

PROTOCOL AND DATA COLLECTION

Table of contents

0. Introduction and summary
1. Technical equipment
2. Procedure
 - 2.1 Wordlists
 - 2.2 Text
 - 2.3 Formal conversation
 - 2.4 Informal conversation
3. Working in the field
4. Further remarks on the selection of informants
5. Readings
6. Documents
 - 6.1 Information sheet
 - 6.2 Notes on the information sheet and the formal interview
 - 6.3 Wordlist 1
 - 6.4 Wordlist 2
 - 6.5 Text
7. Ethical questions
8. Bibliographical references

0. Introduction and summary

PAC (La Phonologie de l'Anglais Contemporain: usages, variétés et structure / The Phonology of Contemporary English: usages, varieties and structure) is a project coordinated by Philip Carr and Jacques Durand. Among other things, we aim at (1) giving a better picture of spoken English in its unity and diversity (geographical, social and stylistic), (2) testing phonological and phonetic models from a synchronic and diachronic point of view, making room for the systematic study of variation, (3) favouring communication between specialists in speech and in phonological theory, (4) providing data and analyses which will help improve the teaching of English as a foreign language. To achieve these goals our project is involved in the construction of a corpus of spoken English from a wide variety of locations in the Anglophone world on the basis of a common protocol. While there are important corpora of spoken English, most of them have been devised on sociolinguistic (rather than explicit phonological) principles and they do not always offer a uniform methodology allowing for a comparison of results and comparable studies of selected problems.

The approach adopted within *PAC* is a well tested one since it has been followed in the international project '*La Phonologie du Français Contemporain*' (*PFC*) coordinated by Jacques Durand (Toulouse II), Bernard Laks (Paris X) and Chantal Lyche (Oslo/Tromsø): for more information see Durand, Laks & Lyche (2002) and the internet site: <http://infolang.u-paris10.fr/pfc>.

The methodology is inspired by the classical work of Labov in that, for each selection of speakers, it involves the reading of a word list and a passage as well as formal and informal conversation (cf. section 2). But in each area surveyed, the speakers (usually groups between 10 and 20 informants) are selected on a network principle well known in the United Kingdom, particularly from the work of the Milroys and their associates (see Milroy 1980).

In the initial phase of the project we favour geographical variation, that is the recording and analysis of cohorts of speakers from as many different locations as possible in the English-speaking world. Within each location, however, we require that the groups include an equal number of men and women and well-defined age ranges (e.g. 70+, 40+, 20+). Social diversity is less easy to achieve with small groups of speakers and it has been found profitable to study family networks which allow for better comparison of age-grading, especially when the social world of the informants has remained stable.

It should be noted that the protocol is neutral as to the selection of informants and it is intended that, after this initial phase, some locations will be analysed from a stricter sociolinguistic perspective. In this initial stage, we control the parameters mentioned above (location, gender, age) and carefully record as much information as possible about the speaker (education, professional status, ethnicity, other languages spoken within the community, etc.). We should stress however that colleagues who, because of limited time and resources, can only study individual speakers on the basis of our protocol are welcome to join the project and contribute to the setting up of our database.

In terms of linguistic study, the recordings obviously lend themselves to various types of exploitation (including syntax and pragmatics). However, all participants in the project commit themselves to studying three areas: 1) Phonological inventories (oppositions and main variants), 2) Rhoticity (Is the accent rhotic / non-rhotic? Does it have /r/ intrusion? How is /r/ phonetically realized in different positions? etc.), 3) T/D (How is the contrast between /t/ and /d/ phonetically realized in different contexts? Is it ever

neutralized? Is there a process of tapping? etc.). Beyond these questions, researchers will obviously pursue their own interests (fast speech processes, stress, rhythm, intonation, etc.).

As in the PFC sister project, once the recordings have been digitalized, our transcriptions of the conversations will be orthographic in the first instance as a basis for more extensive study. The principles and techniques used will be explained in a separate document (see J. Durand and M. Pukli *'The transcription of PAC data: tools and conventions'*, this volume).

The purpose of this document is to outline the methodology we follow to gather the data. Our presentation is intentionally non-technical as our aim is to involve students as well as experienced researchers.

1. Technical Equipment

For the recording of informants the use of a high quality tape recorder is recommended. The ideal is to use a DAT (Digital Audio Tape) recorder which allows us to have digital recordings easily transferable to a computer. Minidisk recorders, which are also digital, are often preferred for obvious financial reasons: they are much cheaper than DATs. However, minidisk recorders involve a heavier compression of data and operate on an analogue output. This can create problems when the data from the recorder is transferred to a computer.

The choice of the microphone is an important one. If possible use a clip-on lavalier microphone. When more than one speaker is being recorded (formal and informal conversation) you can use two microphones, although one of its drawbacks is the increase in the number of connections required. Remember to watch out for possible sources of background noise: avoid open windows, televisions, household appliances, pets; do not record near the window in heavy rain or in a noisy street, and also remember not to move the microphone(s) during the recording.

During or after the recording session you can listen to the tape/cassette/disk (and check sound quality) by using headphones, or by connecting the DAT to an audio system or hi-fi unit. Spend some time getting to know your equipment and make sure that you are comfortable with all technical matters before starting your fieldwork.

Students who are new to this type of fieldwork should inquire about available equipment and borrowing facilities in the audio library or language centre of their universities. Although we have recommended digital equipment, we would like to emphasize that fairly good recordings can be made with high-quality traditional tape recorders, although they are usually less handy to carry around. You can also have a look at the French homepage of the *PFC* project on the internet, where additional information is posted regularly. Note finally that the brief technical discussion offered here is for orientation only. For more detailed discussion, see Tarrier (2003).

2. Procedure

While one of our general aims is to sample a wide variety of contemporary spoken English, individual variation is also at the heart of the *PAC* project. The protocol provides for the recording of various styles for each speaker: (1) reading of two wordlists, (2) reading of a written passage (hereafter called text), (3) formal interview between the fieldworker and the informant, (4) informal conversation. The experience from similar projects is that this approach generally gives rise to a range of registers for each informant. In the ideal case, the recordings show a scale of formality familiar from work in sociolinguistics but the

interpretation of the data needs a great deal of care. In particular, the use of the terms ‘formal’ and ‘informal’ should not be taken at face value. When referring to formal and informal conversations we should not make strong assumptions as to the ‘product’ i.e. the actual style of speech recorded, but rather see these terms as referring to the fundamentally different settings of the two conversational contexts as explained in 2.3.

2.1 Wordlists

Two wordlists are to be recorded for each speaker. They include 192 words altogether, which is rather long but allows for the examination of a wide sample of segmental phenomena. These lists should not be modified since it is imperative to maintain comparability in the project but you can add further items depending on your field of interest. Do bear in mind, however, that additional lists may tax the reader’s patience and goodwill.

For the recording you should have a copy of the two lists in a reader-friendly font style and size, and place all documents to be read under a plastic cover in order to reduce background noise. It is essential to point out that the number preceding each of the words also has to be read. In our experience, informants seem to prefer to read through the lists without any pauses, but you can have a short break between the two lists. The reading over, the informants often want to make comments on our selection of words. This should be encouraged and recorded, as it is a useful source of information on speakers’ attitudes and appreciation of features of their phonological systems.

During the recording of the lists, always have your own copy in order to follow the reader and note any omission, misreading, misinterpretation etc. At the end of the list ask the speaker to read any words which have been missed out or misinterpreted. The same applies to unexpected noises or other interruptions in the course of the reading.

2.2 Text

The text to be recorded, entitled *Christmas interview of a television evangelist*, is a two-page passage originally based on a newspaper article but substantially modified to hide its source and include a number of phonological phenomena worth investigating.

Before recording, ask informants to take a few minutes to read the text to themselves. This gives them a chance to run through the text and know what it is about, and thus be more at ease and perform better for the recording. Never put this task to someone without checking that they will be reasonably comfortable with reading aloud. The speaker’s copy of the text should be printed in a reader-friendly layout with clearly spaced lines and (at least) 14-sized fonts. Also make sure that lines do not break at crucial points (e.g. where linking/intrusive r’s may be possible) and that pages end at appropriate points (e.g. end of a sentence or a paragraph).

2.3 Formal conversation

The formal interview involves the fieldworker and the informant. It has two main objectives: on the one hand, it provides invaluable background information on the speakers, and, on the other hand, it represents one of the four speech styles on the stylistic spectrum aimed to be captured by the *PAC* methodology, i.e. a more formal register than that used in familiar conversation between friends.

After the interview the fieldworker fills in a questionnaire for each informant (cf. the *Information sheet* in 6.1) based on the recorded dialogue. In section 6.2 some topics are suggested that you might want to develop in the conversation, along with useful hints on

some of the techniques to lead such a dialogue successfully, avoiding for example one-word responses from the interviewee.

Please note that there are some essential points we need to know about the informant (for example the parents' mother tongue or whether the interviewee has spent long periods outside his/her region of origin, etc.). You should either incorporate these points into the interview or ask about them afterwards when checking that you have got all details necessary for the information sheet. (However, make sure that informants are willing to be asked about these questions; do not in any way 'force' them to tell you something they are reluctant to disclose.)

Points to remember

- In the formal conversation, the fieldworker takes the lead but it is the informant that should clearly dominate in terms of speaking time.
- Have your questions outlined and prepared before the interview. Interviewees actually expect to be asked (sensible) questions.
- You can dictate the pace and the direction of the dialogue: decide beforehand which items you want to incorporate from the information sheet into the interview proper and which ones you prefer to leave until the end.
- You should plan the conversation to last about 20 minutes.
- Be respectful of interviewees, be careful and tactful with your queries: some might find it embarrassing to talk about their education, family, other personal matters or any random item you are interested in.
- Do not forget that recordings and data collection will be anonymous. Remember to reassure interviewees on this matter.

2.4 Informal conversation

The informal conversation is recorded either with two or more informants without the investigator being present, or with one or more informants and a fieldworker (cf. below and next section). There are no topics or directions imposed on the conversation.

It is best to work in teams of two: one fieldworker who knows the informants well and the other who is a stranger to the target group. This enables the fieldworkers to create two distinct styles in the interviews: a formal dialogue between two persons meeting for the first time, as opposed to the informal conversation between friends.

It is worth noting that taking an active part and participating in the interviews proper facilitates the following phase of the fieldwork, i.e. the transcriptions and analyses. Cooperating with another person is all the more useful as it makes mutual help possible when it comes to listening and transcribing.

Experience has shown that recordings with more than three speakers are rather difficult to exploit. Informal conversations, therefore, should be recorded with two or a maximum of three informants at the same time.

All in all, about 45 minutes of spontaneous speech should be recorded for each informant (appr. 20 minutes of formal and 20 to 30 minutes of informal conversation) – a sufficient basis for having five-minute sequences transcribed orthographically for each stylistic context for each speaker, required for the project at this stage.

3. Working in the field

As just pointed out above, fieldworkers are advised, if possible, to work in pairs, one of them being an inside person who knows the informant well, and the other being an 'outsider', presented as a friend, cf. *Language and Social Networks* by L. Milroy (1980). The formal conversation can thus be led by the person not acquainted, or less acquainted, with the interviewee while the informal conversation will either involve the fieldworker closer to the group under study or two members of the group on their own. Always remember to take some time to get to know the speakers and to put them at ease before switching on the tape recorder.

When working alone with a group of informants, the fieldworker has two options: s/he can do the formal interview and leave the speakers on their own with the tape recorder for the informal conversation, or – if s/he is an 'insider' – s/he might participate in both conversational contexts. In this latter case, a short break should be inserted between the two parts of the interview and an attempt should be made to differentiate neatly the 'formal' and 'informal' contexts (for instance by making a break around a cup of tea). Although being present during both interviews, the fieldworker's roles are clearly different in the two cases: in the 'formal' conversation s/he is 'the fieldworker', while in the 'informal' discussion s/he is a member of the group.

Whether working alone or in pairs, the fieldworkers have to be thoroughly at home in the environment chosen, otherwise the methodology will simply not work.

The collection of data raises important ethical problems which are dealt with in section 7.

4. Further remarks on the selection of informants

As outlined in the introduction, in the *PAC* project we concentrate on small size groups of around twelve speakers, although in specific cases we may want to work with smaller or larger sets of informants. There are manifold reasons for the approach we adopt. First of all, our project is a phonological/phonetic one making room for the study of variation but not studying the latter exclusively. One difference between our project and a fully sociolinguistic one is that we study the whole phonological system characterizing individual speakers rather than a number of variables. Clearly, it is unrealistic to expect to keep track of individual systems once we concentrate on large numbers of speakers (say upwards of one hundred informants). Secondly, as part of our work, we are also attempting to build a reasonably extensive database of orthographic transcriptions and phonological codings for future studies. Again, this is not realistic if our study applies to hundreds of speakers. Thirdly, one of our initial aims is geographical coverage with a common methodology and again this is unrealistic if each area selected involves large numbers of speakers.

In a sense, therefore, our approach is nearer to dialectology than to variationist studies. On the other hand, unlike traditional dialectology, we do not privilege older, male speakers who are supposed to represent the ideal dialect of the area under study. We do build into our methodology a certain amount of variation: equal number of male and female speakers, age-grading and, whenever possible, social differences. Our own approach can therefore be considered as the basis for more extensive sociolinguistic studies for each of the areas we survey.

We note in passing that numbers around thirty have often become standard in recent sociolinguistic surveys. While we cannot but agree that thirty speakers is better than, say,

ten, we are not sure that statistical representativeness can truly be achieved with numbers of speakers which are so small and that reliable trends for the speech of a whole community can always be inferred on this basis. When they are, it is probably because the investigator knows the community and its speech so well that the sample has been constructed in the best way to bring out the relevant parameters of the variety under study. It is not an accident if Labov's well-known study of New York speech (Lower East Side), which was based on firm sociological principles, involved 122 informants. Having said this, we wish to stress once again that our protocol as such is not linked to the number of informants and, if for some areas, the researchers have the resources for studying larger number of speakers or want to concentrate only on sociolinguistic variables on the basis of our methodology, this will be a welcome contribution to our project. Conversely, studies of individual speakers on the basis of our protocol are also a welcome addition to our database.

There are some basic criteria for informants to meet in order to participate in the survey. First of all, informants should clearly belong to the community or area under investigation: this means that they were ideally born and bred there and, if not, that they should have at least come to the area before primary school. Nevertheless, reasonable concessions occasionally have to be made since work within a close network will often favour interviewees who only partially fulfil this geographical criterion. Secondly, a basic level of reading skills is required in order for the speaker to be able to perform the reading task, thus very young and some less educated informants are excluded from the full-scale survey. Thirdly, if possible, distinct social groups should be included within the study. Finally, and most importantly, the target group should be made up of an equal number of male and female speakers falling into at least two (preferably three) well-defined age groups.

5. Readings

Any participant in the *PAC* project should have a good basis in phonology, phonetics, dialectology, sociolinguistics, and corpus work. We will not provide extensive references here but limit ourselves to some essential readings.

There are many good introductions to phonetics. Ladefoged's introductions (1993 and 2001) are a must. At a more advanced level, see Laver (1994).

For a classical approach to the phonology and phonetics of English, Cruttenden's 1994 edition of *Gimson's Pronunciation of English* remains indispensable. For a more recent perspective, see Carr (1999a). Other useful overviews are Giegerich (1992), Roach (2000) and, for American English, Kreidler (1987). Students wishing to explore phonology in more detail are referred to Carr (1993), Durand (1990), Durand & Lyche (2000).

A good knowledge of transcription techniques is required in a project such as ours. The indispensable foundation is provided in *The Handbook of the International Phonetic Association* (1999). In terms of dictionaries, for RP and GA, Roach & Hartman (1997) and Wells (2000) are the standard references. For a recent discussion of some of the issues, further references, and conventions that we use in the *PAC* project for RP and GA, see Durand (2001).

Of course, our description of varieties of English goes well beyond the so-called reference accents. There are many excellent introductions to the subject (e.g. Hughes & Trudgill 1997, Trudgill & Hannah 1985) but the best general overview is still Wells (1982). For the British Isles, Trudgill (1984) and Foulkes & Dougherty (1999) are highly

recommended. Chambers & Trudgill (1998) provide an invaluable discussion of dialectology and its relationship to sociolinguistics.

In terms of sociolinguistics, there are many excellent introductions to the subject: see, for instance, Trudgill (1995) and Romaine (2000). As for the approach adopted in *PAC*, the most useful references are Labov (1972, 1994, 2000) and Milroy (1980, 1987).

The *PAC* project is structured very much on the line of the French *PFC* project. For the latter, see Durand, Laks & Lyche (2002) and all the documents available on the website: <http://infolang.u-paris10.fr/pfc>. Some of the background is provided in Durand, Laks, Lyche (2003). Our combined Labovian/Milroyan approach derives historically from the work on southern French reported in Durand, Slater and Wise (1987). For some personal reflections on variation, see Durand (1993), Carr (1999, 2000). Finally, for an extensive discussion of all matters ranging from recording to the construction and exploitation of corpora, see the various contributions in Delais and Durand (eds.) (2003).

6. Documents

- 6.1 Information sheet
- 6.2 Notes on the information sheet and the formal interview
- 6.3 Wordlist 1
- 6.4 Wordlist 2
- 6.5 Text

6.1 Information sheet

Information sheet on the informant

This sheet is to be filled in by the fieldworker. It consists of two parts: the first one deals with the linguistic and social background of the speaker, the second one is related to the conditions of the recording.

Remember to have the information sheet filled in for all informants participating in the recording, and to note the relationships between speakers when you record an informal conversation with more than one person.

Information sheet

Date of recording:

First name:

Surname:

PAC Identifier:

Age at date of recording:

Place of birth:

Current place of residence (village, town, etc.):

.....

Previous places of residence:

place	number of years	at the age of
.....
.....
.....
.....

Occupation:
Other previous occupations:

Education (specify until what age and what type of education):

Languages spoken:

language	level of proficiency	(basic)	(intermediate)	(fluent)
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

language	frequency of use	(rarely)	(monthly)	(daily)
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Informant's father, year of birth:
Place of origin:
Occupation:
Education:
Languages or local dialect spoken:

Informant's mother, year of birth:
Place of origin:
Occupation:
Education:
Languages or local dialect spoken:

Informant's husband/wife/other:
Place of origin:
Occupation:
Education:
Languages or local dialect spoken:

Number of children, age and education:

People who played an important role during the informant's acquisition of the English language (grandparents, childminder, etc.):

Type of accommodation of the informant (house, flat, in a residential area, housing estate/public housing, block of flats/apartment complex, etc.):

Integration into the area, relationships within the neighbourhood :.....

Ethnic group:

Cultural and leisure activities, travels:

Additional information:

Information sheet on the recording

Interviewer's name (formal conversation):

Interviewer's name (informal conversation):

Length of recording:

Place and setting of the recording:

Location:

Speakers:

Ties between the interviewer and the informants:

Professional:

Friendly:

Family:

Other:

Order of the situations in the recording (e. g.: formal, wordlists, text, informal):

Main topics discussed:

Quality of the recording:

Remarks on the recording (interventions from other people, long telephone interruptions, etc.):

6.2 Notes on the information sheet and the formal interview

A few topics are suggested below that you may want to incorporate into the formal conversation with the informant. Always remember to formulate your questions in a way that the speaker can comfortably develop and expand in response.

You should avoid yes/no questions of the following type (unless you want a short reply to a specific question, e.g. *Are you married?*):

A: Were you born in this part of the world?

B: Yes.

A: Do you like it here?

B: Yes.

A: Were your parents also born here?

B: Yes.

Formulations like the following ones should be preferred:

A: Could you tell me about your childhood/where your parents come from?

A: Do you remember any story from your school years?

A: Could you describe the area you live in?

The conversation should be natural and easy-flowing. You must be tactful not to force people to tell you about events or topics apparently embarrassing to them, especially when you are asking about their family or education (poorly educated speakers).

The themes listed below are useful in drawing a sociolinguistic portrait of the speaker. It is important to obtain information on the following points: at what age the informant and his/her immediate family arrived in the linguistic community under study, the educational level of the informant and his/her parents, time spent abroad or outside the linguistic community, other languages spoken by the informant, his/her parents and grandparents.

Suggested topics (always bear in mind that you need some specific pieces of information, and you should let the informant express him/herself at length at the same time):

- school (good and bad memories, favourite subjects, teachers, etc.)
- family (brothers and sisters, parents, grandparents, origins of the family in the local community)
- languages spoken by informant (level of competence, basic or everyday use, languages spoken by parents)
- occupation (first job ever, current job, place of work, atmosphere, relationship with colleagues)
- time spent away from linguistic community (long holidays, work in other areas or parts of the world)
- place of residence (primary and secondary residences, main features of the area the person lives in)
- spare time activities

6.3 Wordlist 1

Please remember to print your document in a reader-friendly layout, and to ask informants to read the numbers preceding the words.

1. pit
2. pet
3. pat
4. pot
5. put
6. putt
7. sea
8. say
9. sigh
10. sue
11. stir
12. steer
13. stairs
14. err
15. far
16. war
17. more
18. purr
19. moor
20. feel
21. fill
22. fell
23. fall
24. full
25. fool
26. fail
27. foal
28. file
29. foul
30. foil
31. furl
32. bird
33. bard
34. beard
35. bared
36. board
37. barred
38. bored
39. bode
40. bowed
41. bead
42. bid
43. bed
44. bad
45. bard
46. pant
47. plant
48. master
49. afterwards
50. ants
51. aunts
52. dance
53. farther
54. father
55. row
56. rose
57. rows
58. pore
59. poor
60. pour
61. paw
62. paws
63. pause
64. pose
65. wait
66. weight
67. side
68. sighed
69. agreed
70. greed
71. brood
72. brewed
73. fir
74. fair
75. fur
76. four
77. fore
78. for
79. nose
80. knows
81. cot
82. caught
83. meat
84. meet
85. mate
86. naught
87. knot
88. doll
89. dole
90. fierce
91. bird
92. scarce
93. pert
94. start
95. horse
96. hoarse
97. word
98. gourd
99. short
100. sport
101. next
102. vexed
103. leopard
104. shepherd
105. here
106. there
107. weary
108. spirit
109. marry
110. Mary
111. merry
112. sorry
113. story
114. hurry
115. jury
116. bury
117. berry
118. heaven
119. leaven
120. earth
121. berth
122. cook
123. soot
124. look
125. room
126. pearl
127. peril

6.4 Wordlist 2

Please remember to print your document in a reader-friendly layout, and to ask informants to read the numbers preceding the words.

- | | | |
|-------------|-------------|--------------|
| 1. pat | 23. wet | 45. bedding |
| 2. bat | 24. yet | 46. written |
| 3. tuck | 25. witch | 47. ridden |
| 4. duck | 26. which | 48. singer |
| 5. carter | 27. lock | 49. stronger |
| 6. garter | 28. loch | 50. fat |
| 7. fan | 29. earthy | 51. fad |
| 8. van | 30. worthy | 52. lap |
| 9. this | 31. sinner | 53. lab |
| 10. thick | 32. simmer | 54. sack |
| 11. seal | 33. singer | 55. sag |
| 12. zeal | 34. supper | 56. belly |
| 13. bishop | 35. rubber | 57. berry |
| 14. leisure | 36. little | 58. bell |
| 15. heart | 37. middle | 59. bet |
| 16. batch | 38. metal | 60. chutney |
| 17. badge | 39. meddle | 61. kidney |
| 18. rum | 40. bicker | 62. grace |
| 19. run | 41. bigger | 63. graze |
| 20. rung | 42. degree | 64. behave |
| 21. lack | 43. decree | 65. anyhow |
| 22. rack | 44. betting | |

6.5 Text

Christmas interview of a television evangelist

© PAC Project

If television evangelists are anything like the rest of us, all they really want to do in Christmas week is snap at their families, criticize their friends and make their neighbours' children cry by glaring at them over the garden fence. Yet society expects them to be as jovial and beaming as they are for the other fifty-one weeks of the year. If anything, more so.

Take the Reverend Peter 'Pete' Smith, the 'TV vicar' who sends out press releases in which he describes himself as 'the man who has captured the spirit of the age'. Before our 9 a.m. meeting at his 'media office' on Crawshaw Avenue, South London, he faced, he says, a real dilemma. Should he make an effort 'to behave like a Christian' – throw his door open, offer me a cup of tea – or should he just play it cool, study his fingernails in a manner that showed bored indifference and get rid of me as quickly as possible? In the end, he did neither.

'As a matter of fact, John,' he says in a loud Estuary English twang, 'St Francis said, "At all times preach the gospel and speak whenever you have to."' But hey, he didn't mean "Be on

your best behaviour and be happy all the time.” I could have been extra-polite to you, but the real me would have come out as I was talking. You cannot disguise what you are.'

'And what are you then, Pete?'

'Well, I'm a Christian, John. I've been one since I was 14. And I know for sure that Christianity will be judged more on who you are rather than what you have to say about it. Many church leaders don't appear to understand this. They think we can only be really Christian when we are ramming the doctrine of the Creation down people's throats. But if you try to force-feed people they get sick of it and think you're a pain. It's seen as the job of a Christian leader to wear a dog-collar and dress in purple and always be talking about the real meaning of the New Testament. In reality, that turns people right off!'

In many ways, 'Pete' Smith looks exactly how you'd expect a high-profile, born-again Christian to look: tall, handsome, clean-cut and evenly sun-tanned. He has those scarily white teeth that TV evangelists tend to have, and he doesn't wear a dog-collar. In fact, when doing his various religious programmes on Sunday mornings, he has been known to wear a black leather jacket instead, in casual mode. Today, the look is more business-like: metal-rimmed glasses, a grey suit, a blue open-neck shirt, and fashionable black shoes with large buckles. Smith is 44 but he looks a mere 24.

During the whole interview, there wasn't any talk of the poor or the needy but only of his forthcoming trip to China in February and the masses waiting for his message there. I ventured a few questions relating to the charity trust he founded some ten years ago and which, it is generally agreed, employs eight hundred staff and runs schools, hospitals and hostels around the world. And what about the gambling organization he has been willing to advise? Is that a temporary activity or might it be true that he has accepted to be paid to sit on its Board of Directors? Which side is religion on these days? Does money matter? It was as if I had launched a few missiles in his direction. He just sighed in answer: 'I'm only human, John. God knows I do my best and often fail, But it's no skin off my nose if our enemies sneer at some of the good work we do. Truth will out.'

7. Ethical questions

Ethical questions are very important when dealing with data collection and corpora. You cannot constitute, use, share or publish your corpus without the consent and authorisation of the informants. You must explain the scientific purpose of the survey and the way the recordings and personal information will be handled. Make it clear to the participants that in the event of publication or internet use all recordings will be anonymous.

It is best to obtain the interviewee's permission at the end of the session when the speakers are aware of the actual nature of the recordings and the information they contain.

A written form of consent is necessary in order to legally protect the informant and the fieldworker. A model is given below and can be adapted according to the laws and regulations specific to each country of investigation.

Consent form

To all participants in the recordings:

You are being asked to take part in a survey of varieties of English within a project coordinated by the ERSS (CNRS UMR5610) of the University of Toulouse. The recordings are made for scientific and teaching purposes and the results (including recordings) may be

made available in publications (including CDs or other audio formats), shared with other researchers and possibly commercialised and made available on the internet. For the purposes of research and in the event of publication or internet use, all names and information concerning the recorded speakers will be anonymous.

If you give your permission to use the recordings for the purposes indicated above please sign this document below your name and forename.

SURNAME :
FORENAME :
SIGNATURE :

8. Bibliographical references

- Carr, P. (1999a). *English Phonetics and Phonology: An Introduction*. London: Blackwell.
- Carr, P. (1999b). Sociophonetic Variation and Generative Phonology. The Case of Tyneside English. *Cahiers de grammaire* 24: 7-16.
- Carr, P. (2000). Scientific Realism, Sociophonetic Variation and Innate Endowments in Phonology. In N. Burton-Roberts, P. Carr & G. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press.
- Chambers, J.K. & P. Trudgill (1998). *Dialectology*. (2nd Ed.) Cambridge: Cambridge University Press.
- Cruttenden, A. (1994). *Gimson's Pronunciation of English*. London: Edward Arnold.
- Delais, E. & J. Durand (2003) *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse: Presses Universitaires du Mirail.
- Durand, J. (1990). *Generative and Non-Linear Phonology*. London: Longman.
- Durand, J. (1993). Sociolinguistic variation and the linguist. In : C. Sanders (ed.) *French Today : Language in its Social Context*. Cambridge : Cambridge University Press, 257-285.
- Durand, J. (2001). La transcription phonétique de l'anglais : problèmes théoriques et pratiques', in W. Rotgé (ed.) *Agrégation externe d'anglais : mode d'emploi*. Paris: Ellipses, 181-219.
- Durand, J., B. Laks & C. Lyche (2002). La phonologie du français contemporain: usages, variétés et structure. In: C. Pusch & W. Raible (eds.) *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen: Gunter Narr Verlag, 93-106.
- Durand, J. & C. Lyche (2000). La phonologie : des phonèmes à la théorie de l'optimalité. In P. Escudier et alii. (eds) *La parole : des modèles cognitifs aux machines communicantes*. Paris : Hermès.
- Durand, J., C. Slater & H. Wise (1987). Observations on schwa in southern French. *Linguistics* 25, 983-1004.
- Foulkes, P. & G. Docherty (eds.) (1999). *Urban Voices. Variation and Change in British Accents*. London: Edward Arnold.
- Giegerich, H. J. (1992). *English Phonology. An Introduction*. Cambridge: Cambridge University Press.
- Hughes, A. & Trudgill, P. (1997). *English Accents and Dialects*. London: Edward Arnold.
- International Phonetic Association (1999). *The Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

- Kreidler, C. W. (1987). *The Pronunciation of English. A Course Book in Phonology*. Oxford: Blackwell.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1994). *Principles of Linguistic Change. Internal Factors*. Oxford: Blackwell.
- Labov, W. (2001). *Principles of Linguistic Change. Social Factors*. Oxford: Blackwell.
- Ladefoged, P. (1993). *A Course in Phonetics*. (3rd edition.) New York: Harcourt, Brace, Jovanovich.
- Ladefoged, P. (2001). *Vowel and consonants: an introduction to the sounds of languages*. London: Blackwell.
- Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Milroy, L. (1980). *Language and Social Networks*. Oxford : Blackwell.
- Milroy, L. (1987). *Observing and Analysing Natural Language*. Oxford : Blackwell.
- Roach, P. (2000). *English Phonetics and Phonology*. (3rd edition.) Cambridge: Cambridge University Press.
- Roach, P. & J. Hartman (eds.) (1997). [Daniel Jones] *English Pronouncing Dictionary*. Cambridge: Cambridge University Press.
- Romaine, S. (2000). *Language in Society. An Introduction to Sociolinguistics*. (2nd Ed.) Oxford: Oxford University Press.
- TARRIER, J.-M. (2003). L'enregistrement et la prise de son. In J.-M. TARRIER (ed.) *Bulletin Phonologie du Français Contemporain 2*: 5-22. Université de Toulouse-Le Mirail: CNRS ERSS-UMR5610.
- Trudgill, P. (ed.) (1984). *Language in the British Isles*. Cambridge: Cambridge University Press.
- Trudgill, P. (1995). *Sociolinguistics : An Introduction to Language and Society*. (3rd Ed.) London: Penguin Books.
- Trudgill, P. & J. Hannah (1985). *International English. A Guide to Varieties of Standard English*. (2nd edition.) London: Edward Arnold.
- Wells, John C. (1982). *Accents of English*. 3 volumes. Cambridge: Cambridge University Press.
- Wells, John C. (2000). *Longman Pronunciation Dictionary*. (2nd edition.).