THE YOUTUBE REVOLUTION: ENGAGEMENT, PERCEPTION AND IDENTITY

Joanna Smith

Unitec Institute of Technology, New Zealand
jsmith@unitec.ac.nz

ABSTRACT

Advances in technology and social media have caused an avalanche of possible texts for use in pronunciation teaching. This article explores how the use of short digital videos can revolutionize the teaching of pronunciation in a second or foreign language. Their use is first examined pedagogically in relation to three areas: student engagement, speech perception and second language identity. Then, examples of use, student recordings and student perceptions are presented, which illustrate their value. It is suggested that appropriate, informed use of short digital videos has the potential to accelerate the acquisition of the target pronunciation(s).

Keywords: identity, speech perception, language acquisition, digital natives, technology

1. INTRODUCTION

Students who fit Prensky's description of 'digital natives' [14] now demand hands-on activities, graphics before texts and information at 'twitch speed'. For these students, watching videos on mobile phones and sharing them on social networking sites is an everyday occurrence. For them, working with sound files that the teacher has chosen, and is in control of, can be rather boring, and potentially even inappropriate.

While the use of video in the language classroom is not new, the prepared language teacher can tap into the social/technological phenomenon of short digital videos, and make the most of these readily available texts to create what could be considered transformative classroom materials.

This article will first present research that demonstrates how videos increase student engagement, allow for improved speech perception, and create opportunities for new identities to be imagined. Some examples of use are presented, along with a sample of student perceptions about learning pronunciation this way.

2. STUDENT ENGAGEMENT

By exploiting popular digital videos, the teacher not only enters the students' digital world, but also more fully engages them in class.

2.1. Videos and the brain

Berk [2] presents research describing the use of videos and the brain. Videos have been found to benefit students by connecting to multiple intelligences, both hemispheres of the brain, and to the emotional sense of the students. He also suggests that the use of multimedia enhances learners' ability to remember content. He suggests 20 learning outcomes that can be achieved through the use of videos in lessons, including grabbing students' attention, improving student attitudes towards content and learning, and inspiring and motivating students.

2.2. Range of texts

While general increased engagement is desirable in any classroom, for the purposes of learning pronunciation, videos can be exploited for examples of specific phonetic features.

Any video with speech in the target language can be used. Humorous materials provide light relief in between theoretical input, current events clips provide high motivation for listening, and videos capturing real life conversations can be recycled for teaching other aspects of the language, such as pragmatics, the teaching of which relies heavily on having authentic texts [17].

In order to cater for a range of tastes and preferences, students can be encouraged to bring their own texts to class, or send them to the teacher beforehand, and the teacher (or students) can find relevant examples to illustrate the teaching content. In this way, pronunciation lessons can be customized to any mix of students.

3. SPEECH PERCEPTION

Phonetics teaching has always been enhanced by the use of visual cues. More than three decades ago, Brown [3] suggested 'Since foreign learners need all the help they can get in the form of aids to interpreting the spoken message it seems worthwhile drawing to their attention what visual clues to segmental quality they may expect to find in informal speech.' (p. 79). Research in the area of speech perception [8, 9] suggests that visual cues not only enhance, but significantly contribute to understanding, further confirming that audio-visual materials are superior to audio only.

3.1. Motor Theory of speech perception

The motor theory of speech perception [5] suggests that perceiving speech is perceiving gestures, and the motor system is recruited for perceiving speech. Brain-imaging studies [15] show excitation across many different regions of the brain, especially when the talker's face is clear. Perceiving other people's behaviors (here, verbal behavior) triggers imitative motor plans in the perceiver [13, 16]. In other words, during observation, observers form mental representations of how their own bodies would perform those same movements. This mental representation aids in understanding the speech we perceive.

3.2. Perceiving a foreign language

While much of this research deals with speech perception in the first language, some research suggests that second language speakers rely even more heavily on visual cues in the presence of ambiguous audio messages [6]. Further, [1] shows that perception in an unfamiliar accent can improve after production training in that accent.

In light of this research, it seems that the use of audio-only texts denies learners valuable input. This visual information can be easily supplied with the use of video, particularly ones in which speakers' faces are clearly visible, and the task of finding such material is now much easier.

4. IDENTITY

Because accent and identity are so deeply intertwined, many learners have been reluctant (often subconsciously) to change their accent, as there is always a 'choice between the reward of being efficient and the cost of not marking [one's L1] identity' [7]. Teachers, too, have often felt reluctant to teach pronunciation, for fear of

'imposing' an unwanted identity on a learner, or even, specifically in the case of English, 'perpetuating Western imperialism in different parts of the world' [12].

4.1. Allowing choice

One answer to this issue is to suggest that learners themselves need to be responsible for choosing their target accent. The wide range of easily accessible video clips now allows for use of model speakers with a variety of genders, ethnicities, ages and social statuses. The teacher's job can be re-conceptualized as facilitator, empowering students to make an informed choice, and showing learners how to emulate their chosen models. Any reluctance to teach pronunciation due to fear of imposition is now irrelevant.

4.2. Preserving context

In audio-visual texts, speakers' body language, their clothing, the physical context and the reactions of others towards them are all preserved. It is much easier for a learner to judge whether or not they wish to emulate that identity if they can see what a speaker with that accent is like.

4.3. Multiple models

In fact, personal identity is not constant, but rather multiple and dynamic. Just as teachers can promote 'additive bilingualism', which adds, rather than detracts from a learner's identity, learners can be encouraged to experiment with a variety of different target accents, developing multiple L2 identities, and therefore becoming empowered to code switch as necessary in different sociolinguistic contexts.

This is what Jenkins [11] describes as 'the possibility of approaching [non-NS pronunciation norms] as identity-driven, with speakers being aware of what an NS of English would do at a given point, but choosing to do something else, perhaps in order to signal a shared identity with a particular NNS interlocutor.' (p. 202).

4.4. Motivation

Dörnyei's 'L2 Motivational Self System' [4], based on psychological research, outlines three 'possible selves', which can act as guides for learners. The 'ideal self' and 'ought self' can be powerful determinants for learner outcomes, given the right circumstances. One of the requirements for this to be the case is the availability of an

elaborate and vivid future self-image. While much has been written on the role of imagination and mental imaging for this image to emerge, 'the use of role models seen in films or TV or in real life' is a common route to 'igniting the vision' (p. 33).

In other words, when presented with appealing, realistic, real-life models, in the package of vivid audio-visual texts, students can more easily construct their own 'ideal self', which can in turn provide higher motivation for language acquisition.

The number of views, and the 'like' status on YouTube clips can thus become a teacher's guide, and better enable a teacher to 'hit the nail on the head', in presenting possible 'ideal selves'.

5. EXAMPLES AND STUDENT PERCEPTIONS

Pronunciation work is often viewed as a slow process. Studies on the development of L2 vowels, such as [10] illustrate the need to view second language phonological acquisition as a long-term project.

The author suggests, however, that the use of relevant, highly motivating audio-visual texts in the classroom can provide a short cut, empowering students to make dramatic changes to their pronunciation in the shorter term. Below, two videos and their classroom use are described, and student recordings of mimicking these texts are presented to illustrate this idea.

5.1. Examples of digital videos used

The following examples of use come from an advanced English class in a New Zealand tertiary institution, where students mostly had an overall IELTS equivalent of about 6.5. There were 38 students in the class, and the teacher tried to incorporate at least one digital video in each class that semester. The videos were used to illustrate a range of spoken English features, including pronunciation.

5.1.1. Friends, series 9, episode 4

The teacher assumed that many students would have seen this popular TV series, and would be familiar with the characters. This particular scene was chosen initially to illustrate the use of sarcasm, including the accompanying body language. The text was later 'recycled' for a lesson on intonation, as it demonstrated a wide range of pitch. Students were invited to mimic the monologue.

5.1.2. BBC News item about Moko the dolphin

This clip was chosen for its current events and local interest value. The British accent was discussed by analyzing the pronunciation of a few key words, and the students used the text to find examples of weak forms, word linking and pausing. Students were invited to mimic the text.

5.1.3. Student recordings of these texts

Recordings from four students are included below. These recordings were chosen in order to represent a range of abilities, genders, ages and first languages in the class. There are two males and two females, their ages range from early twenties to mid-fifties, and their countries of origin are Korea, Indonesia, China and India respectively.

Student A *Friends*: audio file 1; *BBC*: audio file 2
Student B *Friends*: audio file 3; *BBC*: audio file 4
Student C *Friends*: audio file 5; *BBC*: audio file 6
Student D *Friends*: audio file 7; *BBC*: audio file 8

As can be heard, all of the students display a surprising flexibility in producing quite a different 'voice' for each of these mimicking tasks.

5.1.4. Free speech versus imitation

The texts above are not free-speech samples, and the question of transfer remains. However, the recordings nevertheless illustrate the plasticity of the learners' L2 phonological system. The author suggests that by demonstrating the ability to produce varying target pronunciations, students are showing an ability to accommodate, given the right sociolinguistic circumstances.

5.2. Student perceptions of video use

At the end of the semester, students were presented with an informal survey questionnaire to determine their perceptions of the use of YouTube clips in the class. Most of the questions were open-ended, for example "What was your favorite YouTube video from our class, and why?" and "What do you think are the advantages and disadvantages of using videos for learning speaking?" 20 students returned the survey. Some student comments relevant to the three themes of this article are presented below.

5.2.1. Student engagement

'Gives a break from the tutor talking all the time, excitement because I am watching what had

actually happened, keeps the mind awake and interesting.'

'More interesting than radio'

5.2.2. Speech perception

'Advantage is making the learning speaking vividly and easily'

'I can feel in my mind how people say and some voice could leave in my mind'

5.2.3. Identity

'Videos can provide more background information and can teach us facial expression and body language which are very important in a authentic conversation.'

'I think [interesting videos] are important because if the students really really enjoy the videos, they will never stop copying the way the speakers talk.'

5.3. Student video preference

From the 20 responses, eight different 'favorite videos' were mentioned, with reasons ranging from specific items of learning (9), humor (6), challenge (4), emotion (3) and familiarity (1). This is more evidence for the benefit of using a wide range of videos in a diverse class.

5.4. Student perceptions of improvement

In response to the prompt about whether they thought their speaking skills had improved as a result of spending time working with YouTube videos, 17 out of 20 students indicated an improvement and then elaborated how. While such improvements were not quantified, this can still be considered a positive result, even if only a measure of student satisfaction.

6. CONCLUSION

An unprecedented array of pronunciation models has become available in the form of digital videos, which learners in the net generation watch and share regularly. For reasons of increased student engagement, enhanced speech perception and vivid potential second language identities, pronunciation teachers are encouraged to make use of this valuable resource. The few examples presented here show that informed use of these videos may even accelerate L2 phonological development. Further research is required to quantify such claims, but student indications of satisfaction point to the benefit of teachers joining the YouTube revolution.

7. REFERENCES

[1] Adank, P., Hagoort, P., Bekkering, H. 2010. Imitation improves language comprehension. *Psychological Science* 21(12), 1903-1909.

[2] Berk, R.A. 2009. Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. *International Journal of Technology in Teaching and Learning* 5(1), 1-21.

[3] Brown, G. 1977. *Listening to Spoken English.* London: Longman.

[4] Dörnyei, Z. 2009. *Motivation, Language Identity and the L2 Self.* Bristol: Multilingual Matters.

[5] Galantucci, B., Fowler, C., Turvey. M. 2006. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review* 13(3), 361-377.

[6] Ganström, B., House, D., Beskow, J., Lundeberg, M. 2001. Verbal and visual prosody in multimodal speech perception. *Proc Nordic Prosody 2000,* Trondheim. Frankfurt am Main: Peter Lang, 77-88.

[7] Gatbonton, E., Trofimovich, P., Magid, M. 2005. Learners' ethnic group afficliation and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly* 39(3), 489-511.

[8] Grant, K., Tufts, J., Greenberg, S. 2007. Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impared individuals. *J. Acoust. Soc. Am.* 121(2) 1164-1176.

[9] Greenberg, S. 2007. What makes speech stick? *Proc. 16th ICPhS,* Saarbrücken. *http://www.icphs2007.de/ conference/Papers/1760/1760.pdf*

[10] Gulinello, F. 2010. A longitudinal investigation of vowel acquisition. In Levis, J., LeVelle, K. (eds.), *Proc. 1st pron. In Second Lang. Learning and Teaching Conf* Iowa State University, 2009. 90-104.

[11] Jenkins, J. 2007. *English as a Lingua France: Attitude and Identity.* Oxford: OUP.

[12] Norton, B. 1997. Language, identity and the ownership of English. *TESOL Quarterly* 31(3), 409-429.

[13] Pickering, M., Garrod, S. 2007. Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences* 11(3) 105-110.

[14] Prensky, M. 2001. Digital natives, digital immigrants. *On the Horizon.* MCB University press. 9(5). *http://www.marcprensky.com/writing/*

[15] van Wassenhove, V. 2007. Analysis-by-synthesis in auditory-visual speech perception: Multi-sensory motor interfacing. *Proc. 16th ICPhS,* Saarbrücken. *http://www.icphs2007.de/conference/Papers/1763/1763.pdf*

[16] Wilson, M., Knoblich, G. 2005. The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131(3), 460-473.

[17] Yates, L. 2004. The 'secret rules of language': Tackling pragmatics in the classroom. *Prospect* 19(1), 3-21.