

This article was commissioned by *Language and Linguistics Compass*

<http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291749-818X>

and submitted to them in May 2012. It should be published before the end of 2012.

This is the submitted draft.

Article title **The Linguistic Geography of the French of Northern France:
do we have the basic data?**

Author name **Damien Hall**

Author affiliation **University of Kent**

Author correspondence address **School of European Culture and Languages
University of Kent
CANTERBURY
Kent
CT2 7NF**

Author e-mail address **d.hall@kent.ac.uk**

The Linguistic Geography of the French of Northern France: do we have the basic data?

Abstract

Linguistic geography and dialectology in France have a long and distinguished history. It can be argued that the academic study of the geography of accents and dialects started there in the late 19th and early 20th centuries with the work of Gilliéron and Edmont. Later, such distinguished figures as Martinet, Walter, Tuailon and others continued the tradition. Most of this dialectology, though, has in common the fact that it is done by ear: a trained fieldworker interviews participants and writes down what he / she hears. Meanwhile, in the United States, Labov, Ash and Boberg have pioneered a new method of phonetic and phonological dialectology, where maps are made not from a linguist's impressions of sounds but from actual phonetic measurements of those sounds. This article describes the French dialectological tradition and these more recent techniques from the United States, then goes on to describe Towards A New Linguistic Atlas of France, a project which will apply these American phonetic and phonological techniques to the linguistic geography of the regional French varieties of Northern France.

1 What have we got to study?

1.1 Why is the Northern half of France worth studying?

Why might we want to study just the Northern half of France, not the whole country, in the first place, since the whole country now speaks French?¹ Historical dialectology splits France into two halves, North and South. For ease of reference, the areas are known by their words for 'Yes': the Northern half is the *langue d'oïl* 'language of *oïl*' area, and the South is the *langue d'oc* area, which (apart from the linguistics) is also still the name for a specific region of Southern France today. See LL-MAP's 'The Languages of France' for the basic idea, or Lexilogos' 'Carte des Langues de France' for a slightly more detailed map. The dominant variety in the *langue d'oïl* area became French, while the dominant variety in the *langue d'oc* area became Occitan. Both French and Occitan are descended from Latin, though different scholars have different views of exactly how, and the precise details are not crucial to this review. One view classifies both *langues d'oïl* and *langues d'oc* as descended from Latin into the Gallo-Iberian sub-family, and diverging after that (see Lewis, 'Gallo-Romance' for *langues d'oïl* and 'Gallo-Iberian' for *langues d'oc*). Another view sees them as diverging from (spoken)

Latin much earlier, so that *langues d'oïl* descend from spoken Latin through a Gallic Latin sub-family, while *langues d'oc* descend through an Aquitanian Latin sub-family (see Multitree; in this page, click on 'Details' in the bottom left-hand corner of the screen for further details and bibliography).

1.2 The difference between French and Anglo-Saxon sociolinguistics

Coming closer to the present, many (socio)linguistic studies and textbooks of the last forty or so years allege that, in both halves of France, there is little geographical differentiation between varieties of French. Instead, differentiation across this region is said to be largely social, so that people from Northern France are more likely to be able to pinpoint the social class of an unknown speaker, than his or her precise place of upbringing. Armstrong & Pooley give an extensive survey of sociolinguistic studies of France (mostly done by ear) which take this point of view, and indeed there are experimental results demonstrating that French people are comparatively bad at using accent alone to distinguish between regions of upbringing (Boughton, Boula de Mareüil *et al.*).

On the other hand, though France has much less geographical accent variation which is obvious to the ear than many English-speaking countries do, this does not mean that there is none at all. It is difficult to imagine a study of audible differences between accents of France on the scale of volume 2 (the British Isles volume) of Wells' monumental 1982 study. All the same, Carton *et al.* provide a fairly detailed study of accents of individuals from around France, together with recordings which are now available online, and these do show a lot of variability; however, the age of some of the speakers used in that study, and the fact that they live in rural areas, raises the suspicion that some of the accent variability they show may be due to their speaking local Romance varieties related to but different from French, and not (only) French itself. This is a frequent problem with older dialectological studies in France; there is more detail below in Section 2.

Apart from those studies where it is not actually certain what variety the speakers (thought they) were speaking, there *are* modern sociolinguistic studies which certainly look at accent features which vary geographically within France. For French in general, we have the ground-breaking work of the *Programme 'Phonologie du Français Contemporain'* ("Phonology of Contemporary French")

Programme': Durand, Laks & Lyche, <http://www.projet-pfc.net>). For areas in the North, we can cite at least:

- Reichstein, Deyhime, Léon, Peretz, Lennig, Mettas, Jamin and Berit Hansen & Juillard on Paris (Berit Hansen & Juillard is a follow-up to Peretz)
- (some contributions from) Salmon on Alsace
- Lefebvre's and Pooley's books on Lille
- Hornsby on Avion
- Hall on Normandy

These are sociolinguistic studies, however. The core subject matter of this review is the dialectology of France, to which we now turn.

2 The history of dialectology in France

France can be thought of as the cradle of today's dialectology, thanks to the *Atlas Linguistique de la France* (Gilliéron & Edmont). The *ALF* was not the first European linguistic atlas – that honour goes to Wenker's atlas of north and central German pronunciation, from the 1870s (Schmidt & Herrgen) – but that atlas was composed by postal self-report. The data are still very valuable, but self-report always has the problem that it may not be very accurate, and it cannot be checked. The *Atlas Linguistique de la France*, on the other hand, has given us the enduring image (true or not) of the dialectologist on a bicycle, going from settlement to settlement to collect people's pronunciations (Brun-Trigaud *et al.* 19-20). The result is a monumental work of 1,920 maps, covering pronunciations of words and phrases by 735 people in 639 places: a legacy which is still being exploited to the full a hundred years later. It is now followed and complemented by the long-running series of *Atlas Linguistiques de la France par régions*, a linguistic and ethnographic series which goes into more detail in each region covered by the earlier *ALF* (Séguy). For example, the *Atlas Linguistique et Ethnographique Normand* (Brasseur) has 697 informants over 114 sites, in an area about 5% the size of France – see Figure 1. The same area is covered by 32 *ALF* sites.

2.1 Methodological currents in French dialectology

With this wealth of information, a wealth which is still growing, one might think that the dialectology of France was well covered: surely there would be little room for another study? However, we come here to a major point of difference between what we might call the 'first generation' dialectological methodology, and a 'second generation' way of doing these studies. Broadly, when I say 'first generation' methodology, I mean the one followed by studies published before the Second World War, and by those published since then but constructed by the older methodology. But let me note here that this gross classification in *no way* reflects a judgement as to their relative usefulness: neither way of doing a dialectological survey is better or worse than the other: they are simply different.

2.1.1 The 'first generation'

The aim of the first generation of French dialectological studies was and is to preserve linguistic varieties before they died out. This was (and is) seen as a priority in France, a territory with many autochthonous varieties all descended from Latin, over which French had come to be imposed by decades of highly centralising language attitudes. Centralising has continued to this day, even if more attention is now being paid to the regional varieties of France than used to be the case, for example through funding for teaching of certain varieties and through the renaming and widened remit of the *Délégation Générale à la Langue Française* as the *Délégation Générale à la Langue Française et aux Langues de France* (DGLFLF). All in all, though, the attitude that French is the most advantageous language is greatly at the expense of the other varieties, most of which *are* now in danger of dying out.

As the aim of the 'first generation' of French dialectological studies was preservation, the speakers sought for these studies were naturally the people who spoke these varieties 'best' – that is, in a way that was as traditional as possible, and as little as possible 'contaminated' by French (Chambers & Trudgill 29). Note my inverted commas around 'best' and 'contaminated'! These people were often the type who have come to be known in dialectology as NORMs – Non-mobile Older Rural Males – since youth and personal mobility, in particular contact with towns, would increase the 'risk' that the person's speech would have been influenced by French. The 'male' stipulation reflects the frequent finding that older men's local varieties often seem 'purer' and 'further from French' (again, note the quotes) even than older women's, foreshadowing the later finding in Labovian sociolinguistics that

women are often (subconsciously) more likely to use prestigious linguistic variants, which in this case would usually mean French. Older women *are* occasionally sampled in this type of study, but not nearly as often as men.

2.1.2 The 'second generation'

The aim of the 'second generation' of regional linguistic studies in France is in great part sociolinguistic as well as dialectological. In talking about the aims and results of the 'first generation' of French dialectology, Section 2.1.1 does not use any words etymologically connected with 'society', as social variation was factored out of those studies – whether purposely or not. In the 'second generation', however, scholars are more interested in chronicling variation, and certainly in *French* as opposed to – or as well as – in the minority autochthonous varieties. Thus, for the first time in France, we begin to see explicit interest in linguistic variation in French, conditioned by the 'Big Four' social variables (Preston 15) – gender, age, socioeconomic status and race – as well as *geography*.

Even given this new interest, though, the consensus which is largely accepted in French sociolinguistics is that purely geographical variation in the French of Northern France is much less considerable than social variation in the French of the same area. According to this consensus, a Parisian worker speaks in much the same way as a worker from Lille – both of them probably with the *accent de banlieue*, which is more and more studied (Fagyal) – and a middle-class Parisian speaks in much the same way as a member of the middle classes from Lille. If differences between a Parisian accent and a Lille accent are pointed out, a common explanation is that much of the North is more working-class than much of the Paris area, because of the North's history of heavy industry. What may be thought to be a regional difference, therefore, can on closer inspection be found to be a social difference. What is more, French people may have clear opinions about how accents sound, in terms of how correct and pleasant they are (Kuiper), but they have been shown to perform badly, on the whole, when asked to identify the accent of speakers played to them in recordings (Boughton).

3 Modern dialect geography in France

Careful studies, then, have led to the result that regional accent differentiation in Northern France is not considerable. However, since at least the Second World War there has been a small but

persistent current of dissenting voices, from studies which have carefully controlled for many social variables, and have still arrived at results which suggested variation conditioned by geography. In what follows, I will go through the methods and results of a few of these studies, and outline a new project to add to our data.

3.1 Martinet, *La prononciation du français contemporain*

La prononciation du français contemporain is a classic study of accent variation in France. Taken prisoner by German forces during the Second World War, André Martinet made a study of the French officers in his prisoner-of-war camp, and this was the first hint that all was not as uniform in French, even in the North of France, as we are sometimes led to believe. Martinet divided France into twelve areas – the *Midi*, that is the South, and eleven North of that (Figure 2) – and he *did* find differences between the eleven ‘Northern’ areas for a number of his variables. This was in a study probably conducted almost exclusively among members of a social class whose French is now said not to vary, at least when they are not speaking in a particularly ‘familiar’ way: these were Army officers, at a time when officers in the armed forces were mostly drawn from the middle and upper classes. Martinet’s respondents were also all men. These factors mean that none of the findings of modern sociolinguistics about the interactions between gender, socioeconomic status and language (Labov, ‘The intersection of sex and social class’ and *Principles of Linguistic Change vol. 1*) can be brought to bear – and *this* means that the fact that Martinet’s results *do* fall into geographical delimitations must not be discounted. The high degree of ‘control’ (no pun intended!) of the sample means that it represents the optimum conditions for uniformity to emerge if it is present, but also the conditions under which any geographical differences detected must be thought of as credible, because they *cannot* be ascribed to social class or gender (or ethnic) differences. Of course, we must not forget that Martinet’s data was collected 70 years ago, and that it consisted of self-reported responses to questions like ‘For sentence x, which pronunciation (of three or four given) seems to you the most natural?’, or ‘If you pronounce word x and word y differently, is the difference one of vowel-sound, vowel-length, or something else?’. We know that self-report can be unreliable. Nevertheless, Martinet’s study *did* reveal geographical differences within the North of France, and they must be taken seriously.

3.2 Henriette Walter

In the thirty years which followed the publication of *La prononciation du français contemporain*, there were comparatively few studies of the French of France on a national scale. Honourable mentions are due to scholars like Gaston Tuillon, Georges Straka, Gérard Taverdet and Léon Warnant. But the work on which I wish to stop next in detail is that of Henriette Walter, in particular *Enquête Phonologique et Variétés Régionales du Français*, for which most of the data were collected in the 1970s. She had preceded it with *Dictionnaire de la prononciation française dans son usage reel* (with André Martinet) and *La dynamique des phonèmes*, but these books were criticised for their samples of speakers, which were limited both in terms of numbers – *Dictionnaire de la prononciation française...* had 17 – and in terms of geography: the same dictionary sampled only Parisians. These facts do not in any way limit the value of the studies, especially as one of the points was to demonstrate that Paris is a melting-pot where all the speech is *not* uniform (*Enquête phonologique* 10). However, though they may have sampled people whose *roots* were from all over France, they did not sample people who *lived* and *had grown up* all over France.

Enquête phonologique fills that gap, dividing Francophone Europe into 35 regions – 28 of them in France – covered by 111 interviewees, 99 of them in France (Figure 3). For the first time, a study of the whole European French-speaking population was conducted, which aimed to speak to people brought up in the areas it covered. However, the average age of the informants was still rather high, and many of them were still rural – so the result was still possibly backward-looking rather than forward-looking. However, it was much more sociolinguistically informed than previous studies of the whole of France, at least, had been.

4 The next steps: instrumental dialectology

If these studies, particularly Martinet's *La prononciation du français* and Walter's *Enquête phonologique*, did not provide the 'basic data' about geographical variation across France, then, what would that basic data look like? The data we already have, some of which I have described, is an essential part of the picture, but an important step which we have not yet taken is to get comparable acoustic data from across the country, and to analyse it phonetically. The instrumental evidence – vowels measured in terms of their component frequencies, so that they can be described in a

scientifically precise way – is an important complement to what we already have, because of a crucial fact about geographical changes in accents which is convincingly revealed by work in the Labovian tradition of sociolinguistics. In that tradition, as much recording and instrumental analysis as possible is done (at least where we are talking about accents, which can be described by measuring vowels). The fact that this analysis reveals is that *there are differences which are possibly too small to be heard at a conscious level, and yet are consistent throughout communities and move in an expected direction*. This kind of difference is best illustrated by looking at work on the sociolinguistics of American English, since the techniques used to analyse it were first developed there and are most advanced there; the ‘Towards A New Linguistic Atlas of French’ project, which will be described in the last half of this article, aims to apply these kinds of techniques to the cities of Northern France.

4.1 The Atlas of North American English

Since the end of the 1960s, the US cities around the Great Lakes of North America have been known to be participating in a series of sound-changes known as the Northern Cities Vowel Shift (Figure 4 – see also Gordon). The shift involves a reorientation of many of the vowels of North American English – Penny Eckert’s website has a chart and sound-files (go to the bottom of the webpage at the reference).

In the *Atlas of North American English* (ANAE; Labov *et al.*), one of the variables used to categorise speakers is the front-back distance in Hertz between the vowels /*ɛ*/ and /*a*/. (Labov’s work uses the symbols (*e*) and (*o*) respectively for these vowels; a minimal pair which is often used for them is *Ked* and *cod*.) In non-Northern Cities Shift speech, *Ked* is [kɛd], fronter than *cod* [kad]; however, the more of a Northern Cities accent a speaker has, the closer together these vowels move. In the most advanced speakers *Ked* is pronounced [kʌd], a central mid-vowel, and *cod* is pronounced [kæd], which is at least central, if not actually fronter than the vowel in *Ked*. In these speakers, the normal front-back relationship between these two vowels is reversed. The difference between the vowels of *Ked* and *cod* can therefore be used as a diagnostic of how advanced a speaker is in the Northern Cities Shift. This measure is a real measure of the degree to which a speaker has the Northern Cities Shift, inasmuch as its geographical spread corresponds well with the geographical spread of other Northern Cities Shift features (*cf* Figure 5, adapted from ANAE Map 14.9, Labov *et al.*: 204).² And yet,

when asked about the features of their local accent, no Northern Cities speaker points out that they have the vowels of *Ked* and *cod* closer together than other speakers, and they certainly would not be able to quantify the difference.

4.2 The Regional French of Normandy

In work on Normandy, Hall has found a similarly surprising, subtle difference between the vowels of *pattes* ('animal's paw') and *pâtes* ('pasta'). In the most conservative varieties of European French, and in Canadian French generally, these words (which mean 'animal's paw' and 'pasta' respectively) are pronounced [pat] and [pɑt], but many people pronounce them the same, both [pat]. The difference can be heard in the (Canadian) recordings in Germain-Rutherford's website. It is instructive that apparently no website aiming to teach the sounds of European French now includes the difference between /a/ and /ɑ/, but it can be found in some of the recordings of regional French at the website for Carton *et al.*

In Hall's sample from Darnétal, in the Rouen urban area, speakers younger than 20 or over 69 years old kept *pattes* and *pâtes* separate, but *raised them both towards the centre of the vowel-space* (Figure 6). In the perceptual part of the study (where Darnétal speakers were asked to describe features of their local accent), they had ideas about some features, but no-one mentioned or imitated the raising of *pattes* and *pâtes*. And yet, they must have been able to perceive it at some level, because they were able to reproduce it (unconsciously, of course) in their conversational speech: the finding of raising the *pattes* and *pâtes* vowels was consistent within those age-groups, the youngest and the oldest in the study.

The point of describing these American and French examples is to underline the *importance of phonetic measurement*, alongside other ways of looking at accents: without phonetic measurement, it is likely that neither of these two consistent and measurable accentual features would have been discovered – or at least, they would not have been discovered and quantified as easily.

5 Towards A New Linguistic Atlas of France

This brings us, then, to the study which began in March 2011, 'Towards A New Linguistic Atlas of France'. It has a methodology between those of the 'first' and 'second' generation of French dialectology, which can be thought of as 'sociolinguistically-informed dialectology'. It is certainly dialectology, in that it aims to look at variation in Northern French in its geographical aspect. It is not really fully-fledged sociolinguistics, because social variation is not represented in the sample: it only takes account of social factors in that it controls for them. Most sites in the project are represented by two or four speakers, who will be as socially homogeneous as possible.

5.1 The TANLAF sample

TANLAF will cover the 34 largest towns and cities in the North of France, in a crescent from Brittany to Alsace *via* Paris (Figure 7 and Table 1; see also http://bit.ly/TANLAF_FieldworkProgress for the latest interview progress). The sample size in each urban area is determined by population according to INSEE's 'Recensement de la population 2008' (published in 2011). All the sample areas have populations of 100,000 or more in their *aire urbaine* (roughly zones of economic attraction – see INSEE, 'Aire urbaine'). For urban areas with populations between 100,000 and 200,000, one man and one woman will be sampled; for urban areas with populations larger than 200,000, two men and two women will be sampled. The only exception will be Paris, as it is so much larger than all other French urban areas: in Paris, up to sixteen speakers will be sampled – that is, two men and two women from each of four sites around the area.

Speakers will be aged between 18 and 30, or their early 30s, will have been brought up in the place they represent in the sample, and will be as far as possible what British people and Americans might call 'middle-class'. This criterion is set in order to try to avoid both the *accent de banlieue* on the one hand, and hyper-educated hyper-standardised speech, on the other. TANLAF will therefore also avoid potential speakers who are educators, and most people who have more than a first degree – because these people may well have an impulse (perhaps conscious) to speak in what they consider a standard manner, and so to avoid a local accent even if one does exist in their area. In order not to overstate accent differences which are directly due to the local autochthonous varieties, people who actually speak more than a few words of them will be excluded; of course, avoiding these people is not difficult among the young. TANLAF will investigate the possibility that any regional differences

detected could be because of the local autochthonous variety in the substrate, but no particular speaker should be (unrepresentatively) different from most speakers in their town simply because they personally speak the variety.

TANLAF's sampling methodology therefore departs in at least two major ways from the strong tradition of French dialectology until now – or at least that of the 'first generation' of studies. That tradition concentrates on NORMs – Non-mobile Older Rural Males, though they are not all in fact male – whereas TANLAF will sample non-mobile *younger* people of both sexes, and they will also be *urban*. Again, this is partially due to a methodological innovation of Labovian sociolinguistics – something which is actually present in much work in that tradition since the 1960s, including the *Atlas of North American English*. Not all the speakers sampled in the ANAE are young, but all are non-mobile and urban, because the repeated finding of work in that tradition is that innovations in language are highly likely to begin in towns and spread to their hinterland, through the increased opportunity that towns give for people of different accents to speak to one another. Furthermore, it is most often among young people that language change begins and spreads. TANLAF will therefore test the extent to which processes of linguistic change which have been found to obtain over and over in the Anglophone world, are applicable in France.

5.2 The TANLAF dataset

The analysis will be conducted on a uniform dataset collected according to an extended version of the protocol for the Phonology of Contemporary French project (Durand, Laks & Lyche; <http://www.projet-pfc.net> – referred to as 'PFC', from the word-order of its name in French). This means a passage of text to read, a word-list to read, and two types of conversation, 'guided', where specific demographic questions are asked, and 'free', which, as the name implies, can be about anything. In TANLAF, the PFC project reading-passage is not changed at all; the PFC word-list is also retained, but it is supplemented by a large number of other words, because TANLAF will include at least preliminary investigations on a wider range of phonemes than those concentrated on by PFC. This approach was chosen not only because it avoided the need to reinvent the wheel on which PFC is already turning quite nicely, but also because it will allow at least parts of the TANLAF data to be added to the valuable PFC database.

Table 2 gives details of the phonemes which TANLAF will investigate. Other variables are also studied in the PFC project, and those will not form part of the TANLAF analysis – but, for the record, they are:

- pronunciation or omission of schwa (the unstressed vowel in words like the word for ‘table’, which can be pronounced either /tabl/ or /ta.blə/); and
- liaison (the contexts in which speakers either pronounce or do not pronounce word-final consonants followed by a vowel, as in the sentence *Les soldats arrivent* (‘The soldiers arrive’), which can either be pronounced [le sɔl da a ʁiv] (as is most common) or [le sɔl da za ʁiv] (in careful language). (This is just one convenient example: other liaison contexts are obligatory, and still other potential occasions for liaison are almost never realised.)

To summarise, then, TANLAF will examine urban Northern French data using the phonetic and phonological methodology of the *Atlas of North American English*; when this stage is completed, the data will be used for a dialectological survey, again similar to that of *ANAE*. In *ANAE*, this categorisation and mapping resulted in the division of North America into a small number of gross dialect areas. Studies of potential and possible geographical variation in France (or at least in the North of it) often include, towards the end, a telling phrase such as ‘small differences may be found’ between geographical varieties in France; this has the implication (often also stated explicitly) that differences according to other social variables like class are more telling. This is probably true for many French linguistic variables; but *ANAE* and the studies leading up to it show that small differences, not previously thought to vary systematically, are in fact worth mapping, and TANLAF will investigate whether the same is true for Northern France. Many French people who have already been interviewed for TANLAF have expressed surprise at the idea that geographical accent differences in Northern France are believed to be small – though of course much work has shown that the differences people *believe* they perceive and the differences they can actually use reliably are not the same thing.

5.3 What do we expect to get out of TANLAF?

What, then, is TANLAF expected to produce? In brief, a uniform, comparable set of phonetic and phonological data from the largest urban areas of the Northern third of France. To summarise, what TANLAF sets out to do is sociolinguistically-informed dialectology of France. It isn't sociolinguistics in the full sense, because many of the loci of sociolinguistic variation are purposely being held constant. Specifically, only geography and gender are allowed to vary, while socioeconomic status, age and probably race are being held more-or-less constant. So the interviewees will be men and women, mostly between 18 and 30 years old, with ideally more than the legal minimum education (though not *too much more*), from each of the 34 largest urban areas in the Northern third of France. In this way, subtle phonetic differences between their accents will be isolated – some, by all means, easily audible for any speaker, but transitional zones are also expected, as well as consistent but subtle differences, which would not be so easily isolated, even by a *trained* ear, in the absence of phonetic measurements. A number of studies have found that such subtle differences do exist, and new ones may be emerging. The study will be limited to well-chosen middle-class young people, the social group among whom any changes would be likely to emerge first. This is where the 'sociolinguistically-informed' part comes in. It is expected that this approach will give us a more detailed picture than we yet have of geographical variation in French within the area studied.

It is important to note that this will not be the final word, of course. A point which is not sufficiently appreciated about the *Atlas of North American English* is that it is only intended to provide a basic picture on which other scholars are invited to elaborate. Others can fill in the holes between the urban areas that the *ANAE* covers, cover those urban areas in more detail, and so on. The authors of the atlas state this in its introduction (Labov *et al.* 3, 9). The same will be true of the data-set and analysis that come out of TANLAF: people will be welcome to elaborate on its data, and question and contradict it. Also, TANLAF does not include a perceptual component, which would be the best way of finding out 'which differences are big enough to make a difference' for native speakers. From the data collected here, such a study will certainly be possible in future, but for the moment TANLAF aims to collect a basic, partially uniform data-set on which such experiments and further conclusions can be based.

6 Conclusion





We return to the question asked at the beginning. Do we have the basic data about geographical variation in the North of France? We hope to have demonstrated here that, through the past century's worth of dialectological and sociolinguistic studies in France, we have *some* of the basic data, or – better – admirable basic data about some areas of variation. We know a lot about the language varieties which survive in France from before French was imposed, and we know quite a lot, too, about audible regional variation in French itself, though there have been few studies which covered more than one region, and especially not in the last generation. We do not yet, however, have a collation of basic data which will show us how French itself varies across France *now*, and how it may be changing. The TANLAF project on the Northern third of France is the first step in getting that particular sort of basic data.

¹ This research is funded by a Leverhulme Trust Early Career Fellowship to the author (ECF/2010/0556, held from 2011 to 2013; <http://www.leverhulme.ac.uk/funding/ECF/ECF.cfm>). I would like to thank Leverhulme for their support, as well as my audiences at various conferences and workshops. An initial version of this work was presented at the final workshop of the Research Network 'Language and Social Structure in Urban France' (Cambridge, September 2011, funded by the (UK) Arts and Humanities Research Council; <http://sociolinguistique-urbaine.com/spip.php?article130>). Much helpful feedback was also received from the audiences at the Romance Linguistics Seminar XXXIX (<http://people.pwf.cam.ac.uk/cjp16/rls/rls39/prog39.htm>), the 25th CerLiCO conference (www.risc.cnrs.fr/pdf/programme_cerlico_2011.pdf), Methods in Dialectology 14 (<http://westernlinguistics.ca/methods14/>), New Ways of Analyzing Variation 40 (<http://nwav40.georgetown.edu/>) and Journées PFC Paris 2011 (<http://www.projet-pfc.net/paris2011.html>). Any errors are of course my own.

² **Note to editors and copy-editors (not to be reproduced in the finished article!)** The key to Figure 5 is in the table on the next page of this .doc document. The limitations of MS Word prevent me from inserting the map symbols into an endnote in my manuscript, as I wished to, but I hope it will be possible to put the table in an endnote in the final product online and in print.

2 **Text of Endnote 2**

3 The key to Figure 5 is:

Isoglosses	Explanation	Notes
	/a/ (LOT, in the word-class system of Wells' <i>Accents of English</i>) has been fronted so that it is fronter than /ʌ/ (STRUT), whereas in non-Northern Cities Shift English LOT is backer than STRUT. This means that they are pronounced [ʌ] and [æ] (i.e. STRUT has a quality similar to non-NCS English, and LOT is pronounced 'lat').	One of the most widespread dialect features of the North of the US
	/ɛ/ (DRESS) and /a/ (LOT) are close together on the front-back axis, so they are pronounced [ʌ] and [a]	The element of the Northern Cities Shift used as an example here
	/æ/ (TRAP) is higher and fronter than /ɛ/ (DRESS), so they are pronounced [ɛ] and [ʌ] ('trep' and 'druss')	The most 'iconic' features of the Northern Cities Shift (particularly trap > 'trep')
	<ul style="list-style-type: none"> • /ow/ (GOAT) is not fronted (as it is in many areas South of here) • /a/ (LOT) and /ɔ:/ (THOUGHT) are not merged (as they are in much of the US) • the pronunciation of /æ/ (TRAP) is not determined by the consonant following it (as it is in some Eastern Seaboard areas) 	'The North': the widest definition of the Northern US dialect area, defined by the other dialect areas at its boundaries

4

5

6 Works cited

- 7 Berit Hansen, Anita & Caroline Juillard. 2011. La phonologie parisienne à trente ans d'intervalle – Les
8 voyelles à double timbre. *Journal of French Language Studies* 21(3). 313-360.
- 9 Boughton, Zoë. 2006. When perception isn't reality: Accent identification and perceptual dialectology
10 in French. *Journal of French Language Studies* 16(3). 277-304.
- 11 Boula de Mareuil, Philippe, Bianca Vieru-Dimulescu, Cécile Woehrling & Martine Adda-Decker. 2008.
12 Accents étrangers et régionaux en français: Caractérisation et évaluation. *Traitement*
13 *Automatique des Langues* 49(3). 135–162. <http://perso.limsi.fr/Individu/mareuil/publi.html>. (10
14 May, 2012.)
- 15 Brasseur, Patrice. 1980, 1984, 1997, 2011. *Atlas Linguistique et Ethnographique Normand* (4 vols).
16 Paris, France: Centre National de la Recherche Scientifique.
- 17 Brun-Trigaud, Guylaine, Yves Le Berre & Jean Le Dù. 2005. *Lectures de l'Atlas linguistique de la*
18 *France de Gilliéron et Edmont: Du temps dans l'espace*. Paris: Éditions du Comité des travaux
19 historiques et scientifiques.
- 20 Carton, Fernand, Mario Rossi, Denis Autesserre & Pierre Léon. 1983. *Les Accents des Français*.
21 Paris: Hachette. <http://accentsdefrance.free.fr>. (2 May, 2012.)
- 22 Chambers, Jack & Peter Trudgill. *Dialectology*. Cambridge: Cambridge University Press. 2nd edn.
- 23 Délégation Générale à la Langue Française et aux Langues de France. n.d. Délégation Générale à la
24 Langue Française et aux Langues de France. <http://www.dglf.culture.gouv.fr>. (10 May, 2012.)
- 25 Deyhime, Guiti. 1967a. Enquête sur la phonologie du français contemporain [first part]. *La*
26 *Linguistique* 1: 97-108.
- 27 Deyhime, Guiti. 1967b. Enquête sur la phonologie du français contemporain [second part]. *La*
28 *Linguistique* 2: 57-84.
- 29 Durand, Jacques, Bernard Laks & Chantal Lyche. 2009. *Phonologie, Variation et Accents du*
30 *Français*. Paris: Hermès.
- 31 Eckert, Penelope. 2012. Vowel shifts in Northern California and the Detroit suburbs.
32 <http://www.stanford.edu/~eckert/vowels.html>. (14 May, 2012.)
- 33 Fagyal, Zsuzsanna. 2010. *Accents de banlieue: aspects prosodiques du français populaire en contact*
34 *avec les langues de l'immigration*. Paris: L'Harmattan.

35 Germain-Rutherford, Aline. 2005. Tableau comparatif des voyelles françaises et américaines.
36 http://courseweb.edteched.uottawa.ca/Phonetique/pages/phonetique/tableau_comp_voy.htm.
37 (14 May, 2012.)

38 Gilliéron, Jules & Edmond Edmont. 1902-10. *Atlas Linguistique de la France*. Paris: Champion.

39 Gordon, Matthew J. 2011. Methodological and theoretical issues in the study of chain shifting.
40 *Language and Linguistics Compass* 5(11). 784-794.
41 <http://onlinelibrary.wiley.com/doi/10.1111/j.1749-818X.2011.00310.x/full>. (11 May, 2012.)

42 Hall, Damien. 2008. *A sociolinguistic study of the regional French of Normandy*. Philadelphia, PA:
43 University of Pennsylvania dissertation.

44 Hornsby, David. 2006. *Redefining regional French: Koinéization and dialect levelling in northern*
45 *France* (Studies in Linguistics 3). Oxford: Legenda.

46 Institut National de la Statistique et des Études Économiques [INSEE]. 2011. *Résultats du*
47 *Recensement de la Population – 2008*. Online database, accessible through
48 <http://www.recensement.insee.fr/home.action>. (23 May, 2012.)

49 Institut National de la Statistique et des Études Économiques [INSEE]. 2012. Aire urbaine.
50 <http://www.insee.fr/fr/methodes/default.asp?page=definitions/aire-urbaine.htm>. (14 May, 2012;
51 click on 'English' in the top right corner for a translation.)

52 Jamin, Mikaël. 2005. *Sociolinguistic Variation in the Paris Suburbs*. Canterbury: University of Kent
53 dissertation.

54 Kuiper, Lawrence. 1999. Variation and the Norm: Parisian Perceptions of Regional French. In Dennis
55 Preston, *Handbook of perceptual dialectology: Volume 1*, 243-262. Amsterdam: John
56 Benjamins Publishing Company.

57 Labov, William. 1990. The intersection of sex and social class in the course of linguistic change.
58 *Language Variation and Change* 2(2). 205–254.

59 Labov, William. 1994. *Principles of linguistic change, volume 1: Internal factors*. Malden, MA:
60 Blackwell.

61 Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English:*
62 *Phonetics, Phonology and Sound Change: a Multimedia Reference Tool*. Berlin: Mouton de
63 Gruyter.

- 64 Language and Location – Map Accessibility Project [LL-MAP]. 2012. The Languages of France.
 65 <http://www.lmap.org/viewer.html?maps=c:696547>. (1 May, 2012.)
- 66 Lefebvre, Anne. 1991. *Le français de la région illoise* (Série Homme et Société 18). Paris:
 67 Publications de la Sorbonne.
- 68 Lennig, Matthew. 1978. *Acoustic measurement of linguistic change: The modern Paris vowel system*.
 69 Philadelphia, PA: University of Pennsylvania dissertation (Pennsylvania Dissertation Series
 70 no.1).
- 71 Léon, Pierre. 1972. Étude de la prononciation du “e” accentué chez un groupe de jeunes Parisiens. In
 72 Albert Valdman (ed.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*.
 73 The Hague: Mouton.
- 74 Lewis, M. Paul (ed.). 2009. *Ethnologue: languages of the world*. Dallas, TX: SIL International. 16th
 75 edn. Online version: <http://www.ethnologue.com/>.
- 76 Lewis, M. Paul (ed.). 2009a. Language family trees: Indo-European [...] Gallo-Romance.
 77 http://www.ethnologue.com/show_family.asp?subid=301-16. (2 May, 2012.) In Lewis, M. Paul
 78 (ed.). 2009.
- 79 Lewis, M. Paul (ed.). 2009b. Language family trees: Indo-European [...] Gallo-Iberian.
 80 http://www.ethnologue.com/show_family.asp?subid=300-16. (2 May, 2012.) In Lewis, M. Paul
 81 (ed.). 2009.
- 82 Lexilogos. 2002-2012. Carte des langues de France.
 83 http://www.lexilogos.com/france_carte_dialectes.htm. (2 May, 2012.)
- 84 Martinet, André. 1945 / 1971. *La Prononciation du Français Contemporain: témoignages recueillis en*
 85 *1941 dans un camp d'officiers prisonniers*. Paris: Librairie E. Droz.
- 86 Martinet, André & Henriette Walter. 1973. *Dictionnaire de la Prononciation Française dans son Usage*
 87 *Réel*. Paris: France Expansion.
- 88 Mettas, Odette. 1979. *La prononciation parisienne: Aspects phoniques d'un sociolecte parisien (du*
 89 *Faubourg Saint-Germain à La Muette)*. Paris: Société d'Études Linguistiques et
 90 Anthropologiques de France.
- 91 Multitree. 2009. Indo-European: Composite. [http://multitree.linguistlist.org/trees/Indo-](http://multitree.linguistlist.org/trees/Indo-European%3A%20Composite)
 92 [European%3A%20Composite](http://multitree.linguistlist.org/trees/Indo-European%3A%20Composite). In Multitree: A digital library of language relationships. Ypsilanti,

93 MI: Institute for Language Information and Technology (LINGUIST List), Eastern Michigan
94 University. <http://multitree.org/>. (2 May, 2012.)

95 Peretz, Caroline. 1977. Aspects socio-linguistiques du parler parisien contemporain. In Henriette
96 Walter (ed.), *Phonologie et Société*. Montreal: Didier.

97 Pooley, Timothy. 1996. *Chtimi: The urban vernaculars of northern France*. Clevedon: Multilingual
98 Matters Ltd.

99 Pooley, Timothy. 2004. *Language dialect and identity in Lille* (Studies in Linguistics and Semiotics
100 Vol. 19A). Lewiston, NY: The Edwin Mellen Press.

101 Preston, Dennis R. 1986. Fifty some-odd categories of language variation. *International Journal of the
102 Sociology of Language* 57. 9-47.

103 Reichstein, Ruth. 1960. Études des variations sociales et géographiques des faits linguistiques. *Word*
104 16(1). 55-99.

105 Salmon, Gilbert-Lucien (ed.). 1985. *Le français en Alsace* (Bulletin de la Faculté des Lettres de
106 Mulhouse XIV). Paris: Champion.

107 Schmidt, Jürgen Erich & Joachim Herrgen (eds.). 2001-9. *Digitaler Wenker-Atlas*. Marburg:
108 Forschungszentrum Deutscher Sprachatlas. <http://www.diwa.info/>. (10 May, 2012.)

109 Séguy, Jean. 1973. Les Atlas Linguistiques de la France par régions. *Langue Française* 18. 65-90.
110 doi: 10.3406/lfr.1973.5631. [http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-
111 8368_1973_num_18_1_5631](http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_1973_num_18_1_5631). (10 May, 2012.)

112 Walter, Henriette. 1976. *La dynamique des phonèmes dans le lexique français contemporain*. Paris:
113 France Expansion.

114 Walter, Henriette. 1982. *Enquête Phonologique et Variétés Régionales du Français*. Paris: Presses
115 Universitaires de France.

116 Wells, John. 1982. *Accents of English*. Cambridge: Cambridge University Press.

117

118

Tables119 **Table 1: Urban areas to be covered in TANLAF**

Urban area	Population	Urban area	Population
Amiens	292,047	Lorient	212,317
Arras	128,558	Maubeuge	131,760
Beauvais	123,680	Metz	389,603
Béthune	365,513	Mulhouse	281,520
Boulogne-sur-Mer	134,244	Nancy	434,202
Brest	311,735	Paris	12,089,098
Caen	396,959	Quimper	123,963
Calais	126,618	Reims	313,818
Charleville-Mézières	106,747	Rennes	654,478
Cherbourg-Octeville	118,673	Rouen	649,291
Colmar	126,302	Saint-Brieuc	166,040
Creil	117,825	Saint-Quentin	111,595
Douai-Lens	544,143	Strasbourg	757,609
Dunkerque	259,566	Thionville	134,012
Évreux	110,528	Troyes	188,320
Le Havre	293,851	Valenciennes	366,781
Lille	1,150,530	Vannes	144,353

120

121 *Speakers will be aged between 18 and 30 ...*

122

123 **Table 2: TANLAF variables**

124 **PFC variables also in TANLAF**

‘a family’	Whether or not an informant makes a contrast between /a/ (as in <i>pattes</i> ‘paws’) and /ɑ/ (as in <i>pâtes</i> ‘pasta’)
‘é / è family’	Whether or not an informant makes a contrast in stressed final position between /e/ (as in <i>fée</i> ‘fairy’) and /ɛ/ (as in <i>fais</i> ‘do(es)’)
‘eu family’	Whether or not an informant makes a contrast between words like <i>jeune</i> /ʒœn/ ‘young’ and <i>jeûne</i> /ʒøn/ ‘fasting’
‘o family’	Whether or not an informant makes a contrast between words like <i>côte</i> /kot/ ‘coast’ and <i>cote</i> /kot/ ‘(popularity / financial) rating’
nasal vowels	How many contrasts are made between the vowels in <i>blond</i> ‘blond’, <i>blanc</i> ‘white’, <i>brin</i> ‘sprig’, and <i>brun</i> ‘brown’ (for some speakers all four are different; some pronounce <i>brin</i> and <i>brun</i> the same; and some pronounce <i>blond</i> and <i>blanc</i> similarly or the same)
glides	Whether or not an informant makes a contrast between all three glides, as exemplified in the standard <i>huer</i> /ʁe/ ‘shout out’, <i>yé-yé</i> /jeje/ (French 1960s rock and roll, and its singers and fans) and <i>ouais</i> /wɛ/ ‘yeah’

125

126 **Non-PFC variables in TANLAF**

127 **Other stressed vowels of French**

128	r	Pronunciation of (word-final) /ʁ/
129	k	Amount of aspiration of (particularly syllable-initial) /k/

130

131 *Other variables are also studied ...*

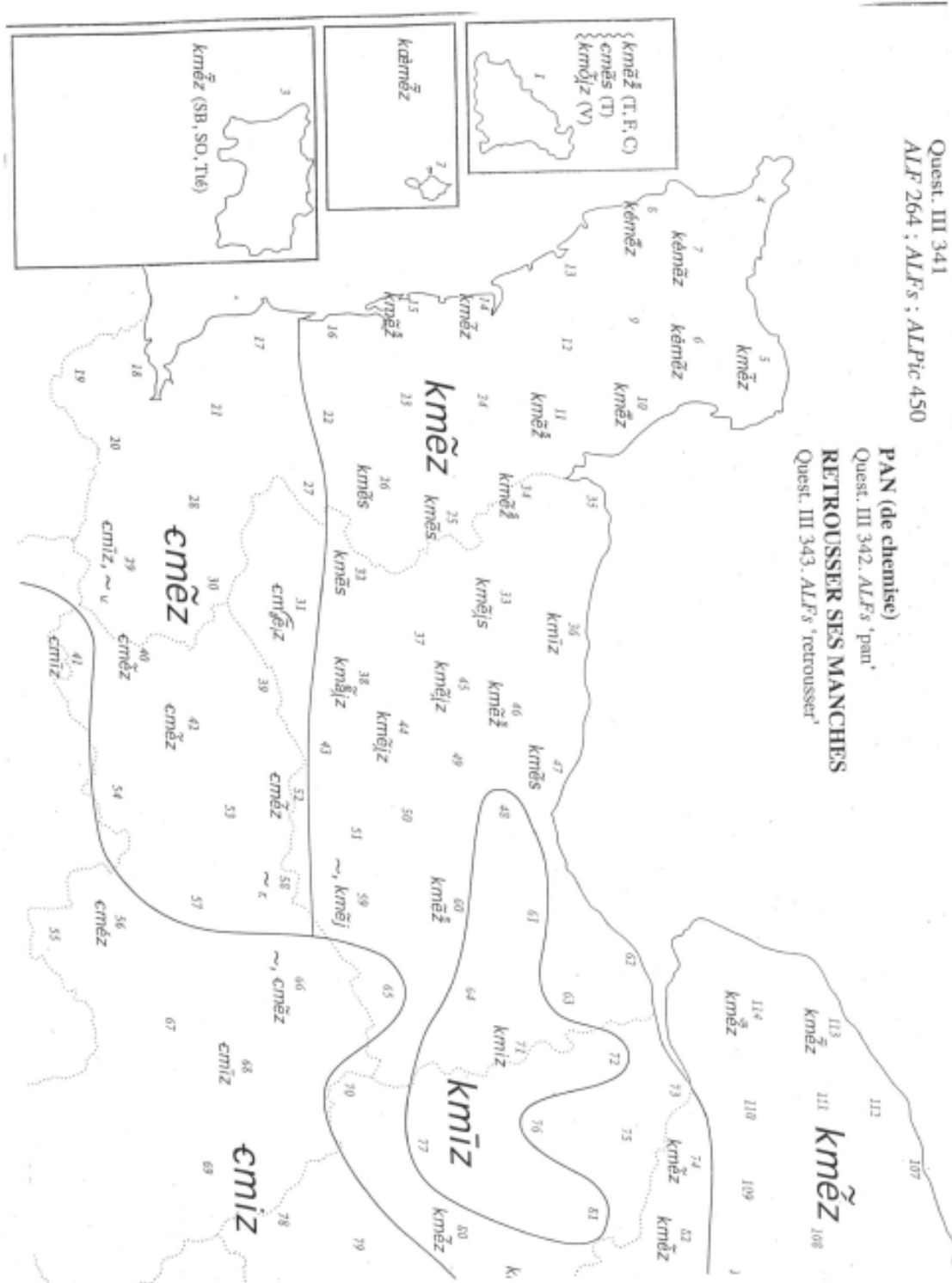
132

133

Figures

134 Figure 1: Map 1099 from Brasseur's *Atlas Linguistique et Ethnographique Normand* (chemise
135 in Normandy)

136

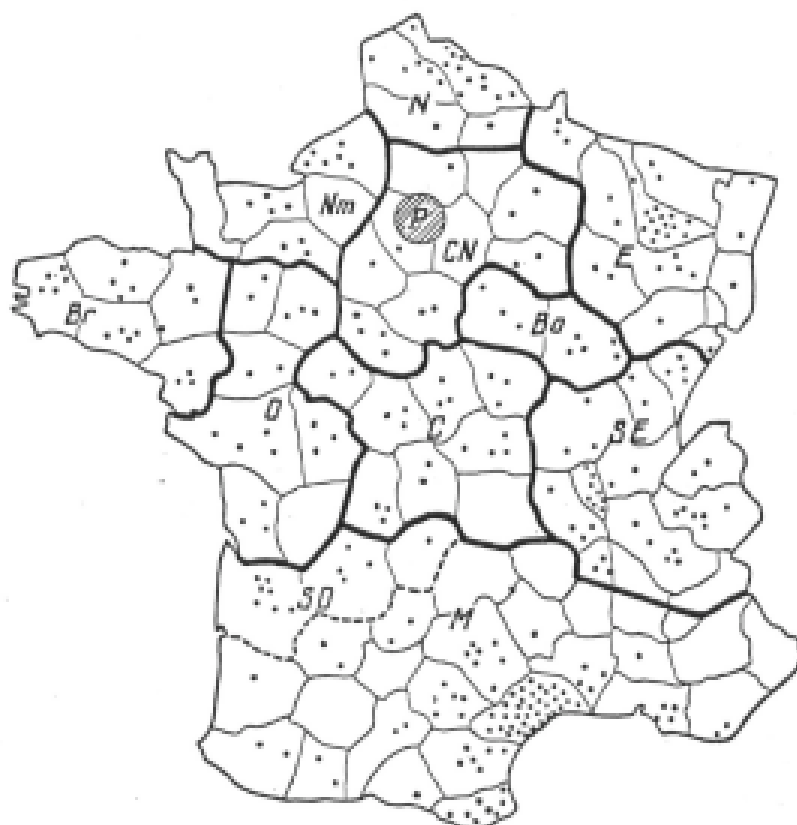


137

138 The same area is covered by 32 ALF sites.

139

140 **Figure 2: Location of subjects for Martinet's *La prononciation du français* (p.30)**



141

142 *This was in a study probably conducted ...*

143

144 Figure 3: Location of subjects for Walter's *Enquête phonologique* (p.253)

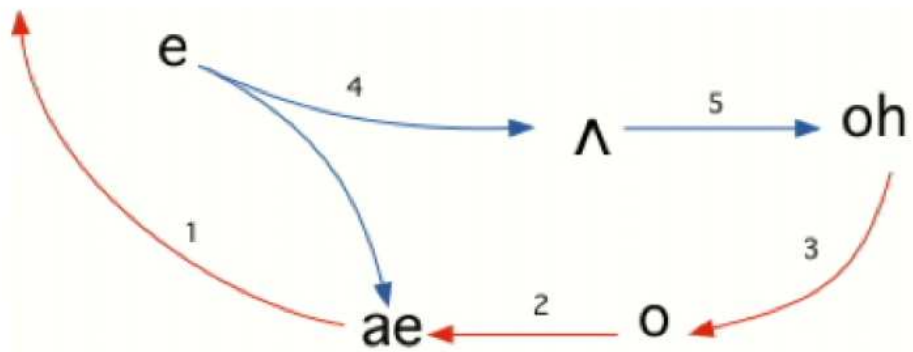


145

146 **4** *The next steps: instrumental dialectology*

147

148 **Figure 4: The phonology of the Northern Cities Shift**



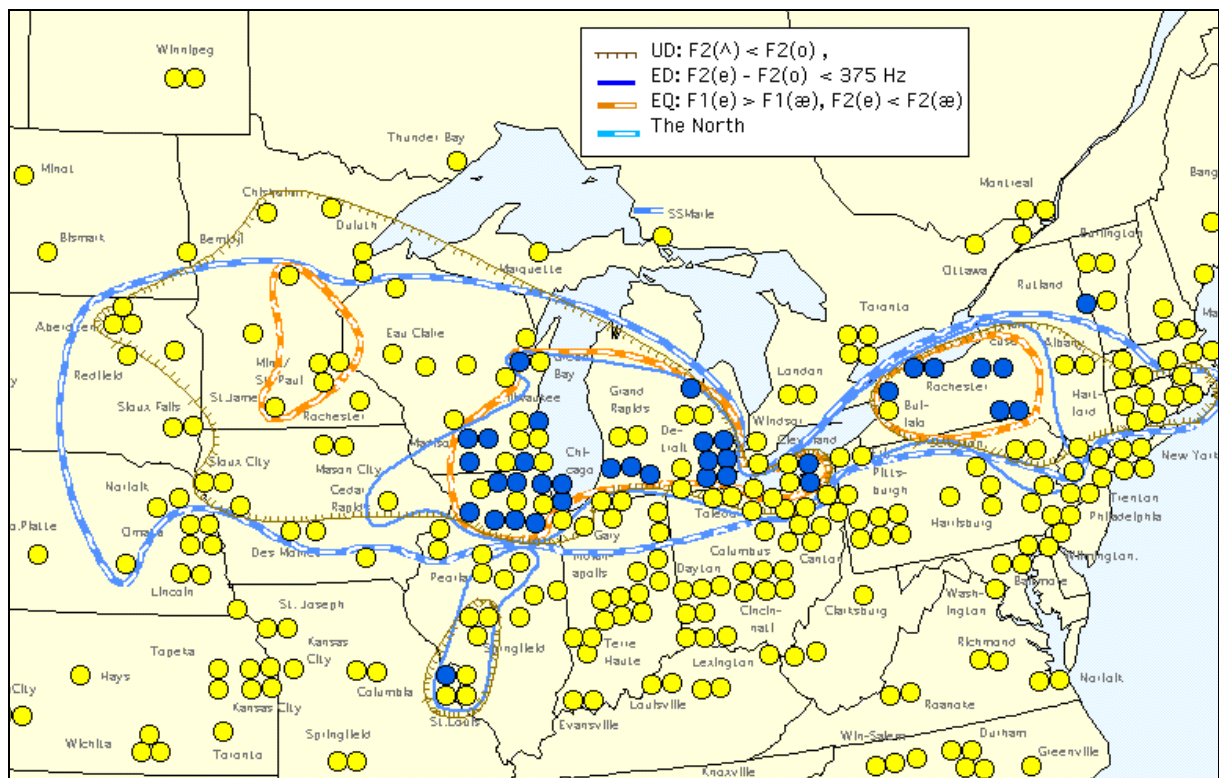
149

150 *The shift involves a reorientation of many of the vowels of North American English ...*

151

152

153 **Figure 5: The geographical spread of the Northern Cities Shift**



154

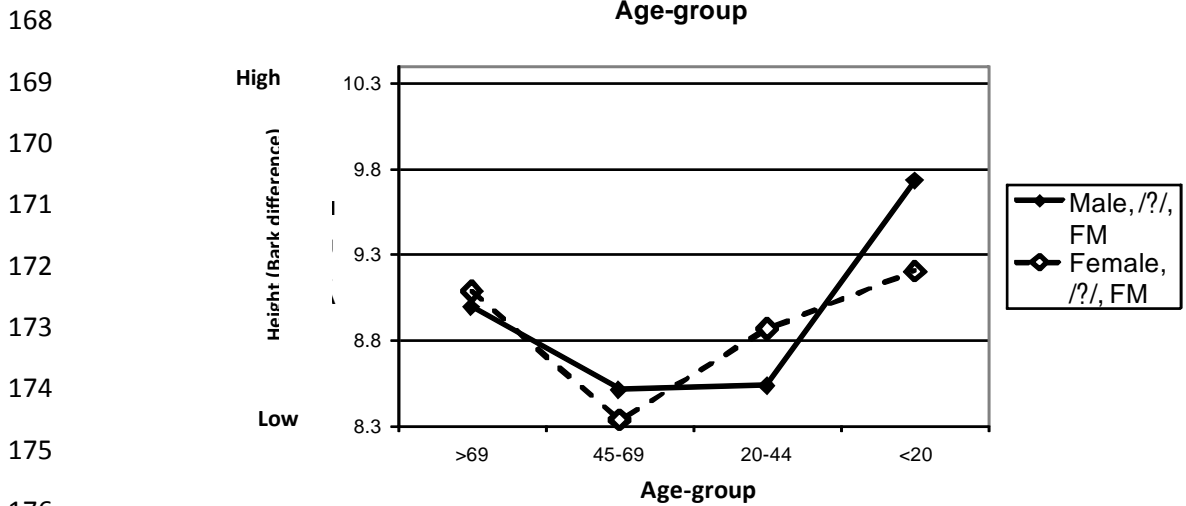
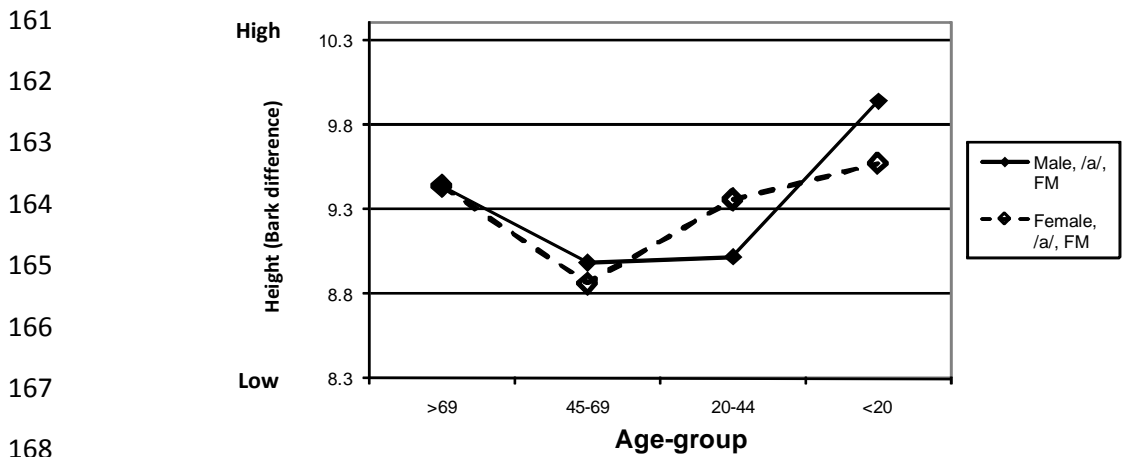
155 **4.2 The Regional French of Normandy**

156

157 **Figure 6: Centralisation of /a/ and /ɑ/ in Darnétal, near Rouen (Upper Normandy) – adapted**
 158 **from Figure 3-10 of Hall**

159 **Top: height of /a/ in formal linguistic tasks**

160 **Bottom: height of /ɑ/ in formal linguistic tasks**

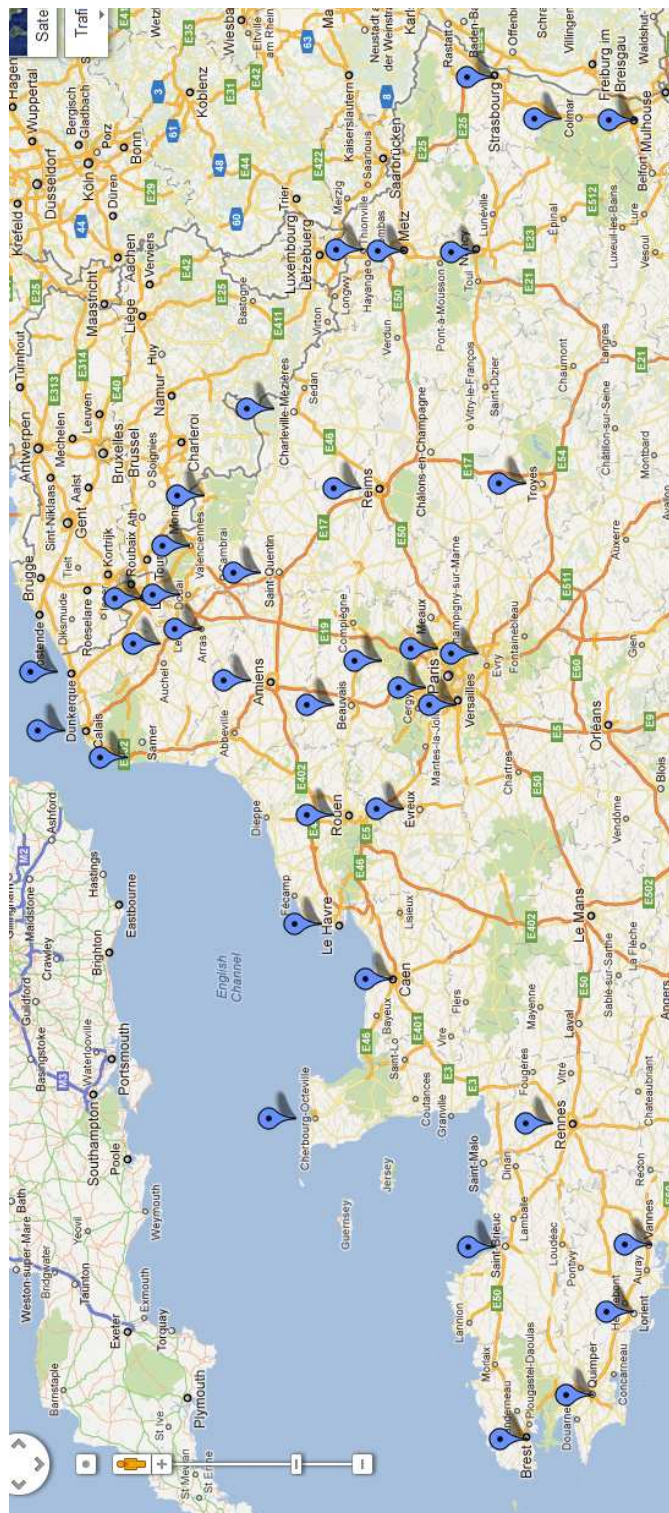


177 *The point of describing these ...*

178

179 **Figure 7: Urban areas to be covered in TANLAF (© Google Maps 2012)**

180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206



Speakers will be aged between 18 and 30 ...