

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/101888>

Please be advised that this information was generated on 2021-10-20 and may be subject to change.

Catia Cucchiariini, Warda Nejjari and Helmer Strik

# My Pronunciation Coach: Improving English pronunciation with an automatic coach that listens

**Abstract:** Individualized tutoring and feedback by trained language instructors are known to be optimal for language learning. Providing them is time-consuming and costly, however, and therefore not feasible for the majority of language learners. This applies particularly to pronunciation, where corrective feedback should ideally be synchronous, which makes it even more difficult to provide it adequately in the classroom. Recent systems for computer-assisted pronunciation training (CAPT) that make use of automatic speech recognition (ASR) offer new ways of providing tailored feedback on second language pronunciation. In this paper, we present our new project, My Pronunciation Coach, in which we are developing an ASR-based CAPT program that specifically caters to learners of English with Dutch as their mother tongue. The pronunciation coach software uses speech technology algorithms to detect pronunciation errors. Feedback on these errors is given through an interface in an easily understandable manner and remedial exercises are provided accordingly.

---

**Catia Cucchiariini:** P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

E-mail: c.cucchiariini@let.ru.nl

**Warda Nejjari:** P.O. Box 9103, 6500 HD Nijmegen, The Netherlands. E-mail: w.nejjari@let.ru.nl

**Helmer Strik:** P.O. Box 9103, 6500 HD Nijmegen, The Netherlands. E-mail: h.strik@let.ru.nl

---

## 1 Introduction

One-on-one interactive situations in which learners receive individualized corrective feedback are known to be optimal for language learning. However, providing this type of tutoring by trained language instructors is time-consuming and costly, and therefore not feasible for the majority of language learners. This applies particularly to pronunciation, where corrective feedback should be given immediately after the utterance has been spoken, which makes it even more

difficult to provide sufficient practice in the classroom. Recent systems for computer-assisted pronunciation training (CAPT) that make use of automatic speech recognition (ASR) offer new ways of providing tailored feedback on second language (L2) pronunciation.

The usefulness of ASR-based CAPT that provides automatic corrective feedback was demonstrated by our Dutch-CAPT project (Cucchiarini et al. 2009), which addressed pronunciation in Dutch L2. Experiments showed that learners of Dutch appreciated the system. Moreover, the experimental group that used Dutch-CAPT for only four sessions of 30–60 minutes exhibited a significantly larger reduction in the number of pronunciation errors compared to the control group.

In our new project, My Pronunciation Coach (MPC), we are developing an ASR-based CAPT program that specifically caters for learners of English with Dutch as their mother tongue. The pronunciation coach uses speech technology algorithms to detect pronunciation errors. Feedback on these errors is given through an interface in an easily understandable manner and remedial exercises are provided accordingly. Since the system has to cope with English spoken with a whole range of Dutch accents, this is a challenging task requiring innovative technology, developed and optimized for this specific task. MPC is web-based, providing language learners, in particular those at Radboud in'to Languages (Radboud University Language and Communication Centre), the opportunity to improve their pronunciation through regular and sustained practice whenever and wherever they want. Firstly, the MPC project background and objectives will be discussed (Section 2). Secondly, MPC target groups and user demands will be outlined (Section 3) followed by the type of pronunciation education that MPC can offer (Section 4). Thirdly, the typical Dutch pronunciation errors selected for MPC and the research it is based on will be discussed (Sections 5 and 6), followed by our conclusions and future prospects (Section 7).

## 2 The MPC project: background and objectives

In general, people working for international companies and organizations need to have a good command of the English language. There is great demand for professionals who are not only able to make themselves understood, but who can also communicate effectively, fluently and confidently with colleagues and business associates. Even though these new global circumstances have led to more learning of English, pronunciation is often an underemphasized theme in language education. Language teachers are often confronted with a lack of contact hours, and therefore prefer to focus on skills such as grammar, vocabulary, writ-

ing, listening and reading. There are good reasons for paying attention to pronunciation improvement in language teaching programmes, but the main problem seems to be that pronunciation correction requires so much time, feedback and practice, that most of the time it is not feasible in a language classroom, at least not with the necessary amount of individualized attention.

Acquiring a comprehensible pronunciation in an L2 is essential for successful interaction in the L2 for all learners, irrespective of their educational level or career (Abercrombie 1991). This consideration calls for innovative solutions that can make language learning more effective, more personalized, less expensive and less time-consuming. Computer-assisted language learning (CALL) programs have become increasingly popular because they can offer intensive training at any time. In addition, the recent advances in automatic speech recognition (ASR) technology open up new opportunities for developing CALL systems that can address pronunciation, one of the most problematic skills in terms of time investment and cost.

English plays an important role in the daily lives of Dutch people. For example, English is a compulsory language in Dutch secondary education, and an increasing number of university students receive their education almost completely in English in order that they can participate fully in an internationally oriented society. The variety of English taught in Dutch schools is Standard British English, also known as Standard English (SE), and the pronunciation standard is Received Pronunciation (RP). Although Dutch learners of English in general reach a relatively high pronunciation level (Wang and Heuven 2003), they appear to have problems with various aspects of the English sound system and many learners would like to achieve a native-like pronunciation (Koet 2007), because pronunciation is important for successful interaction and social acceptance. These typically Dutch pronunciation difficulties have been documented in a number of textbooks (Gussenhoven and Broeders 1976; Collins and Mees 1981; Gussenhoven and Broeders 1997; Collins and Mees 2003) and pronunciation studies (Haagen 1998; Koet 2007). Some also provide information on which aspects of pronunciation appear to be more problematic and also indicate how a native-like pronunciation could be achieved (Gussenhoven and Broeders 1976; Collins and Mees 1981; Gussenhoven and Broeders 1997; Collins and Mees 2003; Koet 2007).

Other authors have addressed the question of whether a native-like pronunciation should be the objective of pronunciation instruction or whether intelligibility should be considered to be sufficient (Wang and Heuven 2003; Lowie 2004). Although many researchers now agree that achieving a native-like pronunciation may not be necessary for all learners, a reasonably intelligible pronunciation is considered to be an essential component of communicative competence (Abercrombie 1991; Morley 1991; Munro and Derwing 1995; Celce-Murcia et al. 1996).

In addition, a native-like pronunciation is also an important requirement at the higher levels of the Common European Framework of Reference (CEFR, Council of Europe 2001), as indicated below:

- B1 Threshold Pronunciation: Is clearly intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur.
- B2 Vantage: Has acquired a clear, natural, pronunciation and intonation.
- C1 Effective Operational Proficiency: Can vary intonation and place sentence stress correctly in order to express finer shades of meaning.

Furthermore, improving pronunciation and trying to sound as native as possible appears to be an important objective for language learners (Koet 2007).

For the reasons mentioned in Section 1, we decided to use the knowledge and expertise gained in our Dutch-CAPT project to develop a similar system for Dutch learners of English, My Pronunciation Coach (MPC), which language learners can use to practice pronunciation whenever and wherever they want. A feasibility study we carried out as a preliminary to the present project corroborated our assumption that employing such dedicated technology for language learning offers many possibilities, and that there is a large, steadily increasing market potential.

The first line of MPC products we intend to develop are:

1. MPC tech – enabling technology modules, which customers can integrate into their existing courses;
2. MPC Radboud – a complete course with content that language learners can use to practice speaking;
3. MPC framework – a framework that customers can use to develop courses.

### 3 Target group and user demands

In order for the MPC project to serve the learning needs of specific learner groups as well as possible, we defined the following two target groups, each with specific learning needs and user demands:

- A. Pupils in final grades of senior general secondary education and pre-university education in Dutch secondary schools, who are taking English as a compulsory part of their curriculum (aiming at CEFR levels B1 and B2 respectively).
- B. Participants in the language courses Certificate in Advanced English and Certificate of Proficiency in English, taught at the Centre for Language and Communication Radboud in'to Languages (aiming at CEFR levels C1 and C2). These participants are university students enrolled in various programmes, university staff, and highly educated professionals.

A survey was conducted to assess the target group user demands. The 132 respondents representing target group A ( $n = 66$ ) and target group B ( $n = 66$ ) provided important information regarding two variables: (1) need for pronunciation training and (2) demands and wishes.

1. *Need for pronunciation training*: Even though the majority of respondents from both groups are convinced that pronunciation training is crucial to learning a foreign language and consider it at least as important as other aspects of language, such as grammar, vocabulary and sentence structure, it is viewed as problematic. Both groups also indicate that group classes do not offer sufficient opportunity to practise pronunciation. Most respondents, particularly the secondary school pupils (target group A), view pronunciation training as a time-consuming activity. However, when confronted with the possibility of pronunciation training via MPC almost all respondents indicate that this can be a very useful extra tool for pronunciation training.

2. *Demands and wishes*: Both groups indicate that the MPC program must take the level of the user into account and should automatically offer new exercises. Yet, the results also show that target group A considers attractive graphics an important element of an online pronunciation tool; this is indicated by 69% of the pupils (target group A) and 28% of the course participants (target group B). The pupils also attach higher value to a gaming aspect than the course participants: 79% vs. 55%. Target group B also values a time indication for each training exercise, more so than target group A: 73% and 44% respectively.

## 4 Dedicated ASR technology for pronunciation training

Within the framework of My Pronunciation Coach (MPC) we intend to develop computer programs that specifically address pronunciation in a second language and that can be used to support and improve language learning at any time and in any place, and in due course for any language. Dedicated technology will be developed to make this possible. Our approach and available technology make it possible to detect pronunciation errors at the level of individual words and sounds so that appropriate feedback and remedial exercises can be provided to the learner.

MPC is intended for Dutch people learning English, which makes it possible to optimize this system for this language pair: focus on errors made by these learners, optimize the technology for detecting these errors, provide suitable exercises for practising the problematic aspects, etc. The tasks the language learners

have to perform are, for instance, reading a sentence aloud, listening to the sentence produced by the system and then repeating (producing) this sentence, and shadowing (i.e. listening to a sentence, and repeating it while it is produced, with only a short delay). The level of difficulty of these tasks will gradually increase and adapt to the proficiency level of the language learner. For these tasks it is known what the learner should say; however, since what they actually produce could be different, the technology should be able to verify whether the learner was making a serious attempt to produce the utterance in the task or whether (s)he was trying to fool the system. To this end utterance verification algorithms will be employed. In all cases mentioned above, the technology should be able to cope with English spoken with many different Dutch accents at different levels. This is a difficult and challenging task that requires dedicated technology optimized for this specific goal.

## 5 Preliminary selection of pronunciation errors to be addressed in MPC

In developing Dutch-CAPT we formulated a number of criteria for error selection in CAPT systems (Cucchiarini et al. 2009), including error frequency, salience and persistence. In addition, the variety of English selected for MPC is Standard English (SE) with RP as the pronunciation standard. Even though many Dutch people might never achieve a native-like RP accent, the variety that is trained in MPC is SE and RP, since this is the official variety taught in Dutch education. It might not be necessary for all learners to achieve a native-like pronunciation, and MPC can be built around the specific pronunciation needs of each target group, for example using the CEFR or other specialized learning needs.

These criteria were also adopted to carry out a first selection of pronunciation errors to be addressed in MPC. This selection was based on Radboud in'to Languages's teaching experience throughout the Netherlands and research data presented in various studies (Collins and Mees 2003; Gussenhoven and Broeders 1997; Doel 2006; Haagen 1998). The relevance of the selected errors depends not only on the effect mispronunciation can have on intelligibility, but also on the possible negative attitude a Dutch English pronunciation can evoke (Munro and Derwing 1995; Doel 2006; Nejari et al. in press). The errors selected on the basis of Doel's error hierarchy (2006: 292) and Gussenhoven and Broeders's substitution tables (1997: 113, 171) are presented in this section.

The Dutch English vowel pronunciation errors selected are displayed in Table 1, where column 2 shows the RP vowels selected, followed by the condition or the

	RP	Condition	Dutch	Example
1	/ɪə/	before /r/	/i/ tien 'ten'	beer, idea
2	/æ/	+fortis consonant +lenis consonant	/ɛ/ pet 'cap' /ɛ:/ serre 'conservatory'	bat bad
3	/ʌ/	spelling with o spelling with u	/ɔ/ sok 'sock' /ʏ/ bus 'bus' /ə/ bedachtzaam 'thoughtful'	other bus unwise
4	/u:/		/u/ soep 'soep'	soup
5	/ʊ/		/u/ goed 'good'	good

**Table 1:** Vowel errors in Dutch English pronunciation

context in which the error occurs. Column 4 describes the Dutch vowel used to replace the original RP vowel, together with an example of a Dutch word and its English translation. Examples of words with the selected RP vowels are shown in the right-hand column. If no condition is specified, then the error can be applied to all conditions. For consonants, pronunciation errors are shown in Table 2, which is organized in the same way as Table 1.

## 6 Data-driven investigations of pronunciation errors by Dutch learners of English

Tables 1 and 2 show a preliminary selection of errors, to be addressed in MPC, based on existing literature and teacher experience. This can be considered a top-down, knowledge-based approach to error selection, which is extremely useful to obtain an initial idea of common errors and provide information on the relevance of the errors in terms of salience, persistence, and stigmatization. However, a knowledge-based approach can provide only an impression of the frequency of each error. For more objective and precise information on this point a bottom-up, data-driven approach is required.

To develop systems like the one envisaged in MPC, both approaches are necessary. In particular, data-driven studies are required to obtain objective and quantitative data on the frequency of occurrence and the contexts of the various errors in different learner target groups. The information obtained from these investigations can in turn be used to develop error detection algorithms and to design useful pronunciation exercises for the different target groups, which can



	RP	Condition	Dutch	Example
1	/b/	word-final	/p/ Rob	hub
2	/d/	word-final	/t/ bad 'bath'	bad
3	/g/	word-final	/k/ lik 'lick'	big
4	[p <sup>h</sup> ] [t <sup>h</sup> ] [k <sup>h</sup> ]	initial voiceless plosives	/p/ pak 'suit' /t/ tak 'branch' /k/ kat 'cat'	pack tap cap
5	/tʃ/		/ʃ/ sjaal 'scarf'	chips
6	/dʒ/		/tʃ/	bridge
7	/dʒ/		/ʃ/ sjaal 'scarf'	jam
8	/w/		/v/ wie 'who'	wine
9	/ð/		/d/ dak 'roof'	the, this
10	/θ/		/s/ sap 'juice'	booth
11	/θ/		/s/ sap 'juice' /t/ tap 'tap'	thirty, three
12	/s/		/ʃ/ sjaal 'scarf'	socks
13	/z/	word-final	/s/ sap 'juice'	jazz

**Table 2:** Consonant errors in Dutch English pronunciation

vary from highly proficient university students of English to secondary school pupils.

To this end, a first data-driven investigation was carried out using speech recordings of Dutch students of English that had been made at Radboud University Nijmegen. The material consisted of short stories read aloud by Dutch students of English, which are used to test the pronunciation level of the students in their first and second year at the Department of English Language and Culture of Radboud University Nijmegen. There were seven different stories varying in length from 200 to 350 words. They were specially selected for the purpose of pronunciation testing. This material can be described as found data, since it was not collected specifically for the purpose of the MPC project, but for another purpose, namely testing pronunciation for a university English course. In total, this database contains speech recordings of 226 students reading 617 stories. The average number of words is 257 and the average length is 98 seconds. The total number of phonemes is about 520,000 (200,000 vowels and 320,000 consonants).

A detailed description of the results of this investigation is provided in Cucchiarini et al. (2011), and will not be reiterated here, since it is beyond the scope of the present paper. However, it is important to underline that this study generally confirmed the expectations presented in Section 5 and provided some additional useful findings concerning the frequency of occurrence of certain errors. In addition, it revealed errors that were not included in the preliminary selection, but that are apparently made even by such advanced students. It seems that these errors should also be addressed in MPC, although further analysis is obviously required to gain more insight into the nature of these errors and the specific contexts in which they occur.

An important result is that in this group of learners the frequency of pronunciation errors is relatively low: about 1.58% of the vowels and 1.82% of the consonants were mispronounced in the full data set. These percentages may appear to be extremely low, and they probably are, but this does not necessarily imply that pronunciation training is completely superfluous in such cases. Especially these learners, who want to sound as native as possible, strive to eliminate any traces of non-nativeness from their speech, including errors that may not be very frequent, but that might give them away as non-natives.

In the meantime, speech recordings have been made of a second target group, viz. secondary school pupils. For this group of learners we expect pronunciation error frequency to be higher than for the university students. Contrary to the speech recordings of university students, which had been made for a different purpose, the speech database of secondary school pupils has been collected especially for the purpose of MPC and the speech material has been selected accordingly. The secondary school pupils have to read aloud words and sentence structures (statements, questions, commands, names) that are commonly used and measure the phonetic and prosodic abilities of the speakers. Examples of these common words are: *the Netherlands, brother, sister, language, classmate, airport, terrible, naughty, be really into* and *horse riding*. The following examples show the types of sentence structures that were recorded:

1. *My name is Josh Roberts.*
2. *If you want, you can leave a message on my blog.*
3. *Where is that hat?*
4. *What I think is most important about family is that they share in your life, not only by doing things together, but also by simply asking how your day was.*

In addition, the secondary school pupils need to supply the recording system with personal information relating to their language background. One session comprises approximately 150 utterances and lasts about 30 minutes.

## 7 Conclusions and future prospects

In this paper we have explained the aim and rationale of the MPC project and have reported on part of the work that has been carried out so far. We have seen that the database, created on the basis of the “found speech” available at the Department of English Language and Culture and the annotations made by the transcribers, constitutes a good starting point for MPC. This database can be used not only for analysing pronunciation errors, as we have done in Cucchiarini et al. (2011), but can also be employed for training the speech recognizer and the algorithms for the detection of pronunciation errors, which are required for the MPC system.

Since we observed that the students recorded in this database make relatively few mistakes, using a relatively limited set of English phonemes, we realized that new speech recordings had to be made in order to develop versions of MPC for target groups with lower pronunciation quality than the university students. For this purpose we contacted a number of schools and thus far managed to make recordings of 130 pupils. These recordings will also serve a dual function, a) as an important knowledge source for studying the pronunciation errors made by secondary school pupils, and b) as training material for the speech recognizer.

The information obtained from these speech databases on pronunciation errors made by the different target groups will in turn constitute the basis for developing tailored pronunciation exercises in MPC. We will try to develop exercises that focus on the specific difficulties of each target group or learner.

To summarize, the work carried out so far has provided valuable language resources and useful information that can be used, not only to develop and improve the MPC system, but also for other research purposes and applications. In the near future we intend to extend the MPC system in different directions. We would like to address aspects of speaking proficiency other than pronunciation, such as morphology, syntax, and vocabulary, and would like to include L1–L2 combinations other than Dutch-English.

## References

- Abercrombie, David. 1991 [1956]. Teaching pronunciation. In Adam Brown (ed.), *Teaching English pronunciation: A book of readings*, 87–95. London: Routledge.
- Celce-Murcia, Marianne, Donna M. Brinton & Janet M. Goodwin. 1996. *Teaching Pronunciation*. Cambridge: Cambridge University Press.

- Cucchiari, Catia, Ambra Neri & Helmer Strik. 2009. Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication* 51(10). 853–863.
- Cucchiari, Catia, Henk van den Heuvel, Eric Sanders & Helmer Strik. 2011. Error selection for ASR-based English pronunciation training in “My Pronunciation Coach”, Proceedings Interspeech 2011, Florence, Italy.
- Collins, Beverly & Inger Mees. 1981. *The sounds of English and Dutch*. The Hague: Leiden University Press.
- Collins, Beverly & Inger Mees. 2003. *The phonetics of English and Dutch* (5th edition). Leiden: Brill.
- Doel, Rias van den. 2006. *How friendly are the natives? An evaluation of native-speaker judgments of foreign-accented British and American English*. Utrecht: Lot PhD Dissertation.
- Gussenhoven, Carlos & Ton Broeders. 1976. *The pronunciation of English: A course for Dutch learners*. Groningen: Wolters-Noordhoff-Longman.
- Gussenhoven, Carlos & Ton Broeders. 1997. *English pronunciation for student teachers* (2nd edition). Groningen: Wolters-Noordhoff.
- Haagen, Monique van der. 1998. *Caught between norms: The English pronunciation of Dutch Learners*. The Hague: Holland Academic Graphics PhD Dissertation.
- Koet, Ton. 2007. *Polder English in Dutch ears*. Utrecht/Amsterdam: PhD Dissertation.
- Lowie, Wander. 2004. De zin en onzin van uitspraakonderwijs [The sense and nonsense of pronunciation education]. *Levende Talen* 5(1). 3–12. <http://www.ltprojecten.nl/sites/default/files/200702120853089197.pdf> (accessed 4 December 2011).
- Morley, Joan. 1991. The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly* 25(3). 481–520.
- Munro, Murray J. & Tracey M. Derwing. 1995. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning* 45(1). 73–97.
- Nejjari, Warda, Marinel Gerritsen, Hubert Korzilius & Monique van der Haagen. 2012. Reactions to Dutch accented English. *World Englishes*.
- Wang, Hongyan & Vincent J. van Heuven. 2003. Mutual intelligibility of Chinese, Dutch and American speakers of English. *Linguistics in the Netherlands 2003*. 213–224.

## Acknowledgements

The MPC project is funded through an STW Valorisation grant from the Netherlands Organisation for Scientific Research (NWO), no. 11795. We are indebted to Liesbet Korebrits, Michiel Sanders, Edith Schouten, and Simone Barnhoorn of Radboud in'to Languages, the language and communication centre of Radboud University Nijmegen, and to Henk van den Heuvel and Eric Sanders of the Centre for Language and Speech Technology (CLST). Finally, we thank the Department of English Language and Culture of Radboud University Nijmegen, in particular Monique van der Haagen and Hedi Kamara, for providing the speech material employed in this study.

## Bionotes

*Catia Cucchiarini* is a senior Researcher in Speech Sciences and Speech Technology, Centre for Language and Speech Technology, Radboud University Nijmegen, The Netherlands.

*Warda Nejjari* is a teacher of English and International Business Communication, Radboud in'to Languages, Radboud University Nijmegen, The Netherlands.

*Helmer Strik* is Associate Professor of Language and Speech Technology, Department of Linguistics, Radboud University Nijmegen, The Netherlands.