

# A multimedia approach to eartraining and IPA transcription

Jo Verhoeven, Robert Davey  
City University London

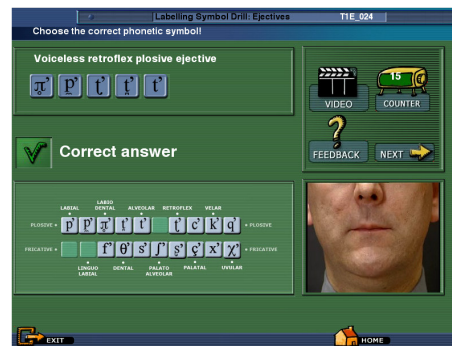
**1 Introduction** Besides theoretical reflection about and a taxonomical interest in the sounds of the world's languages, aspiring linguists and phoneticians require the necessary practical skills to perceptually distinguish sound differences and to produce these themselves. These ear training and performance classes are generally labour intensive, time consuming and consequently relatively expensive. During these practical classes, educational output is extremely high, whereas outside the classroom students are left pretty much on their own. In order to solve this problem, a multimedia package was developed which enables students to practice the perception/transcription of a wide range of phonetic distinctions outside the classroom. During the training process, students receive instant feedback about their progress, whereas teachers will have access to meaningful reports about their students' performance. As a result, classroom teaching becomes highly efficient in that it can concentrate on specific problems experienced by individuals or groups of students.

**2 Exercise types** Eartrainer provides four types of transcription exercises focusing on different subskills necessary in phonetic transcription: phonetic labelling, symbol drill, guided transcription and free transcription. Within each component, the exercises are grouped together according to the different sound classes distinguished by the IPA: plosives, ejectives, implosives, clicks, nasals, trills, fricatives, approximants and vowels. After intensive practice on each of these sound classes, students do mixed exercises to consolidate the acquired skills. The design of the package is intended to reflect the traditional phonetic taxonomic categories as much as possible. This is to say that the ordering of the sound classes on the menus, for example, reflects the various constrictions of the vocal tract in producing the sounds. As such we wanted to stimulate students also to reinforce the traditional phonetic framework in the physical design (the looks) of the package.

**2.1. Phonetic labelling** In the phonetic labelling component, students see a phonetic description of a sound in terms of the phonetic taxonomic labels referring to airstream mechanisms, voicing, place of articulation and degree of constriction. In each exercise, they have to choose the appropriate IPA symbol from a series of five options. This component thus aims to relate the meaning of the traditional phonetic taxonomic concepts to individual IPA symbols. After students have made their selection, they get instant feedback as to the correctness of their choice and they get the option to listen to (and watch) the sound being spoken by a trained phonetician. All the sounds in this exercise type are presented between two open-mid vowels, except for the vowel exercises in which the vowels are positioned after a bilabial oral stop.

**2.2. Symbol drill** In the symbol drill component, students see/hear a sound and they have to choose the correct phonetic symbol from a series of five options. After their choice, students are corrected if necessary. The target sounds occur in simple nonsense words which have a maximum of two syllables. The sounds to be recognized can occur word-initially, medially or word-finally and are never part of a cluster. The aim of this

component is to relate sound perception to the IPA symbols. The graphical interface of this exercise type is illustrated in figure 1 with respect to the plosives:



**Figure 1.** Graphical interface of the Phonetic Labelling exercises.

In the two exercise types described so far, students also have the possibility to obtain more information about the phonetic symbols which are used: after doing an exercise on e.g. plosives, a feedback button lights up which gives access to a survey of all the phonetic symbols for the plosives with their place-of-articulation labels. At the moment, these symbols are presented on the well-known IPA grid, which orders the symbols in terms of place of articulation from left (front articulations) to right (back articulations). We are presently investigating the possibility to make these feedback screens more educational by relating the symbols to a graphical representation of the vocal tract.

**2.3. Guided transcription** The third type of exercise is guided transcription. Students see/hear a video recording of the pronunciation of nonsense words which have been partially transcribed. It is their task to choose the correct symbols for the missing sounds. For this purpose, students have access to all the symbols for a particular sound category, i.e. for the exercise on plosives they have all the plosive symbols at their disposal. The number of sounds to transcribe differs from word to word, as does the phonetic context in which the missing sounds occur. In view of these parameters, it can be assumed that this type of exercise is considerably more difficult than the previous two.

In the student assessment stage of this type of exercise students are given two opportunities to complete the transcription. After completing the transcription in a first trial, they click the evaluate button: the system indicates which symbols are correct and which are not. In the second trial, they can listen to the production of the word again and change the transcription of the incorrect symbols before asking the system for a final assessment of the transcription. Only the results of the final assessment are stored on the student-feedback screen.

**2.4. Free transcription** The fourth type of exercise is free transcription and it comes closest to the real manual transcription of speech. Students listen to a nonsense word which is pronounced by a phonetician and they are given a visual indication of how many sounds they have to transcribe. Students complete the transcription by choosing the correct phonetic symbols from a keyboard containing all the symbols that may be necessary. The transcription does not have to be completed sequentially from left to right, but students can start transcribing anywhere in the word and in whatever order they want. This ensures that they can provide the symbols for sounds they recognise first and then concentrate on the sounds that may find more difficult.

Eartrainer assesses the transcription in two rounds. In the first round it indicates which segments were transcribed correctly. Subsequently, students get the opportunity to revise their transcription before having it assessed for the second time.

It should be mentioned that the focus in the free-transcription exercises is on the identification and transcription of main segments only. Different kinds of secondary aspects of articulation that are usually transcribed by means of diacritics have not been included.

The graphical interface of the free transcription module of Eartrainer is illustrated in figure 2:



**Figure 2.** Graphical interface of the free transcription module of Eartrainer.

**3 Phonetic symbols** Eartrainer has a total of 206 phonetic symbols, which were designed graphically as small individual buttons. This has two important advantages. In the first instance, students do not have to install any phonetic font in order to use the package: this avoids all kinds of cross-platform compatibility problems. Furthermore, students do not have to actually learn to use a phonetic font on their computer. This can be quite tedious, especially in terms of the relationship between the keyboard buttons and the actual phonetic symbols.

The second advantage of a graphical representation is that the phonetic symbols (contained in graphics files) are coded in terms of the phonetic dimensions relevant to the sounds in question. These codes can then be used in the feedback process of the transcription training.

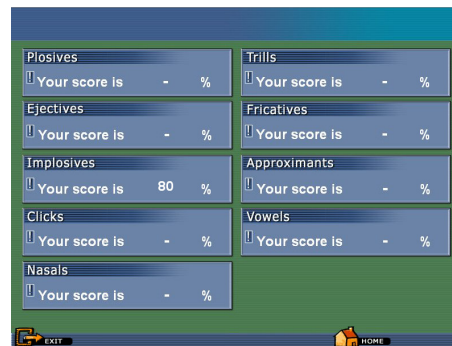
As far as the criteria are concerned on the basis of which it was decided to include certain phonetic symbols and exclude others, it would have been possible to work purely on the basis of frequency of occurrence of sounds in the languages of the world. Calculations in Maddieson (1984) could have provided good and objective guidelines. The result of such statistical approach would have been, however, that less frequent sound categories would not have been included, even though they represent phonetically interesting taxonomic categories. Therefore, it was decided to include as many main segments as possible in order to provide substantial phonetic variety to the learner.

In the near future, all sound segments will be coded for frequency of occurrence in the languages of the world, so that these can be taken into account in the error analysis and the feedback process: mistakes against more frequent sounds will e.g. be penalised more heavily than mistakes against less frequent ones.

**4 Sound recordings** All the sounds and nonsense words were read by a male phonetician who has had extensive training in phonetic ear training and performance: his deliveries were recorded on video in a professional studio. The video recording zoomed

in on the lower part of the speaker's face so that the articulation of the sounds is clearly visible. The video clips for the individual sounds and nonsense words were digitized and converted to a Quicktime format.

**5 Educational tools** Students have a number of educational tools at their disposal in Eartrainer. At present, they have access to information about how well they have been transcribing in that they get an indication of the percentage of correct transcriptions for the different sound classes. All this information is grouped together on a student feedback screen, which is illustrated in figure 3:



**Figure 3.** Student feedback screen in Eartrainer.

One of the planned future developments is fully-fledged phonetic error analysis, which will provide information about the phonetic dimensions which are most problematic to the learner. Errors will be analysed in terms of consonants/vowels, voicing distinctions, manner of articulation, place of articulation and a number of aspects of articulation such as nasal/oral distinctions, central/lateral etc.

**6 Compatibility** At present, Eartrainer is available on cd-rom, compatible with PC and Mac OS. It is being considered to also make the package available as an internet application.

**7 Conclusion** We are convinced that Eartrainer has the potential to significantly contribute to optimising eartraining classes and phonetic transcription. It allows students to practice on an individual basis outside the walls of the traditional classroom and the feedback can be used by teachers to make the teaching process in the classroom more efficient. In addition, a number of future developments are being considered such as the inclusion of native speakers of a variety of languages as well as pathological speech.

## 8 References

Maddieson, Ian (1984) *Patterns of Sounds*. Cambridge: University Press.