

A SYSTEM FOR INDEPENDENT E-LEARNING OF PRACTICAL PHONETICS

Christel de Bruijn^a, Miguel Baptista Nunes^b, Linhao Fang^b, Rigved Pathak^b & Jingchao Zhou^b

^aCentre for Health and Social Care Research, Birmingham City University, UK;

^bInformation School, University of Sheffield, UK

christel.debruijn@bcu.ac.uk; j.m.nunes@sheffield.ac.uk

ABSTRACT

This paper presents a web-based system that supports students in their learning of phonetic transcription and production. It allows students to create their own exercises from an integrated database of recordings, according to their specific learning needs. Every recording in the system has been meta-tagged. For example, a cardinal vowel will have, among others, the tags 'vowel' and 'cardinal'. Exercises are automatically generated by the system by searching the meta-tags which are stored in a database. The paper describes the types of recording in the system (individual phones, English words, nonsense words and short phrases), and the options from which the students can choose to create and customise their transcription and production exercises. It discusses how feedback to the student is handled, and the conventions that have been used in the transcription of the recordings. Finally, it describes the maintenance aspects of the system, as well as some of the technical details.

Keywords: phonetic transcription, ear training, independent learning, phonetics teaching, online

1. INTRODUCTION

This paper reports on the ongoing development of a web-based system that supports students in their training of phonetic transcription and oral production skills, as part of their degree course in speech and language therapy. Phonetic transcription is considered a fundamental aspect of the practice of speech and language therapists, and training in transcription must therefore form part of the curriculum of courses in the UK, as recommended by the Royal College of Speech and Language Therapists (RCSLT) [7].

Providing transcription training is, however, labour intensive; in order to provide feedback to individual students, groups should remain small [3]. As a result, transcription training can be very costly if teacher-student classes are the only way to

deliver this. Another drawback of relying solely on class time for transcription training is the problem of the greatest common denominator. Very few teachers, if any, will be in the lucky position where there is sufficient time to develop a transcription training program tailored to each individual student. Because of this, teaching materials for classes are likely to be selected or developed based on the teacher's perception of difficulties that are commonly experienced by students.

Of course, it is unlikely that any transcription training provision would be solely based on class time only. Indeed, a survey published in 1997 [2] about the state of phonetics education in Europe reveals that a combination of different delivery methods was often used within institutions for transcription training, such as lectures, practical work, self study and (in some instances) the internet. However, a mixture of delivery methods still does not necessarily address the issue of focus on the most common student difficulties. Self-study materials are also at risk of being geared towards the learning needs of the group rather than those of the individual simply because as a teacher one always aims at providing materials that achieve the maximum benefit possible. In order to address the learning needs of the individual, learning materials must be tailored to the individual. One way of doing this is by letting the learner state his learning needs explicitly and provide materials accordingly. In this paper, we describe a system which allows the learner to do just that, and which provides learning materials accordingly.

The system described here creates an environment that allows students to take ownership of their learning process by reflecting on their learning needs and selecting those areas that need further development. For example, students who have difficulty discriminating between particular sounds can practise these by creating their own exercises from an integrated database of sounds. This allows them to work at their own pace and

create and carry out exercises that are adapted to their specific individual difficulties and needs. The identification of these learning needs and creation of exercises by the students emerges from the level of difficulty set in the tutor-led workshops. The system has, therefore, been developed not to replace teacher-student contact time, but to be used by students alongside these classes.

2. PEDAGOGICAL RATIONALE

This e-learning system is embedded in a wider Higher Education blended learning environment that is composed by formal lectures, small group tutorials and web materials in a Virtual Learning Environment (VLE) - Moodle.

The basic pedagogical underpinning for the system is a constructivist approach aiming at fostering independent learning, self-assessment of learning needs and ownership over the learning process. This requires a system that enables learners to engage in meaningful and authentic interactions. Therefore the design of the system described in the following section aimed at providing a rich collection of structured recorded materials that the students could use at their own pace, based on their own learning needs. It is expected that in this way students can interpret and construct meaning based on their experiences and interactions in the whole of the blended learning environment, but that this system can be pivotal in their control and ownership of their own learning.

3. FUNCTIONALITY AND STRUCTURE OF THE SYSTEM

The system consists of three components, namely for administration, transcription and production, which are described below.

3.1. The transcription module

3.1.1. Types of exercises

At the start of the transcription exercise, students can choose from four different types: transcription of individual phones, English words, short phrases and nonsense words. The relevance of the first three types of material is clear. In addition, nonsense words have been included because, as Ladefoged [5] p.459 states: "Part of teaching is a matter of breaking presuppositions". People have expectations about a sound due to the context in which it is produced. By taking away this context, (which, in the case of nonsense words is done by

taking away the language context of that word) one takes away potential biases in perception. It therefore helps students in paying attention to the phonetic details. When starting an exercise in transcription of individual phones, students are presented with a choice between vowels and consonants. Choosing consonants will bring up a chart of pulmonic consonants on the screen, similar to the layout of the IPA chart, from which they can select all consonants they want to include in the exercise. Choosing vowels instead, will give them a further choice between cardinal vowels or English vowels (and the latter again a further choice between monophthongs, diphthongs or both). From here, students can again make their specific selection. Instead of selecting and practicing specific sounds, it is also possible to ask for a random selection of (cardinal or English) vowels or consonants. This may be particularly useful for self-testing purposes and for identifying problem areas. If instead of transcribing individual phones, a student has chosen to transcribe English words, a different choice menu will be presented, starting with a choice between monosyllabic, bisyllabic and polysyllabic words. Therefore, a path of increasing difficulty can be followed. Further choices can then again be made about the sounds those words should contain (e.g. monophthongs, diphthongs or both, as well as specific consonants). As for individual phones, a random selection of words can be requested instead as well. For the transcription of nonsense words, the choice menu is more restricted. Students can choose words that contain specific sounds, or they can ask for a random selection of nonsense words. The database contains words of a variety of consonant-vowel structures, such as CVC, VCV, CCCVCC etc. For transcription of short phrases, students can only get a random selection.

3.1.2. Feedback

Transcription exercises of some systems, e.g. [1, 8] require students to make a selection from a (restricted) set of symbols presented on screen, or to input symbols via a keyboard. In the system described here, however, students are required to carry out their transcription on paper. This allows them to practise the actual writing of the symbols. It also allows students to train their transcription by recalling symbols from memory, without aid, rather than by recognising the correct symbol from

a predetermined set of options. Feedback is given on screen. Students can request to receive feedback after each transcribed item, which may be useful in exercises with short phrases, or they can receive feedback after an entire exercise. This option is provided in order not to interrupt the drill aspect of certain exercises. Take, for example, the case of a student who regularly transcribes [u] instead of [y]. The system allows the student to create a string consisting of a random sequence of these two (or more) vowels. The length of the sequence can be set by the student. In exercises like these, it may be more useful to receive feedback after the entire exercise, rather than after each sound.

3.1.3. *Transcription conventions*

The level of detail in transcription depends on the purpose for which it is required. In the field of speech and language therapy, it can be a thorny issue. Too little detail (such as in a strictly phonemic transcription), and the transcription does not reflect all of the relevant articulatory reality. This might result, in atypical or developmental speech for example, in phonemic contrasts being masked. On the other hand, as Ladefoged [6] p.47 states: "In practice, it is difficult to make a transcription so narrow that it shows every detail of the sounds involved". Some compromise between the two extremes of the narrowness continuum [4] thus needs to be found. As the handbook of the IPA [4] p.29 states: "it is possible (and customary) to be selective about the information which is explicitly incorporated into the allophonic transcription". The system described in this paper does not contain any atypical or developmental speech which makes the issue of what to include and what not, somewhat easier. The main conventions that have guided the transcription of the recordings in the system will now be outlined. Firstly, recordings are transcribed at an allophonic level, with transcription of extrinsic allophones to allow for the representation of assimilation, e.g. the assimilation of an alveolar to a labiodental nasal in words such as 'infinite'. Intrinsic allophones are not transcribed currently. These are variants which are the automatic consequence of the movements of the articulators (such as a nasalised realisation of the vowel in 'pen'). Secondary articulatory features (such as the velarisation of /l/) are also not transcribed. We do not argue here that the transcription of some of these features is not valuable. On the contrary, as

outlined above, transcription of such variants can be highly important. The system is, however, intended to be used by students with no previous experience of phonetic transcription. To provide very narrow detailed transcriptions was deemed counterproductive at this level as the students do not yet possess the knowledge to understand such level of detail. It should also be pointed out that this system was not developed as a replacement for face-to-face transcription classes. Instead, it is intended as a self-study resource for students that can be used alongside transcription classes in which more detailed transcriptions can be discussed when appropriate. Transcription of vowels was carried out according to the lexical set for RP provided by Wells [9]. The only adaptation that was made to this set was the adoption of [eə] instead of [ɛə] for the standard lexical set SQUARE. This is, however, in accordance with the transcription that Wells himself later also uses for this set in the Longman Pronunciation Dictionary [10].

3.2. **The production module**

It is commonly believed amongst those teaching phonetics, that training students in the production of speech sounds is beneficial for their transcription skills. As Ladefoged [5] p.459 states: "It is certainly true that if you can produce a difference between two similar sounds, then you find it easier to hear the difference". The importance of production skills is also reflected in the requirement by the RCSLT to include this in the curriculum [7]. The system provides exercises for the production of individual phones, English words and nonsense words. The choice menus are similar to that in the transcription module. Once an exercise has been generated, students are presented with a screen that allows them to record their pronunciation, listen back to it, and compare it to the pronunciation stored on the system.

3.3. **The administration module**

This (password protected) module allows the administrator to delete, modify or insert recordings, together with transcriptions and attributes for these recordings. Recordings in the system can be searched by means of transcription, or by attribute. For example, a list of all recordings containing the diphthong [eə] can be obtained. Alternatively, a list of recordings by a certain speaker, or containing a certain number of

syllables, or a combination of both can be obtained. The attributes that the system can be searched by depend on the type of recording (i.e. individual phone, English word, nonsense word or short phrase), because the attributes vary for the different types of recordings. Once a single item or a list of items has been obtained, the administrator can either modify some of the attributes or the transcription, or select files for deletion. Inserting new recordings (including transcription and attributes) can be done item by item, or in bulk. Bulk insertion is done per type of recording. It requires the user to select a comma-separated values (CSV) file and a ZIP file containing the MP3 files. The CSV file contains the transcriptions as well as the attributes. Specific instructions for the format of the CSV file are provided during the process.

4. TECHNICAL DETAILS

The system uses a MySQL database that is hosted on an Apache web server, and runs under a Linux operating system. This database contains relevant attributes on a collection of recordings (e.g. in the case of vowels, monophthong or diphthong is one such attribute). The sound files are not stored in the database but in a separate folder structure on the server. Scripting at the server end was done using PHP (for example to handle database queries and search strategies for recordings). However, Javascript was used to handle the user interface functionality (e.g. button functionality). Additionally, the recording facility in the pronunciation module loads the free VIMAS WAV recording applet, which requires installation of the free Java Runtime Environment. The user interface is so far fully functional for Internet Explorer, in combination with Windows XP, Windows Vista and Windows 7. Play back of the sound files is managed by Flash. Transcription answers are displayed in the unicode font Doulos Sil. If this is not installed on the user's machine, transcriptions will appear in the default font Times New Roman.

5. ONGOING AND FUTURE DEVELOPMENT

The system presented here is a work in progress. It is currently being used by the first year students on the speech and language therapy course at Birmingham City University. A formal evaluation by questionnaires is due to be carried out, followed by interviews of a sample of this group of students.

Results of this evaluation will be taken into account in future development of the system. Further developments may also include an expansion of the socio-phonetic variety in the recordings.

6. ACKNOWLEDGMENTS

This work has been supported by grants from the Centre of Excellence in Teaching and Learning and the Centre for Health and Social Care Research from Birmingham City University. The authors would like to thank Helen Jenkins and Jill Clarke as the speakers recorded for the database. We are also grateful to Mark Ashley for his work on the transcription of the recordings.

7. REFERENCES

- [1] Ashby, M., Yanagisawa, K., Kim, Y.S., Maidment, J., Przedlacka, J. 2009. Achieving interactivity in online learning of phonetic skills. *Phonetics Teaching and Learning Conference*, London.
- [2] Hazan, V., van Dommelen, W., Cutugno, F., Fougeron, C., Frid, J., Köster, J., Machuca, M., Mengel, A., Turk, A., Zboril, D. 1997. Phonetics. In Bloothoof, G., van Dommelen, W., Espain, C., Green, P., Hazan, V., Wigforss, E. (eds.), *The Landscape of Future Education in Speech Communication Sciences. 1. Analysis*. Utrecht: OTS publication, 7-50.
- [3] Howard, S. J., Heselwood, B.C. 2002. Learning and teaching phonetic transcription for clinical purposes. *Clinical Linguistics and Phonetics* 16(5), 371-401.
- [4] International Phonetic Association. 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- [5] Ladefoged, P. 1995. Developing phonetic skills. *Proc. 13th ICPhS Stockholm*, 458-461.
- [6] Ladefoged, P. 2006. *A Course in Phonetics*. Boston: Thomson.
- [7] Royal College of Speech and Language Therapists 2010. Guidelines for pre-registration speech and language therapy courses in the UK. http://www.rcslt.org/about/work_with_educators/curriculum_guidelines
- [8] Verhoeven, J., Davey, R. 2007. A multimedia approach to eartraining and IPA transcription. *Proc. Teaching and Learning Conference*, London.
- [9] Wells, J.C. 1982. *Accents of English 1*. Cambridge: Cambridge University Press.
- [10] Wells, J.C. 2000. *Longman Pronunciation Dictionary*. Harlow: Pearson Education.