the region

Glottophobia

The frequency of remarks participants claim to receive concerning their accent is one of the most important factors to explain the status evaluation based on correct and serious scales, see Figure 17.
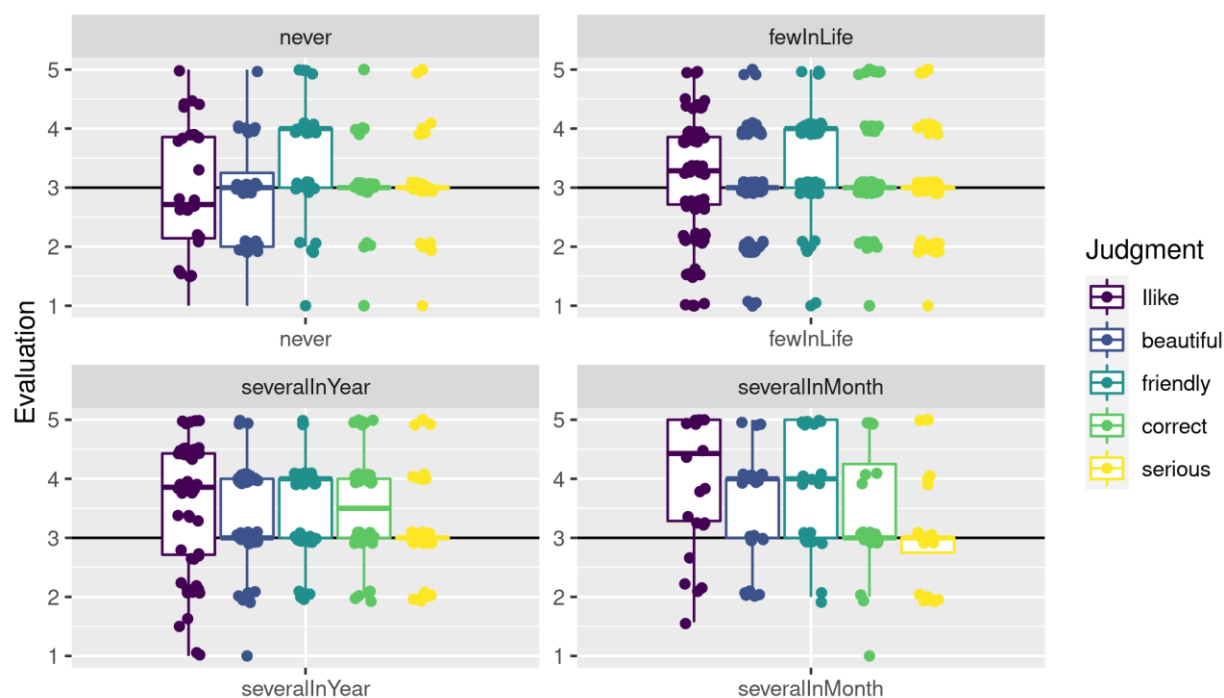
Figure 17: Detail of the evaluation of accent with the frequency of remarks

We note that the nature of the remarks is not significant, see Figure 18. That is to say, it did not matter whether the participants were made fun of, discriminated against or praised for their accent. A lot of participants who receive a lot of remarks mention that even though the comments are often said with a good intention, they are still annoying to them, as condescending. One participant sums up: *"Les remarques sont généralement positives pour ceux qui les expriment mais pas forcément pour moi, car remplies de clichés et caricaturales"* ["Remarks are usually positives for those who express them, but not necessarily for me as they are full of clichés and caricatural"]. Others relate quite traumatizing experience of teachers, from primary school to higher education, humiliating them to pronounce words as the Parisian norm, or being asked after a work presentation where they are from instead of questions related to their presentation topic.
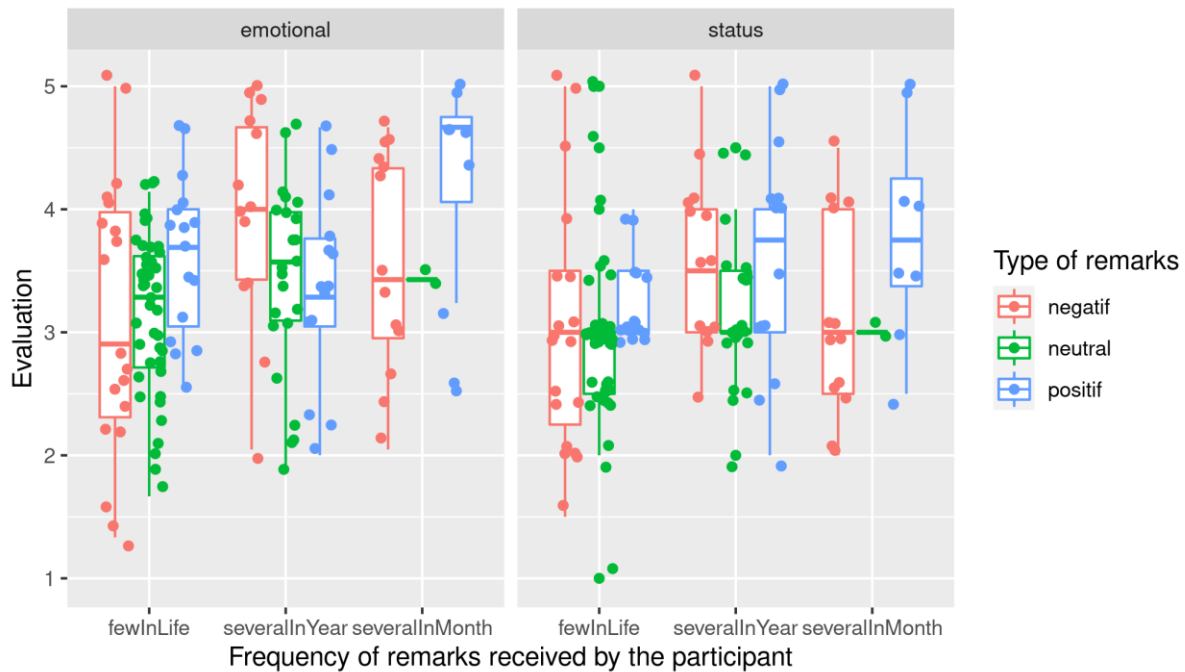
Figure 18: Evaluation of accent increases with the frequency of remarks

I argue that by rating the local accent high, participants are making a stance in reaction to the micro-aggressions and blatant glottophobia (Blanchet 2016). They disalign with French national ideology of centralization and prescriptivism towards standard French in order to reclaim pride in their region. In doing so, they revalue their local variety and their own identity.

On the other end of the scale, some participants had very low evaluations of accents. Most of them received few or no remarks related to their accent in their lives. When looking at the comments they wrote concerning the accent evaluation, it becomes clear that they do not want to be associated with people who have this accent. One participant even wrote that they change their accent *"pour m'éloigner d'une population qui ne me correspondait plus"* ["to distance [themselves] from a population that no longer matched [them]"].

51

Finally, it is interesting to note that even though the different regions are rated differently and do not all benefit from the same prestige, the region of origin, which I assume is where the participants got their accent if they have one, does not strongly influence the type of remarks received, see Figure 19. We nevertheless note that Bretagne, Normandie and Hauts-de-France have no or few positive remarks, all of which are poorly rated for status and emotional evaluations, while Nouvelle-Aquitaine, Occitanie and PACA receive more positive remarks. A closer investigation of those factors could be of interest for a later study.
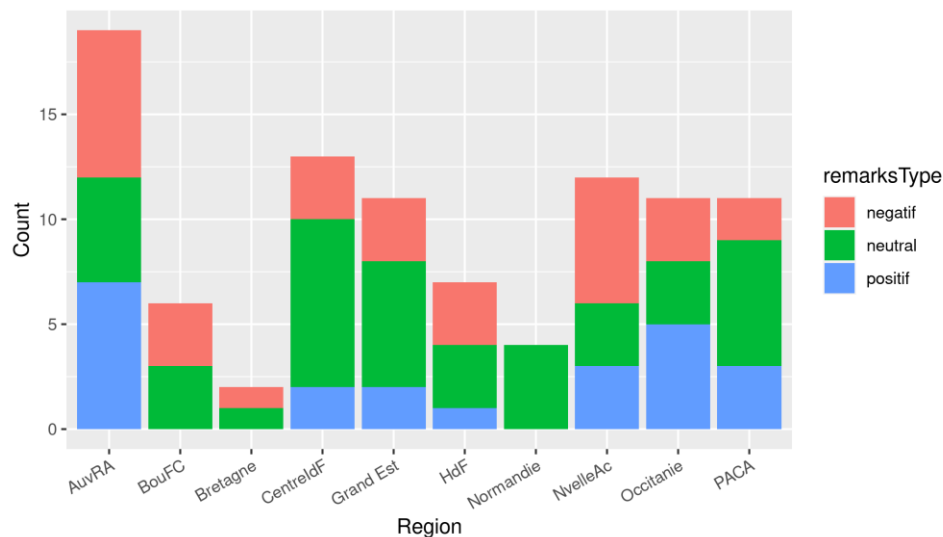

Figure 19: The type of remarks received is not exclusively specific to any region. Still North, Brittany and Normandy receive less positive remark while Southern regions receive more positive remarks

Power dynamics

I studied how participants evaluation of accent reflect their identity as member of a geographical area. The evaluations reflect also a reaction to the national ideologies towards non-standard accents. In this final section, I discuss the socio-economic implications of the accent evaluations.

Many participants mention changing their accent for work, or outside of their home in general, e.g. *"pour certaines professions, il est même conseillé de le perdre parce que 'pas assez sérieux' "* ["for some occupations, it is even advised to lose [the accent] because it is 'not serious enough' "]. They also talk about the fact that their accent "comes back" in situations when they are not in control, e.g. drunk or angry. That is to say that they need to control their speech and "soften their accent" in order to gain symbolic power on the linguistic marketplace (Bourdieu 1977).

Figure 20 shows that the lowest group on the social scale rates higher both on emotional and on status, while the lowest emotional ratings come mainly from the highest social group. The upper-class "cadres" reproduce the ideologies that benefit the standard variety speakers they might be part of. On the contrary, the lowest group negotiates the national ideology through their evaluation of accent. They refute the superiority of the standard or at any rate claim that the local variety is positive emotionally and status wise.
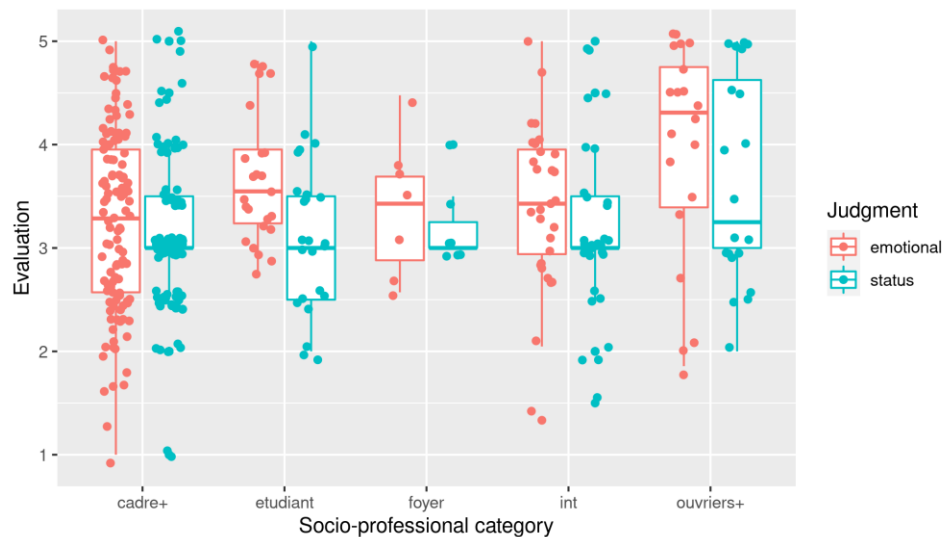

Figure 20: Evaluation of accent in the socio-professional categories

Figure 21 shows that there is a slight decrease of the status evaluation with age, also visible in Figure 13 (leaving aside Grand-Est region). This is unexpected given that younger speakers are less conservatives. This could be a sign of a general trends towards the devaluation of regional accents. Given the large range of evaluations, it is difficult to make any strong claims. Figure 22 shows nevertheless that it would be potentially interesting to explore the interaction of socio-professional categories with age.
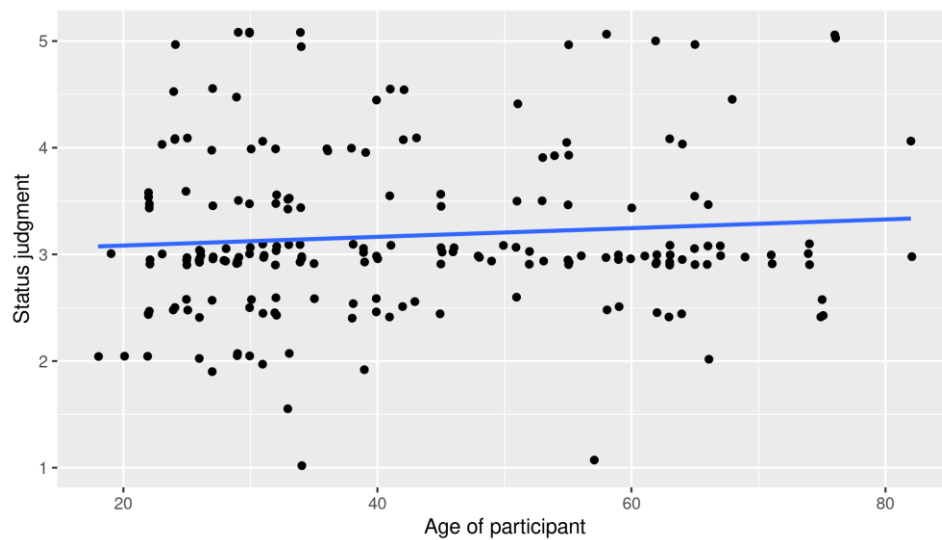

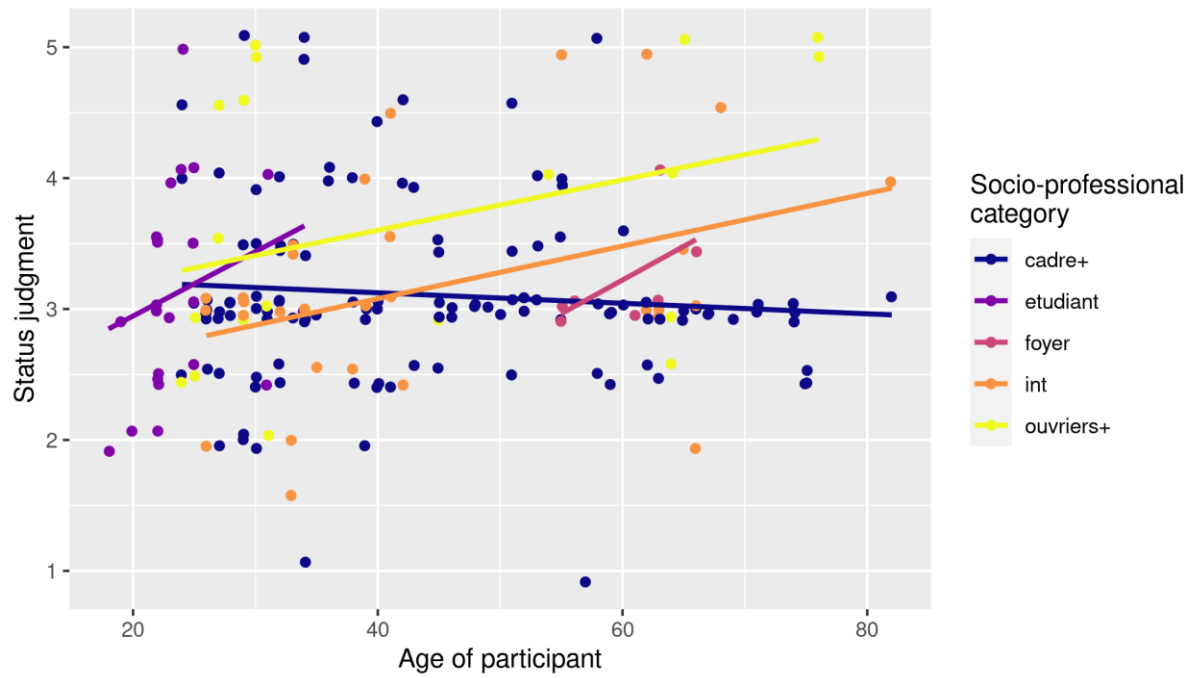Figure 21: Younger participants rate higher on emotional scale but lower on status scale

Figure 22: Evolution of the evaluation of accent with age in the socio-professional categories

Conclusion

My project consisted in looking at attitudes towards regional accents in metropolitan France. I focused on regions where people had lived during their lives. For this research project I built a survey distributed online.

In this report I first reviewed the existing literature on accents and attitudes towards accents, with a special focus on French speakers in metropolitan France. I then detailed the methodology I used to conduct the survey: how I built it, how I distributed it, and how I extracted variables to be able to run statistical analysis on the data I collected. In order to group some variables together, I conducted an initial data processing via Principal Component Analysis. In particular, it resulted in the creation of two dependent variables. The evaluations are indeed divided in two groups: status and emotional judgments. I presented the main statistical results and a discussion followed including additional results.

I found that the main factors driving the evaluations of accents are the regions where the accent is spoken, the amount of remarks received and social class elements. Concerning the region, the finding of Pustka et al. (2019) are confirmed: regions other than Paris benefit from overt prestige, and peripheries benefit from high emotional ratings. I noticed that there might be some changes concerning the region Grand-Est, while it still holds very negative stereotypes for some participants.

An interesting outcome of this study is the correlation between the number of remarks, but not their nature (positive or negative), on one's accent and the self-evaluation of this accent. When participants are made aware of their accent, they hold higher evaluations as a stance in reaction to the remarks they get. As the anecdotal quote from Gasquet-Cyrus (2013) says: *"pour avoir un*

accent, il faut être deux" ["it takes two to have an accent"]: in a (fictional) homogeneous group nobody has an accent.

Finally, the evaluations reflect power negotiation on the linguistic market: "ouvriers" and employees rate higher on the emotional scale, while negative judgments are mostly coming from "cadres".

Results of the survey on French accent suggest that there is no need to focus on native speakers of a variety in attitudes study. It could be interesting to expand the study to non-native speakers of French to check whether they are subject to the same ideologies than native French speakers.

The present report only started to explore the data that was obtained. There are several further possibilities to investigate. A different delimitation of the regions, closer to linguistic areas and based on Figure 22 and Figure 23, might be interesting to check. Some responses of the survey were not used in the results, such as the hobbies or the reason of the accent change. Examining the responses could bring more information on what drives the evaluations of accents in France. The comment sections are sometimes quoted in the Discussion section, but it would be worth analysing them more thoroughly. Numerous participants provided substantial comments, sharing personal experiences and thoughts, that I believe to be extremely valuable. This, with comments received outside of the online survey, indicates that there is a general public interest toward this question.
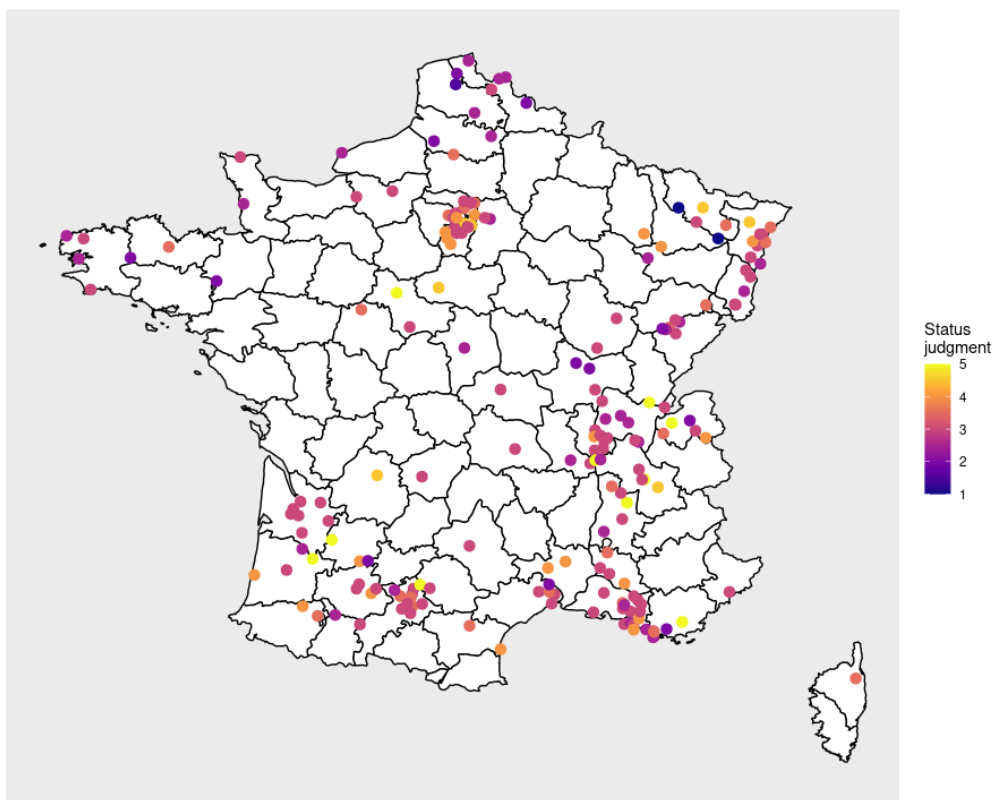
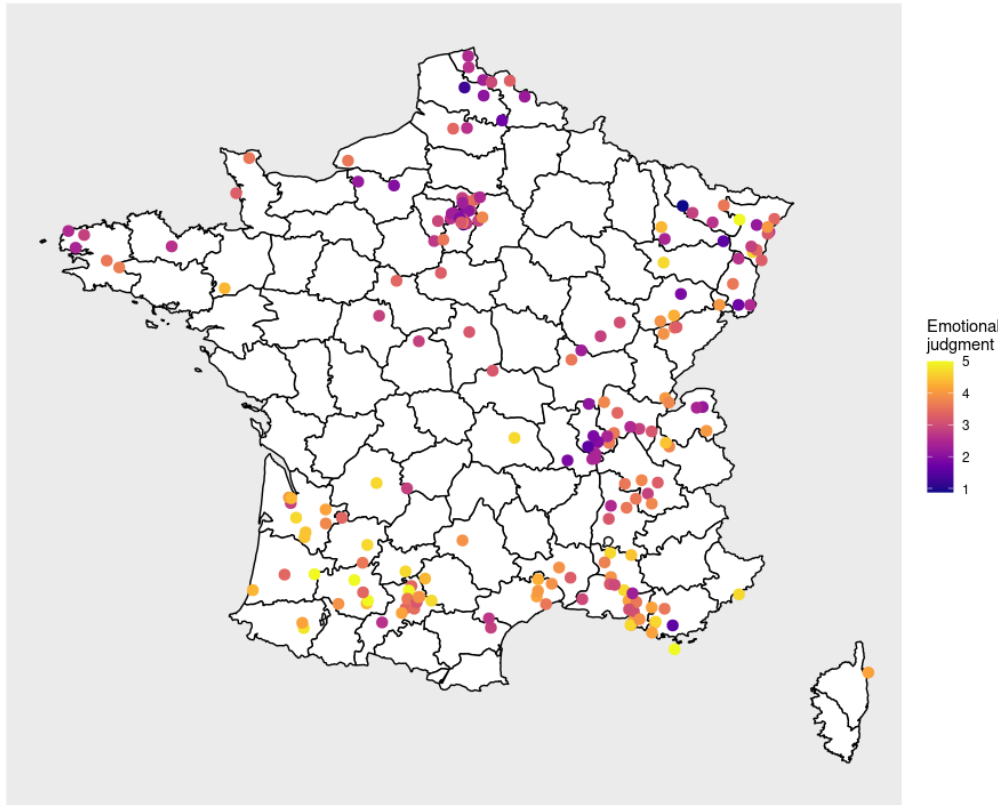Figure 22: Map of the participants status evaluation, for a later study

Figure 23: Map of the participants emotional evaluation, for a later study

Even focusing on what was taken into account in the results, many further analyses could be pursued. Only a few interactions were considered for the statistical models. Gender and age specifically could be interesting to study in interaction with other variables. This could help uncover changing hierarchies for instance: the change would potentially be reflected in the age group, as well as gender. Looking at estimated means might show that some variables have a significant effect only for one age group for instance. Many other interactions could potentially be significant, such as the interaction between the socio-professional categories of the participants and their parents. Indeed, as shown in the Discussion section, the evaluation can be explained as a reaction to remarks. The evaluation might not be similar for higher social class participants who

come from privileged environment compared to those who come from deprived environment. Similarly, as regional mobility, social mobility might polarise evaluations one way or the other.

Finally, the audio recordings would be interesting to analyse by following the PFC project process (Adda-Decker et al. 2006).

Thus, there are many avenues to explore based on the data collected by the means of the survey I built for this MA dissertation.

References

Adda-Decker, Martine, Philippe Boula de Mareüil, Cécile Woehrling, Gilles Adda, Eric Bilinski, L. Lamel & Olivier Galibert. 2006. PI-TCAN Programme interdisciplinaire CNRS "Traitement des connaissances, apprentissage et NTIC" - projet VARCOM.

Anwyl-Irvine, Alexander, Jessica Massonnié, Adam Flitton, Natasha Kirkham & Jo Evershed. 2019. Gorillas in our midst: gorilla.sc, a new web-based experiment builder. *Behavior Research Methods*.

Armstrong, Nigel & Timothy Pooley. 2010. *Social and linguistic change in European French*. Basingstoke: Palgrave Macmillan.

Armstrong, Nigel & Sharon Unsworth. 1999. Sociolinguistic variation in southern French schwa. *Linguistics,* 37 (1), 127–156.

Avanzi, Mathieu. 2018. Cartographier les régionalismes du français en Suisse romande et en France voisine à l'ère des sciences participatives. *Cahiers internationaux de sociolinguistique*, (2), 43–103.

Avanzi, Mathieu & Philippe Boula de Mareüil. 2019. Peut-on identifier perceptivement huit accents régionaux en français? La réponse des sciences participatives. *Glottopol,* 31, 53–73.

Baayen, R. Harald 2013. Multivariate statistics. *Research methods in linguistics* (pp. 337–372). Cambridge: Cambridge University Press.

Banta, Joshua A., Martin H. Stevens & Massimo Pigliucci. 2010. A comprehensive test of the 'limiting resources' framework applied to plant tolerance to apical meristem damage. *Oikos,* 119 (2), 359–369.

Baran, Dominika. 2014. Linguistic practice and identity work: Variation in Taiwan Mandarin at a Taipei County high school. *Journal of Sociolinguistics*, 18 (1), 32–59.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software.*

Blanchet, Philippe. 2016. *Discriminations: Combattre la glottophobie*. Paris: Éditions Textuel.

Blanchet, Philippe & Nigel Armstrong. 2006. The sociolinguistic situation of 'contemporary dialects of French' in France today: An overview of recent contributions on the dialectalisation of standard French. *Journal of French Language Studies*, 16 (3), 251–275.

Bouchet-Valat, Milan. 2014. Les évolutions de l'homogamie de diplôme, de classe et d'origine sociales en france (1969-2011): Ouverture d'ensemble, repli des élites. *Revue française de sociologie,* 55 (3), 459–505.

Bouchet-Valat, Milan. & Cyril Jayet. 2019. La mesure des classes sociales par les nomenclatures: Enjeux, problèmes et débats. *L'Année sociologique,* 69 (2), 311–331.

Bouchet-Valat, Milan, Camille Peugny & Louis-André Vallet. 2016. Inequality of educational returns in France: Changes in the effect of education and social background on occupational careers. *Education, occupation and social origin*. Cheltenham: Edward Elgar Publishing.

Boughton, Zoë. 2006. When perception isn't reality: Accent identification and perceptual dialectology in French. *Journal of French Language Studies*, 16 (3), 277–304.

Bourdieu, Pierre. 1977. The economics of linguistic exchanges. *Information (International Social Science Council),* 16 (6), 645–668.

Boyer, Henri. 2015. ''Vous venez d'où pour avoir cet accent?" : La communauté linguistique comme marché soumis à l'imaginaire de la communauté. *Lengas*, 77, 2–7.

Candea, Maria. 2017. La notion d''"accent de banlieue" à l'épreuve du terrain. *Glottopol*, (29), 13–26.

Candea, Maria, Gaëlle Planchenault & Cyril Trimaille. 2019. Accents du français : Approches critiques. *Glottopol*, 31.

Cavallaro, Francesco & Ng Bee Chin. 2009. Between status and solidarity in Singapore. *World Englishes,* 28 (2), 143–159.

Chambers, J. K. & Peter Trudgill. 1998. *Dialectology* (2nd). Cambridge: Cambridge University Press.

Chauvel, Louis. 2001. Le retour des classes sociales? *Revue de l'OFCE,* (4), 315–359.

Cheshire, Jenny. 2019. Taking the longer view: Explaining Multicultural London English and Multicultural Paris French. *Journal of Sociolinguistics*.

Cheshire, Jenny & Maria Secova. 2018. The origins of new quotative expressions: the case of Paris French. *Journal of French Language Studies*, 28 (2), 209–234.

Coupland, Nikolas & Hywel Bishop. 2007. Ideologised values for British accents. *Journal of sociolinguistics*, 11 (1), 74–93.

Dörnyei, Zoltan. 2003. *Questionnaires in second language research: Construction, administration, and processing.* Mahwah, NJ: Lawrence Erlbaum Associates.

Drummond, Rob. 2013. The Manchester polish STRUT: Dialect acquisition in a second language. *Journal of English linguistics*, 41 (1), 65–93.

Du Bois, John W. 2007. The stance triangle. *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction,* 164 (3), 139–182.

Dubois, Sylvie & Barbara M. Horvath. 1998. Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change,* 10 (3), 245–261.

Durand, Jacques, Catherine Slater & Hilary Wise. 1987. Observations on schwa in southern French. *Linguistics,* 25 (5), 983–1004.

Durand, Jacques, Bernard Laks & Chantal Lyche. 2009. Le projet PFC (phonologie du français contemporain): une source de données primaires structurées. *Phonologie, variation et accents du français*, 19–61.

Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school.* New York: Teachers College Press.

Eckert, Penelope. 1997. Age as a sociolinguistic variable. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 151–167). Oxford: Blackwell.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41, 87–100.

Encrevé, Pierre & Michel Braudeau. 2007. *Conversations sur la langue française*. Paris: Editions Gallimard.

Gadet, Françoise. 2007. *La variation sociale en français*. Paris: Editions Ophrys.

Garrett, Peter, Nikolas Coupland & Angie Williams. 1999. Evaluating dialect in discourse: Teachers' and teenagers' responses to young English speakers in Wales. *Language in Society,* 28 (3), 321–354.

Gasquet-Cyrus, Méderic. 2012. La discrimination à l'accent en France: Idéologies, discours et pratiques. In Trimaille Cyril, Eloy Jean-Michel (ed.) *Idéologies linguistiques et discriminations*. Paris: L'Harmattan, Carnets d'Atelier de Sociolinguistique, 6, 227–245.

Gasquet-Cyrus, Méderic. 2013. Perspectives dynamiques sur la ségrégation sociolinguistique en milieu urbain: le cas de Marseille. *Glottopol*, 21, 9–21.

Giles, Howard. 1970. Evaluative reactions to accents. *Educational review*, 22 (3), 211–227.

Gilliéron, Jules & Edmond Edmont. 1902-1910. *Atlas linguistique de la France* (Vol. 13). Paris: H. Champion.

Hall, Damian. 2018. Dialects of French. *The handbook of dialectology* (pp. 474–485). Oxford: Wiley Online Library.

Hall, Damian. 2013. The linguistic geography of the French of northern France: Do we have the basic data? *Language and Linguistics Compass,* 7 (9), 477–499. Kroskrity, Ed. Regimes of Language: Ideologies, Polities, and Identities.

Irvine, Judith T. & Susan Gal. 2000. Language ideology and linguistic differentiation. In Paul V. Kroskrity, (ed.) *Regimes of Language: Ideologies, Polities, and Identities.* Santa Fe: School of American Research Press.

Kuiper, Lawrence. 1999. Parisian Perceptions of Regional French. *Handbook of Perceptual Dialectology*, 1, 243–262.

Kuiper, Lawrence. 2005. Perception is reality: Parisian and Provençal perceptions of regional varieties of French. *Journal of Sociolinguistics*, 9 (1), 28–52.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software,* 82 (13), 1–26.

Labov, William. 1966. *The social stratification of English in New York city*. Washington, D.C.: Center For Applied Linguistics.

Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, 60 (1), 44.

Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English dialects app: The creation of a crowdsourced dialect corpus. *Ampersand*, 5, 1–17.

Levon, Erez, Devyani Sharma, Dominic Watt & Christina Perry. 2020. *Accent bias: Implications for professional recruiting.*

Lippi-Green, Rosina. 1997. *English with an accent: Language, ideology, and discrimination in the United States.* Abingdon-on-Thames: Psychology Press.

Lippi-Green, Rosina. 2012. The myth of non-accent. *English with an accent: Language, ideology and discrimination in the united states* (pp. 58–67). New York: Routledge.

Moyer, Alene. 2007. Do language attitudes determine accent? A study of bilinguals in the USA. *Journal of multilingual and multicultural development*, 28 (6), 502–518.

Niedzielski, Nancy A. & Dennis R. Preston. 2010. *Folk linguistics* (Vol. 122). Berlin and New York: Walter de Gruyter.

Oakes, Leigh & Jane Warren. 2007. *Language, citizenship and identity in Quebec*. Berlin and Heidelberg: Springer.

Ochs, Elinor. 1992. Indexing gender. *Rethinking context: Language as an Interactive Phenomenon,* 11, 335.

Ostby, Katherine Asla. 2016. *Les voyelles moyennes à double timbre dans le parler de la haute bourgeoisie parisienne : analyse acoustique et diachronique* (Doctoral dissertation). Université Paris Ouest Nanterre La Défense.

Paltridge, John & Howard Giles. 1984. Attitudes towards speakers of regional accents of French: Effects of regionality, age and sex of listeners. *Linguistische Berichte*, (90), 71–85.

Payne, Geoff. 2013. Models of contemporary social class: The great British class survey. *Methodological Innovations Online,* 8 (1), 3–17.

Peirce, Charles S. 1932. The icon, index, and symbol. Collected Papers of Charles Sanders Peirce. C. Hartshorne and P. Weiss.

Penissat, Étienne. 2012. La difficile production d'une nomenclature socioprofessionnelle à l'échelle européenne. *Revue française de socio-économie,* (2), 251–257.

Podesva, Robert. J. & Janneke Van Hofwegen. 2014. How conservatism and normative gender constrain variation in inland California: The case of/s. *University of Pennsylvania Working Papers in Linguistics*, 20 (2), 15.

Pustka, Elissa. 2009. A prototype-theoretic model of Southern French. *Kate Beeching, Nigel Armstrong and Françoise Gadet, eds*, 77–94.

Pustka, Elissa. 2011. L'accent méridional: Représentations, attitudes et perceptions toulousaines et parisiennes. *Lengas. Revue de sociolinguistique,* (69), 117–152.

Pustka, Elissa., Jean-David Bellonie, Marc Chalier & Luise Jansen. 2019. "C'est toujours l'autre qui a un accent" : Le prestige méconnu des accents du Sud, des Antilles et du Québec. *Glottopol*, 31, 27–52.

Sankoff, Gillian. 2004. Adolescents, young adults and the critical period: Two case studies from 'Seven Up'. *Sociolinguistic Variation: Critical reflections*, 121–39.

Savage, Mike, Fiona Devine, Niall Cunningham, Mark Taylor, Yaojun Li, Johs Hjellbrekke, Brigitte Le Roux, Sam Friedman & Andrew Miles. 2013. A new model of social class? Findings from the BBC's Great British Class Survey experiment. *Sociology,* 47 (2), 219–250.

Schleef, E. 2014. Written surveys and questionnaires in sociolinguistics. *Research methods in sociolinguistics: A practical guide*, 42–57.

Secova, Maria, Penelope Gardner-Chloros & Frédérique Atangana. 2018. 'Il parle normal, il parle comme nous': Self-reported usage and attitudes in a banlieue. *Journal of French Language Studies,* 28 (2), 235–263.

Sharma, Devyani. 2017. Scalar effects of social networks on language variation. *Language Variation and Change,* 29 (3), 393–418.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23 (3-4), 193–229.

Sobotta, Elissa. 2006. *Phonologie et Migration : Aveyronnais et Guadeloupéens à Paris* (Doctoral dissertation). Paris 10 - Sciences du Langage.

Tagliamonte, Sali A. & R Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language variation and change,* 24 (2), 135–178.

Tagliamonte, Sali A. & Alexandra D'Arcy. 2007. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change*, 19 (2), 199–217.

Trudgill, Peter. 1974. *The social differentiation of English in Norwich* (Vol. 13). Cambridge: Cambridge University Press.

Trudgill, Peter. 1983. Sex and covert prestige: Linguistic change in the urban dialect of Norwich. *On Dialect*, 169–185.

Vallot, Pauline. 2019. Le diplôme et la profession: Deux composantes de la classe sociale en France et en Allemagne. *L'Année sociologique,* 69 (2), 425–450.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. *Directions for historical linguistics* (pp. 95–188). Austin: University of Texas Press.

Wickham, Hadley 2016. *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wolfram, Walter A. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, D.C.: Center For Applied Linguistics.

Woolard, Kathryn A. 1998. Introduction: Language ideology as a field of inquiry. *Language ideologies: Practice and theory*, 3 (11).

Annex A: Survey Screenshots

Enquête sur les accents en France

Bonjour et bienvenue !

Merci pour l'intérêt que vous portez à notre étude.

Cette étude est menée par Chloé Vincent dans le cadre d'un projet de recherche à l'Université Queen Mary de Londres, encadré par Prof. Devyani Sharma.

L'objectif de cette étude est de recenser ce que les gens dont la langue maternelle est le français pensent des accents qui les entourent dans les lieux où ils vivent.

L'enquête est organisée en 3 parties:

1. Nous vous demandons dans un premier temps de **commenter les accents de lieux en France où vous avez vécu** pendant trois ans ou plus.
2. Vous aurez ensuite la possibilité de **vous enregistrer en lisant** un texte à voix haute, via le micro de votre ordinateur. *Cette partie est optionnelle.*
3. Enfin nous vous demanderons quelques informations sur votre **profil**.

L'enquête dure entre 20 et 30 minutes, selon si vous souhaitez commenter un ou plusieurs accents et si vous souhaitez participer à la lecture à voix haute.

Toutes les données transmises dans le cadre de cette enquête sont **anonymisées** : nous n'établissons aucun lien entre vos réponses et votre identité. Aucune information vous concernant ne sera transmise à un tiers. Si vous ne finissez pas l'enquête, aucune donnée ne sera gardée.

Pour plus d'informations, si vous souhaitez connaître les résultats de l'étude, ou pour rapporter des problèmes liés à l'enquête, vous pouvez envoyer un mail à Chloé Vincent, à l'adresse enquete.accents.france@gmail.com.

En participant à cette expérience, je consens aux points suivants :

1. Je suis d'accord pour participer à cette étude
2. J'ai plus de 18 ans
3. Ma langue maternelle est le français
4. J'ai lu et compris les informations ci-dessus

☐ J'ai vécu plus de 3 ans dans un endroit où je trouve qu'il y a un accent

Si vous ne cochez pas la case ci-dessus, la participation à l'enquête ne sera pas possible, un message s'affichera en anglais vous informant que le consentement est obligatoire pour poursuivre l'enquête ("You must consent to continue")

<div align="center">

J'accepte et je participe !

</div>

Partie 1 : Commentaires sur les accents français

Merci de participer à cette étude. Cette première partie dure entre 5 et 10 minutes.

Il n'est pas possible de revenir à la page précédente lors de l'enquête. Un retour en arrière annulera toutes les réponses.

Il n'y a pas de bonne ou de mauvaise réponse, répondez le plus honnêtement possible. Nous souhaitons simplement savoir ce que les gens pensent des différentes façons de parler le français.

Remarque spéciale confinement COVID-19 : Nous nous intéressons à vos pratiques et commentaires habituels, répondez en considérant vos pratiques hors-confinement.

Pour rappel, toutes les données sont anonymisées.

**Question obligatoire*

Lieux où vous avez vécu

Pouvez-vous renseigner les endroits où vous avez vécu 3 ans ou plus, en indiquant le code postal suivi de la période de votre vie où vous y avez habité.

Remarque: l'exactitude des périodes n'est pas essentielle, des tranches d'ages approximatives nous suffiront.

Codes postaux et périodes *, *par exemple:*
73290 0 - 6 ans
71000 de 6 à 17 ans
38000 entre 19 et 25 ans ① ③
...

[ ]

[ ]

[ ]

[ ]

[ ]

[ ]

[ ]

[ ]

[ ]

[ ]

Dans les pages suivantes, nous allons vous demander vos commentaires sur l'accent d'un endroit **où vous avez habité pendant plus de 3 ans**. Vous pourrez commenter jusqu'à 3 endroits, si vous le souhaitez.

Choisissez la région du lieu que vous voulez commenter en premier dans la liste déroulante ci-dessous : *

[Choisir une région française ▾] ②

Si vous n'avez pas répondu à toutes les questions obligatoires, un message en anglais vous demandera d'y répondre ("Please complete all the fields")

[Commencer l'enquête !]

Commentaires sur la région Normandie

Nous allons vous poser quelques questions sur ce que vous pensez de l'accent qui vous entourait dans la région Normandie.

Il n'y a pas de bonne ou de mauvaise réponse, nous souhaitons recueillir vos impressions personnelles ! Rappelez-vous, cette enquête est anonyme.

*Question obligatoire

L'accent et vous

Où habitiez-vous dans la région Normandie ? *(par exemple : quartier populaire de Lyon, ou dans la campagne autour de Besançon)* *

Pensez-vous que les gens de cet endroit ont un accent ? (1= pas du tout d'accent, 6= accent très fort) *
Vous ne le parlez pas forcément mais vous l'entendiez quotidiennement

Pas du tout d'accent	1	2	3	4	5	6	Accent très prononcé

Décrivez approximativement l'aire géographique où on parle l'accent auquel vous pensez *(par exemple: "il est parlé par les gens qui habitent Marseille / à l'ouest de Rennes / dans la campagne autour de Reims ...")* *

Pouvez-vous décrire cet façon de parler à l'aide de plusieurs adjectifs ? *(par exemple: chantant, lent...)*

Parlez-vous avec cet accent ? (1= pas du tout d'accent, 6= accent très fort) *

Pas du tout d'accent	1	2	3	4	5	6	Accent très prononcé

Avez-vous parlé avec cet accent dans d'autres contextes ? *(par exemple: quand vous étiez plus jeune, quand vous allez dans certains lieux, quand vous êtes ému.e)*

Quelle proportion de vos proches (amis, famille) dans cette région parlaient avec l'accent ? *

Aucun	Quelques uns	La moitié	La plupart	Tous

Que pensez vous de cet accent ?

Quelle est votre impression vis à vis de cette façon de parler avec accent ? Complétez :

*Cette façon de parler est ... *

très incorrecte	plutôt incorrecte	ni correct ni incorrecte	plutôt correcte	très correcte

très ridicule	plutôt ridicule	ni ridicule ni sérieuse	plutôt sérieuse	très sérieuse

très antipathique	plutôt antipathique	ni antipathique ni sympathique	plutôt sympathique	très sympathique

très belle	plutôt belle	ni belle ni laide	plutôt laide	très laide

très marquée socialement	plutôt marquée socialement	neutre	plutôt pas marquée socialement	pas du tout marquée socialement

Pour les questions suivantes, il s'agit de dire si vous êtes d'accord avec les affirmations proposées. (1= pas du tout d'accord, 8 = tout à fait d'accord)

*J'aime cette façon de parler *

Pas du tout d'accord	1	2	3	4	5	6	7	8	Tout à fait d'accord

*Les français en général aiment cette façon de parler *

Pas du tout d'accord	1	2	3	4	5	6	7	8	Tout à fait d'accord

Facultatif : Avez-vous des commentaires sur cet accent, et votre rapport à celui-ci ?

Si vous n'avez pas répondu à toutes les questions obligatoires, un message en anglais vous demandera d'y répondre ("Please complete all the fields")

Page suivante

74

Votre rapport à la région Normandie

Dans cette partie du questionnaire nous nous intéressons à votre rapport à la région Normandie où vous avez vécu.

Pour les questions suivantes, il s'agit de dire si vous êtes d'accord avec les affirmations proposées. (1= pas du tout d'accord, 6 = tout à fait d'accord)

Il n'y a pas de bonnes ou de mauvaise réponse. *Tout est anonymisé.*

*Question obligatoire

J'ai des attaches dans cette région * (10)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

Je me sens vraiment faire partie de cette région * (11)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

J'ai de bons souvenirs dans cette région * (12)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

J'aime les gens de cette région * (13)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

J'aime cette région * (14)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

Les français en général aiment cette région * (15)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

Facultatif : Avez-vous des commentaires à faire sur la période que vous avez passé dans cette région, ou sur votre rapport à cette région ?

[text box   ]

Si vous n'avez pas répondu à toutes les questions obligatoires, un message en anglais vous demandera d'y répondre ("Please complete all the fields")

[Page suivante]

Merci d'avoir commenté un accent de la région Normandie.

Vous pouvez maintenant, au choix:

1. Commenter l'accent d'une autre région dans laquelle vous avez habité
2. Passer à la partie lecture à haute voix (3-4 minutes)
3. Aller directement à la dernière partie (3-4 minutes)

En poursuivant vers la page de **lecture à voix haute**, vous nous autorisez à garder et traiter l'enregistrement de la lecture du texte dans le cadre de ce projet de recherche. *L'enregistrement reste anonyme.*

[Commenter un autre accent]   [Lecture à voix haute]   [Dernière partie directement]

Partie 2 : Lecture à voix haute

Dans cette section, on vous demande de lire un petit article de journal fictif. Il n'y a pas de difficulté de lecture particulière.

Une première étape consiste à tester que votre micro fonctionne bien.

Ensuite, le texte vous sera présenté et vous aurez le temps de le lire silencieusement si vous le souhaitez, avant de vous enregistrer.

Il sera peut-être nécessaire d'autoriser l'utilisation du micro par votre navigateur web.

Si vous souhaitez quitter la partie lecture à voix haute, sans perdre les informations que vous avez renseignées jusqu'ici, vous pouvez toujours passer aux pages suivantes en cliquant sur le bouton Continuer ! Ne revenez pas à la page précédente, tout serait perdu.

Rappel : Toutes les données récoltées sont anonymisées

Continuer !

Test du micro

Quand vous serez prêt.e, enregistrez-vous en cliquant sur "Start recording". Vous avez 3 secondes pour dire "Test 1 2 3" ou ce que vous voulez.

Au bout des 3 secondes, vous pourrez vérifier que l'enregistrement s'est bien passé avec le bouton lecture, et si besoin vous pouvez tester votre micro à nouveau en cliquant sur "Re-record".

Une fois le test terminé, vous pouvez passer à la lecture du texte.

Si votre micro ne fonctionne pas, vous pouvez passer aux pages suivantes directement en cliquant sur le bouton Continuer !, mais ne cliquez pas sur le bouton retour de votre navigateur.

Quand vous cliquez sur "Start Recording", le bouton disparaît,
l'enregistrement a commencé

● Start Recording

Continuer !

Le Premier Ministre ira-t-il à Beaulieu ?

Le village de Beaulieu est en grand émoi. Le Premier Ministre a en effet décidé de faire étape dans cette commune au cours de sa tournée de la région en fin d'année. Jusqu'ici les seuls titres de gloire de Beaulieu étaient son vin blanc sec, ses chemises en soie, un champion local de course à pied (Louis Garret), quatrième aux jeux olympiques de Berlin en mil neuf cent trente-six, et plus récemment, son usine de pâtes italiennes. Qu'est-ce qui a donc valu à Beaulieu ce grand honneur ? Le hasard, tout bêtement, car le Premier Ministre, lassé des circuits habituels qui tournaient toujours autour des mêmes villes, veut découvrir ce qu'il appelle « la campagne profonde ».

Le maire de Beaulieu – Marc Blanc – est en revanche très inquiet. La cote du Premier Ministre ne cesse de baisser depuis les élections. Comment, en plus, éviter les manifestations qui ont eu tendance à se multiplier lors des visites officielles ? La côte escarpée du Mont Saint-Pierre qui mène au village connaît des barrages chaque fois que les opposants de tous les bords manifestent leur colère. D'un autre côté, à chaque voyage du Premier Ministre, le gouvernement prend contact avec la préfecture la plus proche et s'assure que tout est fait pour le protéger. Or, un gros détachement de police, comme on en a vu à Jonquière, et des vérifications d'identité risquent de provoquer une explosion. Un jeune membre de l'opposition aurait déclaré : « Dans le coin, on est jaloux de notre liberté. S'il faut montrer patte blanche pour circuler, nous ne répondons pas de la réaction des gens du pays. Nous avons le soutien du village entier. » De plus, quelques articles parus dans les journaux locaux indiqueraient que des activistes des communes voisines préparent une journée chaude au Premier Ministre. Quelques fanatiques auraient même entamé un jeûne prolongé dans l'église de Saint Martinville.

(16)

Quand vous cliquez sur "Start Recording", le bouton disparaît, l'enregistrement a commencé

● Start Recording

Continuer !

Dernière partie : informations sur votre profil

Nous vous rappelons que ces informations restent strictement anonymisées

Avant de terminer l'enquête, merci de renseigner les informations suivantes pour nous permettre d'avoir une idée de votre profil.

N'oubliez pas de valider à la fin de la page.

Question obligatoire

Pensez vous qu'il est important d'être bien intégré dans la vie locale ? (1= pas du tout important, 6= extrêmement important) * (17)

Pas du tout important	1	2	3	4	5	6	Extrêmement important

Comment aimez-vous passer votre temps libre - hors confinement ? *(par exemple : sport, séries, cuisine, sorties, associatif... N'hésitez pas à préciser)* *

Diriez-vous que vous êtes plutôt social ou solitaire, extraverti ou introverti ?

Pensez-vous avoir changé d'accent au cours de votre vie ? *

○ Oui
○ Non   (18)

Facultatif : Si oui, pour quelles raisons ?

Vous fait-on souvent des remarques sur votre façon de parler ? * (19)

Jamais	Quelques fois dans ma vie	Plusieurs fois par an	Plusieurs fois par mois	Plusieurs fois par semaine

Si c'est arrivé, ces remarques sont-elles positives ou négatives, sont-elles liées à votre accent ? *Vous pouvez donner des exemples d'anecdotes que vous avez vécues, si vous le souhaitez.*

(20)

Êtes-vous d'accord avec l'affirmation suivante ? (1= pas du tout d'accord, 6 = tout à fait d'accord) *

J'aime entendre différents accents ? (21)

Pas du tout d'accord	1	2	3	4	5	6	Tout à fait d'accord

Vous êtes : *

○ Une femme
○ Un homme   (22)
○ Autre
○ Ne souhaite pas répondre

Quel est le plus haut diplôme que vous ayez obtenu ? *

Choisir le diplôme   (23)

Facultatif : Précisez l'intitulé de ce diplôme : *(par exemple: CAP de boulangerie)*

(24)

Quelle est votre catégorie socio-professionnelle : *(si vous êtes à la retraite ou au chômage, indiquez votre activité précédant la retraite ou le chômage)* *

Choisir la catégorie   (25)

Annex B: Variables

The number and letters following "Q" in brackets refer to the question number from the screenshots in Annex A.

Dependent variables

Judgment towards accent was rated thanks to Likert scales. The participants were asked to complete the following sentence: "This way of speaking is ..." ("*Cette façon de parler est...*"). The scales gave the following options "very (negative term)", "quite (negative term)", "neither (negative term) nor (postive term)", "quite (positive term)", "very (positive term)" (*"très", "plutôt", "ni ... ni"*).

correct Likert scale with five options from very incorrect (*"très incorrect"*) to very correct (*"très correct"*)   (Q B)

serieux (serious) Likert scale with five options from very ridiculous (*"très ridicule"*) to very serious (*"très sérieuse"*)   (Q C)

sympa (friendly) Likert scale with five options from very unfriendly (*"très antipathique"*) to very friendly (*"très sympathique"*)   (Q D)

beau (beautiful) Likert scale with five options from very ugly (*"très laide"*) to very beautiful (*"très belle"*)   (Q E)

marque (socially-marked) Likert scale with five options from very socially marked (*"très marquée socialement"*) to not socially marked at all (*"pas du tout marquée socialement"*) (Q F)

The participants were asked to rate their agreement on the sentence "I like this way of speaking" (*"J'aime cette façon de parler"*), on an eight points Likert scale from "do not agree at all" (*"Pas du tout d'accord"*) to "completely agree" (*"Tout à fait d'accord"*). This resulted in the variable **jAime** (Q G)

As mentioned in Methodology section, a final accent judgment was based on the description of the accent (*"Pouvez-vous décrire cet [sic] façon de parler à l'aide de plusieurs adjectifs? (par exemple: chantant, lent...)"*) (Q A). I coded those description as positive, negative or neutral. This variable was dropped from the analysis, as explained in the Methodology section.

The PCA analysis, see Initial Data Processing section for details, resulted in two dependent variables:

emotionalJudgment as a combination of sympa, beau and jAime

statusJudgment as a combination of correct and serieux

Independent variables

accentArea based on the description of the geographical area where the accent is spoken (Q 6). Levels are as follow:

- loc: very localised description of accent (a city, a valley)

- dep: scale of a department size-wise (even if not corresponding to an official administrative department)

- reg: scale of multiple department - might be a former region, or a list of departments

- area: a broad area, such as "South-West", "all cities"

accentStrength.audio based on the audio recordings (Q 16), I rated the accents on a scale from 1 to 4, as follow — 20 % of the recordings were blind checked by another researcher (Célia Richy), only one recording was rated differently:

1. cannot hear any accent

2. very subtle, on a few words

3. audible accent

4. strong accent

accentStrength.region six point Likert scale from "no accent at all" ("*Pas du tout d'accent*") to "very strong accent" (*"Accent très prononcé"*), rating whether people from the area have an accent (Q 5)

accentStrength.self six point Likert scale from "no accent at all" ("*Pas du tout d'accent*") to "very strong accent" ("*Accent très prononcé*"), rating whether the participant think they have an accent (Q 7)

age deduced from the year of birth (Q 27)

csp Socio-professional category from INSEE (Institut National de la Statistique et des Études Économiques), classic way to get socio-economic status in French survey – student and stay at home parent categories were added as they recover different realities that "long term unemployed" in term of social class" (Q 25). Levels are reduced to the following:

- "cadre+": managers, executives, independent workers

- "int": intermediate professions

- "ouvriers+": working class

- "foyer": stay at home parent

- "etudiant": student

csp.dad see csp, of the father (Q 29)

csp.mum see csp,of the mother (Q 30)

diplomeType type of degree obtained (Q 24). Levels are:

- lit: litterature, linguistics

- hum: other humanities

- eco: economy

- pol: politics, law

- ser: care, doctors, nurses

- eng: engineers, scientists

frenchLikeAccent eight point Likert scale to rate the agreement with the sentence "French people in general like this way of speaking" ("*Les français en général aiment cette façon de parler*") (Q 9)

frenchLikeRegion eight point Likert scale to rate the agreement with the sentence "French people in general like this region" ("*Les français en général aiment cette région*") (Q 15)

friendsFamilyAccentDiversity proportion of friends and family who speak with the accent in this region (Q 8): "None", ("*Aucun*"), "a few" ("*Quelques uns*"), "half" (*"La moitié"*), "most" ("*La plupart*"), "all" ("*Tous*")

gender response to the question "You are: a woman, a man, other, does not wish to answer" (*"Vous êtes: une femme, un homme, autre, ne souhaite pas répondre*") (Q 22)

hasAccentChangedInLife response to the question "Do you think you have changed your accent during your life?" ("*Pensez-vous avoir changé d'accent au cours de votre vie ?*") (Q 18)

highestDegree highest obtained degree (Q 23). Levels are as follow:

- BacOrLess: highschool degree of less
- Bac2: two years degree after highschool
- Bac34: undergrad equivalent
- Bac5: master degree equivalent
- PhD

ID unique ID by participants provided by Gorilla

isCountryside based on the location details of where participants lived, indicates whether is was in a rural (TRUE) or urban (FALSE) environment (Q 4)

isSameAsOrigin based on participants life trajectories given with postcodes and periods (Q 3), I deduced the region were each participant had grown up. This variable indicates whether the origin region is the same as the one commented. When the origin region was not deductible, because the participant moved a lot for instance, isSameAsOrigin is FALSE.

likeHearingAccents corresponds to the sociolinguistic diversity scale proposed by Coupland (2007). Participants had to rate their agreement, on a six point Likert scale, with the sentence "I like hearing different accents" ("*J'aime entendre différents accents"*), from "do not agree at all" (*"Pas du tout d'accord"*) to "completely agree" ("*Tout à fait d'accord*") (Q 21). Level 1, 2, 3 were grouped together under "mostly not".

localIntegrationImportance six point Likert scale to answer the question whether it is important to be integrated in the local life ("*Pensez-vous qu'il est important d'être bien intégré dans la vie locale"*), from "not important at all" (*"Pas du tout important")* to "extremely important" ("*Extrêmement important*") (Q 17). Levels 1 and 2 were grouped together under "not important"

otherLanguageSpoken deduced from the languages the participants listed as fluent (Q 28). Levels are "None", "one", "two" and "three or more"

region administrative region of France (Q 2), only the metropolitan regions, Corsica excepted, remain:

- Hauts-de-France
- Grand Est
- Auvergne-Rhône-Alpes
- CentreIdF
- Bourgogne-Franche-Comté

- Provence-Alpes-Côte d'Azur

- Occitanie

- Nouvelle-Aquitaine

- Normandie

- Bretagne

regionBelonging build from the variables **attachesRegion** (Q 10) and **fairePartieRegion** (Q 11), which are based on six point Likert scales to rate the agreement to the following sentences "I have attaches in this region" ("*J'ai des attaches dans cette région"*), and "I really feel like I am part of this region" ("*Je me sens vraiment faire partie de cette région*").

regionEmotionalLink build from the variables **bonSouvenirsRegion** (Q 12) **jaimeLesGensRegion** (Q 13) and **jaimeLaRegion** (Q 14), which are based on six point Likert scales to rate the agreement to the following sentences "I have good memories in this region" (*"J'ai de bons souvenirs dans cette région"*), "I like the people of this region" ("*J'aime les gens de cette région*"), and "I like this region" ("*J'aime cette région*").

remarksFrequency how often participants get remarks on the way they speak (Q 19). Levels are:

- "Never" ("*Jamais"*)

- "Few times in life" (*"Quelques fois dans ma vie"*)

- "Several times a year" ("*Plusieurs fois par an"*)

- "Several times a month" (grouping of "*Plusieurs fois par mois*" and *"Plusieurs fois par semaine"*)

remarksType detail on the remarks they get: positive, negative or neutral (Q 20).

wage Income of participant household (Q 26). Levels are:

- 0 to 15000 €

- 15000 € to 25000 €

- 25000 € to 35000 €

- 35000 € to 50000 €

- more than 50000 €

Annex C: Missing values

In the data set constructed from the survey, some variables had missing values because some questions were not compulsory. One participant responded: "do not want to answer", to the job question, which was considered as missing value. The following table gives the number of missing values for all variables with at least one missing value.

Variable name	NA count
accentStrength.audio	132
remarkType	46
diplomeType	37
wage	21
gender	3
csp	1
csp.dad	3
csp.mum	3