

THE TRANSCRIPTION OF PAC DATA: TOOLS AND CONVENTIONS

Table of contents

- 0. Introduction
- 1. *PRAAT* – an introduction
- 2. *Transpraat*
- 3. The orthographic transcription – some basic principles
 - 3.1 Transcription conventions
 - 3.1.1 Simplified punctuation
 - 3.1.2 Turn taking
 - 3.1.3 Overlapping turns
 - 3.1.4 Truncation of words
 - 3.1.5 Repetition
 - 3.1.6 Parentheses
 - 3.1.7 Reported speech
 - 3.1.8 Some features of spoken English in relation to spelling
 - 3.1.9 Acronyms
 - 3.1.10 Dialectal expressions
 - 3.1.11 Reference orthographic systems
 - 3.2 Transcription sample (converted by *Transpraat*)

0. Introduction

This paper introduces the reader to the transcription conventions adopted in the *PAC* (*Phonologie de l'Anglais Contemporain*) project. We first offer a brief presentation of the tools employed in *PAC* with particular attention to *PRAAT*, a speech analysis program which we use for our transcriptions. We then examine the principles that have been developed to facilitate and unify the transcription of the data.

1. *PRAAT* – an introduction

PRAAT ('*doing Phonetics by computer*') is a speech analysis software developed by Paul Boersma and David Weenink at the Institute of Phonetic Sciences at the University of Amsterdam. It is a shareware program, downloadable for free use from the *PRAAT* homepage (www.praat.org or <http://www.fon.hum.uva.nl/praat/>), running under, among others, the most common operating systems – Windows, Macintosh and Linux. (The software is regularly up-dated; the versions referred to in this document are 3.8.64 (2000) and 4.0.26 (2002).)

PRAAT is a highly sophisticated tool for speech analysis with manifold functions in a simple, user-friendly layout. To mention just a few of its basic operations, with *PRAAT* a speech signal or a 'speech object' can be:

- recorded
- read from the hard disk (or other drives) if recorded elsewhere
- handled practically (even relatively long sound files)
- segmented, labelled and transcribed, with the waveform and its duration (as well as, optionally, the spectrogram with formant and pitch contours) displayed in the same window as the text being typed.

Since the *Help* tutorial in *PRAAT* itself ('*Introduction to Praat*'), the *Praat beginners' manual* and the *Praat short tutorial* (in pdf format) are all available on-line on the *PRAAT* homepage and offer a thorough and complete introduction to the program in English, it seems unnecessary to provide a proper introduction here.¹

In order to transcribe with *PRAAT* you need to be familiar with the following procedures (a detailed description of these steps can be found in French in Delais-Roussarie, Durand, Lyche, Meqqori & TARRIER (2002) :

- how to create a sound file in .wav form
- how to open a long sound file
- how to segment and label a speech waveform (creation of 'textgrids')
- how to duplicate, copy, rename and remove tiers
- how to edit, listen and transcribe a speech segment
- how to save a textgrid.

Among the most obvious advantages of working with *PRAAT* are the following:

- you can listen and transcribe simultaneously

¹ Useful presentations in French of transcription tools in general and *PRAAT* in particular can be found in Delais-Roussarie (2003 b), Delais-Roussarie, Meqqori & TARRIER (2003).

- the text is aligned with the corresponding audio sequence
- you can listen to the sequence you transcribe as many times as you wish, and for portions as short as you like
- you can locate any given sequence in the text (from single words to sentences)
- you can move easily within the speech signal by zooming in and out of the sequence
- you can directly extract sounds from the passage transcribed.

And if you wish to do more advanced tasks you can also extract labelled segments (e.g. all segments labelled ‘voiceless stop’) and perform a number of sophisticated phonetic and acoustic measurements, including durational, spectrographic, intensity, formant and pitch analyses.

2. *Transpraat*

Transpraat (a program developed by A. Meqqori for the French *PFC* project) converts the coded text files (or ‘textgrids’) created with *PRAAT* into ‘normal’ text files that you can edit, format and print with your customary word processing tool, such as *Microsoft Word*. Note, however, that the document will no longer be linked to the original audio signal and *PRAAT* textgrid. Thus all eventual corrections or modifications in the transcription should be done and recorded in the original textgrid in *PRAAT*.

In a *PRAAT* textgrid the transcriptions appear within *PRAAT*-specific codes as illustrated in the following 34-second sequence from a conversation below. In this ‘raw’ textgrid format, the codes link each ‘interval’ to the original sound file:

```

intervals [10]:
  xmin = 1177.5648270086172
  xmax = 1181.1347845804989
  text = "F: And how did you find Afghanistan, I mean the
country?"
  intervals [11]:
    xmin = 1181.1347845804989
    xmax = 1191.169782935204
    text = "RM: Er, it's, it's, er, yeah, it was quite a nice place
er, (XX) smelly in some places, the (XX) particularly, er it's very run
down and er, "
    intervals [12]:
      xmin = 1191.169782935204
      xmax = 1196.0005442176871
      text = "RM: (stammering) most of the places, like the, the
mountains or the farmers and stuff, they're, "
    intervals [13]:
      xmin = 1196.0005442176871
      xmax = 1202.334784580499
      text = "RM: they're generally quite friendly but er, very, very
poor, it's like er, stuck in er, the Middle Ages I think."
    intervals [14]:
      xmin = 1202.334784580499
      xmax = 1211.4604081632654
      text = "RM: But er, yeah, it was not bad, (XX) not bad. Very
rocky, the mountains, some of the, the mountains we were up were p/
particularly high."
    intervals [15]:
      xmin = 1211.4604081632654
      xmax = 1216.8312416315391

```

The same sequence can be seen below, converted by *Transpraat*: there are no codes and speaking turns are automatically adjusted.

F: And how did you find Afghanistan, I mean the country?
RM: Er, it's, it's, er, yeah, it was quite a nice place er, (XX) smelly in some places, the (XX) particularly, er it's very run down and er, (stammering) most of the places, like the, the mountains or the farmers and stuff, they're, they're generally quite friendly but er, very, very poor, it's like er, stuck in er, the Middle Ages I think. But er, yeah, it was not bad, (XX) not bad. Very rocky, the mountains, some of the, the mountains we were up were p/ particularly high.

The program can be freely downloaded from the *PFC* homepage: <http://infolang.u-paris10.fr/pfc>, where a manual on its use (which is extremely simple) is also available².

3. The orthographic transcription – some basic principles

The conversations recorded within the *PAC* project are transcribed in standard spelling. This is clearly a necessary first step for any further analysis, which can then be followed up by phonemic or phonetic transcriptions, and other codings of the discourse.

An orthographic transcription:

- helps us assess the linguistic content of the sequences transcribed in the corpus, and mechanically find occurrences of the linguistic phenomenon we are interested in;
- remains neutral as to the phonemic or phonetic inventory of a given language; transcribing without presuppositions about the phonemic inventory is especially useful, if not crucial, since one of the goals of the analysis is precisely to establish phonemic/phonetic distributions for the variety under study;
- is (almost always) straightforward; it is unnecessary for the transcriber to spend "hours" deciding the actual quality of various sounds, etc. at a stage when the specific range of problems to be tackled have not yet been fully defined;
- is undoubtedly less subjective as a first step than a phonemic/phonetic transcription; the latter is, in fact, an advanced form of analysis in which the transcriber's personal judgement is inevitably involved, and the (eventual) inaccuracy of which can easily impede further research;
- last but not least, for the reasons enumerated above, it is simpler and therefore less time-consuming to do than a phonemic/phonetic transcription. For large-scale projects, orthographic transcriptions have in fact become the norm in the initial stage of work on the raw acoustic data³.

3.1 Transcription conventions⁴

There are important basic conventions established within the project in order to have uniform transcriptions. They should be strictly adhered to if we wish to have comparable data.

² See also Meqqori & Durand (2003).

³ For justification of this position, see Delais-Roussarie (2003), Durand & Lyche (2003).

⁴ Adapted on the basis of the original principles proposed for the French *PFC* project; cf. *Transcription des données: Outil et conventions* by E. Delais-Roussarie, J. Durand, C. Lyche, A. Meqqori & J.-M. Tarrier (2002).

General note:

We transcribe under *PRAAT*, on a single tier, regardless of the number of speakers involved. Interval boundaries are added according to the logic of turn taking (a new interval for each new turn). However, if stretches within boundaries are too long, a true phonemic/phonetic alignment may prove difficult at a later stage. We therefore request that interval units should not normally exceed 15 seconds. No carriage returns are used, and the speaker is identified at the beginning of each interval.

3.1.1 Simplified punctuation

The punctuation system is simplified: the full stop, the comma and the question mark are the only symbols from traditional spelling used for the transcription of discourse in the project.

JV: I don't know what to do with it, I mean I've never looked at a language that way, which is sort of going out and not knowing anything.

Commas indicate a brief pause in the discourse, or a 'non-final', 'continuing' intonation contour marked by a shift in pitch or other cues.

TB: So I was home. I won the airline tickets.

Full stops stand for a relatively long pause in the discourse, or for a 'final' intonation contour.

DH: How many of these are you going to have?

A question mark is inserted at the end of a question.

NB:

- Pauses and intonation contours do not always coincide with expectations based on syntax.
- Pauses and intonation units are not distinguished along rigorous lines in the orthography employed here; such a finer supra-segmental transcription remains an optional subsequent task.
- Commas are used between repeated words or expressions (cf. 3.1.5).
- An exceptionally long pause in an otherwise logically/syntactically coherent sequence will be indicated by a parenthetical remark (cf. 3.1.6).

LC: but overall I'd say, (silence) a little less than half, of those who apply.

3.1.2 Turn taking

At the beginning of each turn the speaker is identified by his/her initials, which are followed by a colon (a space is inserted on its right, but none on its left). The fieldworker is designated by the letter **F**.

F: So, do your parents agree with you?

JF: Well, not really.

As mentioned above, there is no carriage return to mark the end of a sentence or paragraph. The discourse of a single speaker is transcribed continuously under *PRAAT* (with regularly added interval boundaries, each unit being headed by the initials of the speaker).

3.1.3 Overlapping turns

Turns often overlap in a conversation; three types of interventions are distinguished in the transcription:

- Background responses, typical fillers such as ‘yeah’ ‘really’, laughter, vocal and other noises uttered by the listener to maintain interaction are ignored.
- Short interventions – i.e. when the listener interrupts the speaker but does not initiate a new turn, and the speaker goes on speaking – are transcribed within angled brackets in the following manner:

LC: So it's, it's that the approach <F: The approach.> is different.

DR: I mean he may get uh, <F: But Nixon came back. I think if I remember he was beaten once and then.> yeah, yeah that's pretty unusual, pretty unusual.

F: So it's really your grandparents who are Japanese speakers? <JF: Yeah.> Your mum and dad are really English speakers <JF: Yeah.> their, their first language is English?

NB: Here ‘yeah’ has a real ‘response’ value, and, therefore, is transcribed.

- When a listener interrupts the speaker and then ‘takes over’ the conversation, his/her words uttered at the same time as those of the previous speaker are transcribed between angled brackets as indicated above, and a new turn is marked by a new interval (under PRAAT).

F: Do you feel American above all or what do you feel? <TS: Sure I,>

TS: I guess I don't know what that really means, (laughter) I've, you know, I'm an American but, I don't, I'm not like, ‘yeah I'm an American’, you know?

3.1.4 Truncation of words

A slash (followed by a space) indicates unfinished words:

TS: You think you have this demo/ democratic freedom but it's, not really there.

DH: Well they used to, what ab/ what ab/ what about those uh, the uh, the Concord jets.

LC: the col/ the faculty are looking for a good fit.

JF: My lo/ (laughter) uh it's like non-existent.

Truncated intonation units (when speakers do not finish their train of thought, are interrupted, or hesitate, etc.) are marked by a comma or a full stop:

TS: you know, I am an American but, I don't, I'm not like, ‘yeah I'm an American’, you know?

3.1.5 Repetition

Repeated words or expressions are separated by a comma.

DR: I, I like to go skiing in the snow, but I don't want to have to dig my way out of it every day.

JF: I think it's true that, that, there is racism in, racism in, in California but it's really well-hidden.

NB: Commas mark repetition and short pauses in the discourse. Thus in the following example the first comma stands for a short pause, the second for a repetition, the third indicates a repetition that coincides with a short pause at the same time, and the fourth one marks a short pause:

JF: Uh, it's okay it's you know it's, it's really, it's really weird teaching you know, I don't know.

3.1.6 Parentheses

Observations made by the transcriber on non-linguistic aspects of the interaction (noises, stammering, laughter, etc.) and on the recording (background action, quality problems) are placed between parentheses.

DG: That's at the beginning of the week so it's hard to remember. (laughter) Uh, we read a couple of theoretical texts comparing irony to allegory,

TB: My father, he is from Canada. (door opens, F returns) Actually he was born in Massachusetts.

Unintelligible words are indicated by the capital letter X in parentheses. The number of Xs inserted (ideally) corresponds to the number of incomprehensible syllables:

JV: because not (XX) all the cases are uh, show up in the pronoun system,

Words are often hard to decipher due to noise or other interferences, in this case the commentaries are inserted in separate parentheses:

RF: kicked everyone out of the airport and made to go you know (noise) (X) shoot the bag and see if it blows up, and uh,

In cases where the transcriber thinks s/he has probably recognized a word (or sequence of words) but is not fully sure, the word is put in parentheses:

JG: Maybe I'll stay in the technology sector, and uh hopefully do something with creativity, like maybe product design, or writing you know (maybe) marketing oriented, something like that. (laughter)

3.1.7 Reported speech

Reported speech is transcribed between inverted commas (' '):

DR: And then when Bush said 'read my lips no new taxes' and then, you know,

TB: And there was a woman at the other line and she said, 'oh no message', and so I was

TB: and she said I had won the prize and I said 'didn't you just call'

3.1.8 Some features of spoken English in relation to spelling

- Obviously, many reductions and contractions occur in spontaneous speech. Contracted forms are used in our transcriptions only in so far as they are allowed in standard spelling. Note, in the following example, the co-occurrence of a non-contracted and a contracted form, the former bearing a slight emphasis.

JG: Yeah I have heard that and also I've heard that he seems to be very needy of getting votes.

Sometimes non-contracted forms appear in a more formal style:

F: And were your parents from there?

TS: My mom has lived in Los Angeles all her life.

- Word internal ellipsis is an equally frequent feature of spoken English. To avoid a waste of energy at the initial stage of transcription, such deletions are not transcribed. The examination of these features is left to the phonological/phonetic stage of the analysis.

LC: Some very, very intelligent young people, will apply but not do well here because they needed more structure. (and not 'cause)

LC: a portfolio for music, you know original music compositions (and not 'riginal)

- But note that we do not reintroduce words (or word sequences) which appear to have been missed out (in relation to normative grammar). Thus if what we hear is:

F: Was she there?

LC: Think so.

We do not transcribe:

F: Was she there?

LC: I think so.

- Realizations for which standard orthography offers distinctions will be transcribed accordingly. Thus the distinction between *yes* and *yeah* is systematically respected in the transcription.

TS: I don't know. Yeah. <F: It's confusing.> It's confusing. (laughter) Yes there, there's a lot involved and I think, to be, to say a real opinion on it you have, I have to be, really informed.

F: But do you feel now you're from California ? <TS: Yes.> That you're Californian? <TS: Yes. I guess. (laughter)>

- Interjections are another characteristic feature of conversations, employed to express pain, surprise (*ouch, oops*) etc., or simply to provide feedback and to signal active participation towards the other party in the discourse (*uh huh, oh,*

ah, hm). For these speech forms, we use the conventions put forward in the OED.

Most often, however, the speaker is simply using a filler to gain time while thinking, hesitating, or searching for an expression (*hm, uhm, uh, er*) etc. Regardless of the actual sound pronounced, this type of intervention will always be described as ‘*er*’ for British, and ‘*uh*’ for American speech.

RM: Er, it's, it's, er, yeah, it was quite a nice place er, (XX) smelly in some places, the (XX) particularly, er it's very run down and er

DG: Uh, let's see, uh, I uh, I'm from L.A. and I let's say I've been moved uh always to magnet schools which are like schools that kind of specialize in one thing or another

3.1.9 Acronyms

Acronyms – pronounceable words made up from the initial letters of a multi-word name like, for example, UNESCO for the United Nations Educational, Scientific and Cultural Organization – are written in the usual way: capital letters with no separation of any kind if the word is pronounced as a unit. If on the other hand it is spelled out letter by letter, this is indicated by writing a full stop after each letter of the word: U.N.E.S.C.O.

Any unexpected form of actual pronunciation will be indicated in parentheses after the word in SAMPA transcription. SAMPA (Speech Assessment Methods Phonetic Alphabet) is a machine-readable phonetic alphabet developed by speech researchers from many different countries in the late eighties. It is to date the best international collaborative basis for a standard machine-readable encoding of phonetic notation mapping the symbols of the International Phonetic Alphabet onto ASCII codes. As with the ordinary IPA, a string of SAMPA symbols does not require spaces between successive symbols.⁵ (For an example of SAMPA transcription cf. next section.)

3.1.10 Dialectal expressions

Words or expressions that do not belong to either standard British or American English will be transcribed by using SAMPA symbols.

LC: If you want to tell someone to shut up, you say '(hod j@r wiSt)'.

However, if there is a longer stretch of discourse in dialectal speech, "normal" spelling will be employed. If there is a reference dictionary of the dialect being described, its conventions should be used.

3.1.11 Reference orthographic systems

In our transcriptions, we apply the spelling system normally used in the country where the speakers live or come from. Thus, if we transcribe British varieties of English, we use standard British English conventions (adopted in the OED). If we transcribe American English, we use the conventions adopted in Webster's (cf. hesitation 3.1.8). Examples transcribed according to the British and the American conventions, respectively:

⁵ Wells, J. Online. <http://www.phon.ucl.ac.uk/home/sampa/home.htm> (2002 December 16).

Standard British English:

DR: he can't honour the guidelines of the debate for even ninety minutes

RF: when I was, I think, maybe thirteen, just travelling with my mum, and er

American English:

DR: he can't honor the guidelines of the debate for even ninety minutes

RF: when I was, I think, maybe thirteen, just traveling with my mom, and uh

3.2 Transcription sample (converted by *Transpraat*)

Recording made at the University of Santa Barbara in California:

F: So you're in which year of your, of your course now at the moment?

JG: Fourth year senior year. <F: Fourth year.> Senior. <F: And>

F: what do you, what's your, what's the future for you, what are you going to do this year and next year as far as you can see, and beyond ?

JG: Right. Well I'm doing a lot with computers right now, and I really, love music, especially singing which I think I probably got from my grandmother. Uh, my father has nothing to do with music, (laughter) interestingly enough, but uh, in his genes they were passed on. (laughter) Uh, yes I like computers and music and obviously, I'm an English major so, writing as well so, hopefully some sort of fusion of, of all three. You know I think there, I mean this is the College of Creative Studies so, we're sort of gearing up for, nontr/ non-conventional lifestyles I'd say you know rather than, because I have done some corporate work, work which is really, a grind you know, (laughter) forty or fifty hours a week of uh, doing you know sort of middle management things can just be, real tiring so hopefully, s/ s/ hopefully, my livelihood will be something less conventional, <F: How old are you now?> more creative. Twenty-one.

F: Twenty-one. So you're not sure what you will do in life? What do you see yourself if you, project yourself ahead, what do you see yourself, (X) say in four, three or four years' time ?

JG: Well, possibly uh. Again maybe, you know I'm really interested in creativity and computers and writing so, I think uh. Maybe I'll stay in the technology sector, and uh hopefully do something with creativity, like maybe product design, or writing you know maybe marketing oriented, something like that. (laughter)

Bibliographical references

- Boersma, P. & Weenink, D. *PRAAT: doing Phonetics by computer*. Online. <http://www.fon.hum.uva.nl/praat/>
- Delais-Roussarie, E. & J. Durand (eds.) (2003). *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.
- Delais-Roussarie, E. (2003a). Constitution et annotation de corpus : méthodes et recommandations. In E. Delais-Roussarie & J. Durand (eds.) *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.
- Delais-Roussarie, E. (2003b). Quelques outils d'aide à la transcription et à l'annotation de données audio pour constituer des corpus oraux. In E. Delais-Roussarie & J. Durand

- (eds.) *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.
- Delais-Roussarie, E., A. Meqgori & J.-M. Tarrier (2003). Annoter et segmenter des données de parole sous PRAAT. In E. Delais-Roussarie & J. Durand (eds.) *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.
- Delais-Roussarie, E., J. Durand, C. Lyche, A. Meqgori et J.-M. Tarrier (2002). Transcriptions des données : outils et conventions. *Bulletin PFC* 1 : 21-34. CNRS ERSS-UMR5610 et Université de Toulouse Le Mirail.
- Durand, J. & C. Lyche (2003). Le projet 'Phonologie du Français Contemporain' (PFC) et sa méthodologie. In E. Delais-Roussarie & J. Durand (eds.) *Corpus et variation en phonologie du français : méthodes et analyses*. Toulouse : Presses Universitaires du Mirail.
- Meqgori, A. & J. Durand (2003). Manuel d'utilisation de Transpraat. *Bulletin PFC* 2 : 39. CNRS ERSS-UMR5610 et Université de Toulouse-Le Mirail.
- Wells, J. (2002 December 16). *SAMPA computer readable phonetic alphabet*. Online. <http://www.phon.ucl.ac.uk/home/sampa/home.htm>