Fernando Rubio

9 Teaching Pronunciation and Comprehensibility in a Language MOOC

Abstract: This chapter reports on a study comparing the gains in comprehensibility of students enrolled in a traditional face-to-face (F2F) pronunciation course with those enrolled in a Language MOOC (LMOOC). The study analyzes potential correlations between types of feedback and pronunciation gains. The two courses followed a similar structure and had comparable goals, but differed crucially in the types and amount of feedback provided to learners. Students produced a sample of controlled speech at the beginning and the end of each course. The samples were judged by two native speakers that assigned them comprehensibility ratings. Results showed significant improvements in both course formats, but larger effect sizes in the case of the LMOOC. The findings are discussed taking into consideration the different affordances of the two learning formats.

Keywords: MOOC, oral, feedback, pronunciation

9.1 Introduction

Since their inception, and particularly with the increase in their popularity since 2011, Massive Open Online Courses (MOOCs) have been touted by many as a disruptive force in education. The expectation, as Friedman (2012) put it, was that through MOOCs anyone with access to a computer and a fast Internet connection could enjoy the benefits of high quality education at almost no cost. The proliferation of free and massively enrolled courses offered by some of the world's top universities brought increased media attention to MOOCs to the point where the New York Times declared 2012 the “Year of the MOOC” (Pappano, 2012).

Our profession was initially very slow in embracing the concept of massive courses. The caveats were both intellectual and methodological. On the one hand, as members of the Humanities used to dealing with small groups of learners in brick-and-mortar academic settings, many language teaching professionals were skeptical of the feasibility of engaging large numbers of students in effective learning experiences without the one-on-one and small-group types of human interaction that are typical of humanities disciplines. Marianne Hirsch, President of the Modern Language Association and professor of English at Columbia University, expressed this concern in a letter to the editor of the New York Times responding to a column on MOOCs published by Thomas Friedman: “If we want a better-educated global citizenry, should we not foster a multitude of professors with different views who can

share deep critical thinking in a community of learners *such as only the embodied experience of the classroom can yield* [emphasis added]?” (Online Courses, 2013). The prevailing assumption in our field is still that face-to-face (F2F) teaching is the golden standard and anything other than traditional classroom teaching will result in an inferior experience for learners.

But language-teaching professionals are more concerned about the practical and methodological implications of teaching language online to very large numbers of learners. If we accept the assumption that some of the crucial ingredients for language acquisition are ample opportunities to interact synchronously with other learners and with experts (native speakers or instructors) and, in the case of instructed learning, access to corrective feedback, the promise of scalability that is essential to the appeal of MOOCs runs counter to their plausibility as a valid format for language learning.

A crucial question asked by many in our profession is whether Language MOOCs (LMOOCs) are a viable option when the goal of a course is to elicit ample oral production in the target language facilitated by frequent and targeted feedback. For example, courses that focus on improving pronunciation, like the ones described in this chapter, are not seen as the most logical candidates for success in this model. In order to advance our understanding of the potential or the limitations of LMOOCs, the present chapter reports on a study of comprehensibility and the role of feedback in oral production in the LMOOC “Improving your Spanish pronunciation”, taught between January and March of 2103. The study looks at the effectiveness of the LMOOC compared to a traditional F2F course by looking at the gains in comprehensibility experienced by students enrolled in both courses and analyzing the potential relationship between types of feedback and course effectiveness. The following sections provide an overview of relevant literature and a description of the methodology employed in the study.

9.2 The Role of Feedback

Gass, Behny and Plonsky (2013) offer a good working definition of feedback as “An intervention that provides information to a learner that a prior utterance is correct or incorrect” (p. 524). Feedback has occupied a prominent role in Second Language Acquisition (SLA) research for the past 40 years, especially following Hendrickson’s (1978) review of the issue of error correction in the classroom. The questions asked in the literature have to do with whether errors need to be corrected and, if so, which errors, by whom, when and how. There has been considerable discussion as to whether feedback or, at least, corrective feedback, is necessary for acquisition (Krashen, 1981; Truscott 1999). However, most research provides empirical evidence that in the case of adults and for certain features, being exposed to negative evidence

is indeed essential for acquisition (see Loewen, 2012; Long, 2007 for detailed reviews of the issue).

Perhaps more relevant to the study presented in this chapter is the discussion about *how* and *when*; that is, the specific qualities and contexts that make certain types of feedback more effective than others. Different types of feedback can vary in terms of their degree of explicitness. Explicit feedback provides an open signal that an error has been made, while implicit feedback does not. While some research points to the preponderance of implicit feedback in language classrooms and to its potential advantages as a non-intrusive method (Long, 1996; 2007), most of the literature emphasizes the advantages of explicit forms of feedback and often criticize the unsystematic and arbitrary nature of implicit approaches (Carroll, 2001; Ellis, 2007; Ellis, Loewen & Erlam, 2006; Lyster & Ranta, 1997; Spada & Lightbown, 1999).

9.2.1 Feedback and Pronunciation Instruction

In the case of pronunciation instruction, the research on pedagogical approaches to date does not paint a clear picture. Typically, instruction in Phonetics in a Second Language (L2) context consists of a combination of explicit information about the L2's phonological system—including the most salient differences between the First Language (L1) and the L2—ample opportunities to practice, and a good dose of feedback. Different approaches to instruction, with the resulting differences in types of feedback, seem to be effective depending on the context. For example, Celce–Murcia, Brinton, and Goodwin (1996) argue that an implicit instructional approach that focuses on imitation is more appropriate at lower levels of proficiency, while more explicit instruction that focuses on the analysis of phonological features may be more effective for learners at higher levels of proficiency. Some authors advocate for integrating pronunciation instruction into an overall communicative focus without making it a separate, explicit target of instruction (Isaacs, 2009), and other research suggests that instruction in L2 Phonetics does not have a positive effect on learners' pronunciation (Kissling, 2013). However, still others argue that implicit instruction alone is not sufficient. For example, several recent studies (Chung, 2007; Dłaska & Krekeler, 2013; Saito & Lyster, 2012) have found that, in addition to implicit feedback, the presence of explicit, individualized corrective feedback has an added beneficial effect on the acquisition of pronunciation. In a study on the acquisition of the English retroflex /ɹ/ by Japanese speakers, Saito and Lyster (2012) found that those who received explicit individualized corrective feedback significantly outperformed those who did not. Dłaska and Krekeler (2013) studied the immediate effects of instruction on the level of comprehensibility of adult learners of German and found that those who had received individualized corrective feedback made significantly more improvements.

As is frequently the case, the different research design and methodologies employed and the presence of potentially confounding variables makes it difficult

to draw definitive conclusions from the existing research on pronunciation instruction. The mixed results may also have to do with the fact that the effectiveness of a teaching approach depends, in part, on the phonological feature being studied (see Kissling, 2013, for a review). Yet, there seems to be strong agreement that pronunciation instruction has a beneficial effect on adult learners' comprehensibility, intelligibility and accentedness (de Bot & Mailfert, 1982; Kennedy, Blanchet & Trofimovich, 2014; Saito & Lyster, 2012; Trofimovich, Ammar & Gatbonton, 2008). The question that remains is exactly which components of a pronunciation course can be identified as being beneficial in which contexts. Until that question can be unequivocally answered, the only certainty is that learners benefit from exposure to ample feedback of a variety of kinds from a variety of sources. However providing adequate feedback is challenging in pronunciation instruction, among other things, because explicit, individualized and teacher-generated feedback, which seems to be a requirement for improvement, is notoriously time-consuming. Can the affordances of an LMOOC facilitate this challenge?

9.2.2 Feedback and Online Affordances

Hattie and Timperley (2007) summarize the qualities of effective feedback as follows:

Effective feedback must answer three major questions asked by a teacher and/or by a student: Where am I going? (What are the goals?), How am I doing? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?) (p. 86)

According to this view, an effective pronunciation course, online or F2F, should provide learners with appropriate models of correct pronunciation (*where am I going?*), opportunities to judge their own pronunciation (*how am I doing?*), and the necessary information to address any existing problems (*where to next?*). Since MOOCs and, particularly LMOOCs, are a relatively new phenomenon, many in our profession are still wondering whether massive language courses can provide the necessary affordances to facilitate effective feedback of all three kinds to the same or, perhaps, a larger degree than would be possible in an F2F classroom.

The issue of online affordances has received a significant amount of attention recently. In general, the particular qualities of online technologies have been linked directly to the open and participatory nature of Web 2.0. Discussing online learning theory, Anderson (2004) claims that “the greatest affordance of the web for educational use is the profound and multifaceted increase in communication and interaction capability” (p. 42). Of course, the fact that a particular technology provides certain opportunities for learning is no guarantee that the learner will take advantage of them. Ultimately, it is the instructional design of the course, not the technology

or the opportunities that it provides that will have an impact on the educational experience.

In MOOCs in particular, the massive nature of the course represents two sides of the same coin by creating both a challenge and an affordance. The challenge of assessing and providing feedback to large numbers of learners has been discussed in the recent literature (Clougherty & Popova, 2013) and is typically addressed taking advantage of the other side of the coin: giving peers an increased role in the feedback and assessment process. The LMOOC described in this chapter was designed so that students would have unlimited opportunities to access models of correct pronunciation through digital recordings with examples of native speaker talk that were available to them throughout the course. Since all their pronunciation practice was recorded and archived, they also had the chance to judge their own pronunciation and compare it with the models and thus engage in self-evaluation as often as they wanted. Of course, these are not affordances that are exclusive to an LMOOC or even to an online context. What is crucially distinctive in an online course and, particularly, in a massive one, is the access to feedback. The native speaker models and the chance to self-evaluate repeatedly would be of little use without the proper guidance. In order to bridge the gap between their actual performance and their target, learners need to be given specific, individualized and frequent feedback on their performance. In this course, learners received individualized feedback on all their recordings from an expert (the instructor or one of two assistants) and/or from a peer.

9.3 Structure of the Courses

This study compares students in two versions of a Spanish Pronunciation and Phonetics course: a traditional F2F course and an LMOOC both taught by the researcher. The two courses were taught following a very comparable structure. Both included a similar component of explicit phonetic instruction, although the theoretical component was more prominent in the F2F course. Access to recorded models of native speaker pronunciation of target phones and supra-segmental features, lots of opportunities to practice, and frequent feedback were common features of the two course formats.

The F2F course was taught as a 15-week course with typical enrollments that ranged between 25 and 35 students. Classes met twice a week for 1 hour and 20 minutes. Class time was normally spent on a discussion of the readings and some practice listening to and analyzing native-speaker as well as learner speech. The F2F course placed more emphasis on the Phonology component of the course at the beginning of the semester and then moved to application and practice during the second half. Also, since the vast majority of students enrolled in the course were native speakers of English, a significant amount of time was devoted to the analysis and practice of features of the Spanish system that are typically challenging for English

speakers. The course used a Learning Management System (LMS) that allowed students to access recordings of native speaker pronunciation and to record their own speech. The main difference with the LMOOC was the amount and type of feedback in the F2F course. There was plenty of feedback of the first type identified by Hattie and Timperley (2007): *what are the goals?* This happened during in-class analysis of examples of learner pronunciation and during discussions of the main differences in the phonological systems of English and Spanish. This feedback was provided to the group as a whole. There was a limited amount of individualized corrective feedback. During the course of the semester, students had to do four recordings of their own pronunciation that were uploaded to the LMS. They received a grading with specific feedback on their performance and suggestions for improvement.

The LMOOC “Improving your Spanish Pronunciation” was offered for six weeks starting in January 2013. It was one of the courses in the first round of MOOCs offered by Canvas Network, a massive open online course platform developed by Instructure (<http://www.instructure.com>). All courses included in this first offering were capped at 500 students. There was no cost to participate in the course and it was open to anyone with an elementary knowledge of the language. This course was taught by the author with the help of two assistants, both native speakers of Spanish. This LMOOC was organized as a series of twelve individual modules that covered the basics of Phonetics and Phonology and the specifics of the Spanish phonological system. Each module started with an introduction to the topic using a series of short videos followed by brief quizzes that were automatically graded. Students had to view the videos and complete the quizzes with a minimum score in order to advance to the application component of the module. The practice section of the modules started with audio files of native speakers reading fragments of text that included the phonological features emphasized in each module. Then students were asked to record themselves reading the same sentences or paragraphs and to evaluate themselves using a rubric provided. The free audio editor and recorder Audacity was used to record all the audio files in the course. One of the advantages of this software is that it allows dubbing over existing tracks. This way, students were able to listen to a native speaker’s pronunciation and record their own voice on the same track. When listening to recordings, students were also able to visualize the speech patterns by looking at a spectrogram and compare their own production to the native speaker’s. This practice section of the module was not graded and students could record themselves as often as they wanted. As an element of pronunciation instruction, self-assessment has been identified as a useful tool that can enhance learners’ awareness of their performance, increase their motivation and make the assessment process more learner-centered (Dlaska and Krekeler, 2008).

Students then had to listen to peers’ recordings and provide feedback using the same rubric they used to evaluate themselves. Figure 1 shows an example of one of these rubrics. Each student was required to provide feedback to a minimum of two peers. Finally, each module ended with an assignment for students to record a

paragraph or a series of sentences that included the features practiced in the module. This recording received a grade assigned by the instructors following a rubric shared with the students. The instructors also provided individualized corrective audio feedback to every student who completed an assignment. The feedback identified problems that the student may have had, provided correct models, and included specific suggestions for improvement.

Voiced stops				
Criteria	Ratings			Pts
Allophones of /b/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /d/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /t/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /g/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /p/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /v/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
Allophones of /k/ are produced correctly	Always 2 pts	Sometimes 1 pts	Rarely or never 0 pts	2 pts
				Total Points: 14

Figure 10.1: Rubric for a graded recording

As is obvious from the description above, students in this LMOOC had access to large quantities of both implicit and explicit feedback. Crucially, the massive and online nature of the course provided the type of feedback that addressed the three goals indicated by Hattie and Timperley (2007) and explained above. Students self-evaluated when they listened to their own pronunciation and recorded it (*how am I doing?*); they received implicit feedback in the form of a model provided (*where am I going?*); and, finally, they were given explicit individual corrective feedback (*where to go next?*) both by their peers in the practice activities and by the instructors in the graded activities.

Based on the author's experience teaching the same Phonetics course in an F2F format, the author wanted to find out if this type of frequent, explicit, individualized feedback that is not readily available in a traditional course could make a difference in a learner's ability to reach the target performance goals. To that effect, the students' progress in the LMOOC was measured and compared to that of students enrolled in the traditional F2F pronunciation course. A quasi-experimental study with a pretest and posttest design was carried out to try to answer the following two research questions:

- Do students in an LMOOC experience the same degree of gains in comprehensibility as students enrolled in a face-to-face course of similar content?

- What components of the LMOOC result in the strongest gains in students' comprehensibility?

The following sections describe the methods used and report on the results of the study.

9.4 Method

9.4.1 Participants

The data for this study come from 50 adult students chosen from a larger population enrolled in two courses described above. The procedure for selecting the students from each group is explained in the next section. Students in both courses completed a background questionnaire at the beginning of the course. They responded to questions about their linguistic background, academic experience in Phonetics, if any, reasons for taking the course, etc. The F2F data come from three sections of the same undergraduate course taught by the researcher and offered at a public university in the United States. The total enrollment for the three sections was 91. The majority of the students enrolled in this course were native speakers of English who were taking the course to satisfy an academic requirement. Most were pursuing a major or minor in Spanish and many of them had had extensive experience abroad. The second course was the LMOOC, "Improving your Spanish Pronunciation", described above, which enrolled a total of 500 students. Participants in the LMOOC came from a variety of linguistic backgrounds. According to the background questionnaire, around 45% had graduate degrees and a majority, almost 85%, indicated that they were taking the course for personal interest.

9.4.2 Procedures

Students in both courses were given a paragraph to read in Spanish (see Appendix A) on the first day of instruction and were asked to record themselves and submit the recordings to the instructor. At the end of the course, both groups were given the same paragraph to read and record themselves again. Students were allowed to record themselves as many times as they wanted until they were satisfied with their performance. They could choose which version of their recording to save and submit. The first recording will be referred henceforth as the pretest and the second recording as the posttest.

The data in this study come from a subset of the total enrolment in both courses. First, any student who did not submit either of the two recordings was eliminated.

Students who indicated in the background questionnaire that they had previous academic experience in Phonetics, even if it was in another language, were not considered for the study. This reduced the number of potential participants to 135 (75 from the face-to-face course and 60 from the LMOOC). Two native speaker judges were asked to rate a random sample of students from both groups. The total number of samples was 100 (60 from the face-to-face course and 40 from the LMOOC). In order to control for possible bias, two judges were chosen who were not highly proficient in a second language. It has been noted that native speakers' familiarity with a certain foreign accent may affect their judgment of non-native speakers' 'accentedness' (Derwing & Munro, 2005; Rajadurai, 2007). Both judges were experienced teachers of Spanish who were accustomed to dealing with learners.

It is important at this point to clarify that the goal of this study was to measure learners' gains in comprehensibility, which is a construct associated and often confounded with intelligibility and accentedness. Normally, intelligibility is defined by how well a listener is able to understand a speaker's utterance and it is typically assessed by having the listener transcribe the learner's speech. Comprehensibility and accentedness are defined as "listeners' perceptions of how easily they understand an utterance and how closely the pronunciation of an utterance approaches that of a native speaker, respectively" (Kennedy and Trofimovich, 2008). There is a general agreement in our profession that the goal of pronunciation instruction is not to help learners achieve a native-like accent, but rather to help them be understood. Therefore, a rating of accentedness would not be an appropriate measure of the success of this LMOOC.

The difference between comprehensibility and intelligibility as Isaacs and Trofimovich (2012) point out, has more to do with whether the notion of intelligibility is understood in its narrow, or broad sense. In its narrow sense, mentioned above, intelligibility is assessed by the accuracy of a written transcription of a learner's utterances. In its broader sense, though, intelligibility is almost synonymous with comprehensibility in that it refers to the listener's ability to comprehend someone's speech. Taking into account the goals of this course and the sample of oral production that is being used as measurement, it seems appropriate to use the construct of comprehensibility.

Following Kang, Rubin and Pickering (2010), the author designed a measure of comprehensibility that consisted of five 7-point bipolar Likert scales as illustrated in Figure 10.2. The judges had to assign a numerical rating in each of the categories. The average of these five ratings was used as a final composite rating for further analysis.

Hard to understand	1	2	3	4	5	6	7	Easy to understand
Incomprehensible	1	2	3	4	5	6	7	Highly comprehensible
Needed little effort	1	2	3	4	5	6	7	Lots of effort to understand
Unclear	1	2	3	4	5	6	7	Clear
Simple to grasp the meaning	1	2	3	4	5	6	7	Difficult to grasp the meaning

Figure 10.2: Comprehensibility rating scale

In order to practice using the rating scale and to test out its reliability, the judges were given fifteen recordings from students who had been eliminated from the study. To facilitate the calculation of reliability, any time the two judges rated a sample within two tenths or less the rating was considered equivalent. For example, if a sample received a rating of 4.2 by the first judge and 4.4 by the second, the two judges were considered to be in agreement. The level of inter-rater reliability in this test run was 0.78 (Cronbach's alpha), which is a strong level for a sample of this size.

Once they familiarized themselves with the scale, the judges were asked to listen to and rate all 100 pretests and posttests over the course of three days to guarantee consistency in the application of the scale. The total listening time was five hours (average length of recording was 54 seconds). The judges were also encouraged to use the whole range of the scale as much as possible and to take frequent breaks to minimize fatigue. The rating was randomized and blind; that is, the raters were not told which samples corresponded to which group or which recordings were done for the pretest or posttest. Inter-rater reliability for the full round of 100 samples was also strong at 0.86.

To control for intra-rater reliability in the full round of ratings and after all the ratings were completed, 15 of the 100 samples were randomly selected and the judges were asked to rate them again. The degree to which their second ratings were consistent with their first was fairly high (0.90 Cronbach's alpha).

9.5 Results

The rating of the pretests revealed that the two groups differed in their level of comprehensibility at the beginning of the course. As expected, the students in the F2F course, the majority of whom were Spanish majors and minors with several years of academic experience in the target language and often experience abroad, received significantly higher scores than the students in the LMOOC (there was no required level of proficiency to register in the LMOOC). The F2F students had an average comprehensibility rating of 5.7, while the LMOOC students' average rating was 4.4. An independent sample t test showed that the difference in average scores between the

F2F group ($M = 5.7$, $SD = .62$) and the LMOOC group ($M = 4.4$, $SD = .66$) was statistically significant ($p = .007$).

In order to guarantee comparability and to be able to accurately assess the effect of the two different formats of instruction, the researcher randomly selected 25 students from each group that would form the basis for further analysis. The validity of the human rating method is widely supported by research and considered to be the benchmark for determining the comprehensibility of speech samples. However, the risk of judge fatigue is a real possibility that puts a limit on the number of samples. This was the justification for setting a relatively low limit for this study. The samples selected had received comprehensibility ratings between 4 and 5 on the pretest and their ratings by the two judges differed by no more than 4 tenths. An independent sample *t* test confirmed that the two groups of students selected were comparable. The difference in average scores between the 25 students selected from the F2F course ($M = 4.52$, $SD = .35$) and the 25 from the LMOOC group ($M = 4.48$, $SD = .36$) was not significant ($p = .75$).

There are several reasons that justify focusing the analysis on students in this range of comprehensibility. First, the 4-5 range effectively represents the middle range of comprehensibility in the scale used in this study, since no student in either course received an average rating lower than 2. Secondly, there were very few students in the F2F course whose average rating was lower than 4, so it would not have been possible to compare students at the lower end of proficiency because they were found almost exclusively in the LMOOC. Finally, students in the middle range are more likely to benefit from the effects of instruction than those at higher levels. Prior research has shown that once learners reach a certain level of comprehensibility, their progress towards native-like norms tends to slow (Martinsen, Alvord & Tanner, 2014). Choosing to study the effects of instruction on a relatively uniform group (in terms of their comprehensibility) also responds to the concern that the effectiveness of instruction in Phonetics may differ depending on the level of the students (Arteaga, 2000). Studying learners at similar levels of proficiency at the time of the pretest eliminates a potential confounding factor.

Once the two samples of pretests were determined to be equivalent a statistical analysis was performed to measure the students' gains in comprehensibility between the pretest and the posttest. This analysis would provide an answer to the first research question. A matched-pairs *t*-test was performed to determine if the difference in comprehensibility ratings after instruction was significant. The results suggest that both groups benefited from the instructional intervention and improved significantly with the greater improvement shown by the students in the LMOOC. Table 10.1 summarizes the results of this analysis.

Table 10.1: Comprehensibility score means for F2F and LMOOC students

	N	Pretest Mean	Posttest Mean	t	df	p
F2F	25	4.52 (.35)	4.70 (.60)	1.71	24	.04
LMOOC	25	4.48 (.36)	4.99 (.47)	1.71	24	<.01

Note: Standard Deviations appear in parenthesis below means

However, when comparing the means of both groups on the posttest, there were still no significant differences. An unpaired *t* test showed a two-tailed *p* value of 0.067, which is considered to be not quite statistically significant.

Because the groups being compared were relatively small, there is always a chance that the sample may not be large enough for the *t* test to have high power of detecting a meaningful difference. Therefore, to fully understand the possible effects of the two different pedagogical interventions, the average pretest to post-test effect sizes were calculated using Cohen's *d*, a standard measure of effect size frequently used in Second Language Acquisition (SLA) studies. Cohen's *d* is calculated as the difference between the means of two groups divided by the pooled standard deviations. Normally, a value of .20 is considered a small effect size, a medium effect size is .50, and a large effect size is .80. The F2F group showed a small average effect size ($d = .37$), while the effect size was very large for the LMOOC group ($d = 1.21$).

In order to answer the second research question (which components of the LMOOC are more effective?), effect sizes for individual students were also calculated. The intention was to identify those students who had made the most significant progress and then analyze their interaction with the different components of the course for possible correlations. Table 10.2 below shows the mean comprehensibility scores and individual effect sizes for the seven students in the LMOOC who showed the most improvement. The analysis of their interaction with the course content is included in the discussion section below.

Table 10.2: Pretest and posttest scores and effect sizes of students in the LMOOC

Student	Pretest comprehensibility rating	Posttest comprehensibility rating	Effect size d
1	4	5.2	2.79
4	4.2	5.2	2.33
6	5	6.2	2.79
13	4	5	2.33
14	5	6	2.33
19	4	5	2.33
21	4.8	5.8	2.33

9.6 Discussion

The first of the two research questions in this study asked whether a pronunciation LMOOC could be as effective as a traditional F2F pronunciation course. The results suggest that the format of this LMOOC was effective in providing students with the necessary tools to improve their level of comprehensibility. The analysis shows that, although both groups made significant gains after instruction, the LMOOC students showed a much larger effect size. The most immediate and general conclusion that can be drawn from the results is that the absence of face-to-face interaction and the large enrollment in the course did not prove to be an obstacle for acquisition.

A somewhat peripheral issue not directly addressed by this study is the fact that a stronger emphasis on the more explicit analysis of the linguistic structure of the target language, which was a characteristic of the F2F course, did not result in an advantage for those students. Previous studies suggest that explicit phonetics instruction is a beneficial, perhaps even crucial, ingredient to improve pronunciation accuracy (Chung, 2007; Derwing & Munro, 2005; Fullana, 2006; Venkatagiri & Levis, 2007). The results of this study, however, do not warrant that claim and this is corroborated by some recent research (e.g., Kissling, 2013). If explicit instruction is not the determining factor, what accounts for the difference in performance between the two groups in this study? Of course, there are a number of factors that were not considered in this study and may account for the findings. However, a plausible interpretation of the results is that the LMOOC provided learners with more quantity and variety of feedback of the type that is identified by Hattie and Timperley (2007) as effective. That is, feedback that would give learners information about how they pronounce, how they should pronounce and what they should do to address any possible deficiencies.

Although both courses gave students plenty of information about the goals, the LMOOC provided learners with more and a wider variety of opportunities to judge their own pronunciation, and the necessary information to address any existing problems. Crucially, and in contrast to the F2F course, the LMOOC offered learners significantly more individualized corrective feedback, both instructor- and peer-generated.

Based on the results of the analysis of individual effect sizes (see Table 2), the researcher decided to study the interaction that the most successful students had with the course content. This analysis would provide the answer to the second research question regarding the most effective components of the LMOOC. The author selected the seven students who showed the largest effect sizes ($d = 2.33\text{--}2.79$). These students' patterns of interaction with the course content were studied using the learner analytics tool included in the LMS platform. What was found was that all seven of these students had completed all the recordings assigned and had therefore received more individualized feedback on their pronunciation than the majority of their peers (only 10 of the 25 students in the study submitted all the recordings). Also, the most successful students accessed their feedback more frequently than the rest. However, these successful students did not differ from the rest in terms of how often they accessed the theoretical explanations at the beginning of each module or the recorded samples of native speaker pronunciation. This finding suggests that the crucial factor responsible for the difference in performance was increased practice and, most likely, frequent access to individualized corrective feedback. This conclusion is consistent with those reached by Dłaska & Krekeler (2013) and Saito & Lyster (2012), based on studies also conducted with intermediate-level learners of German and English respectively.

The current study had a number of limitations, the most important of which is, perhaps, the fact that it looked at gains experienced by a small group of learners at a specific level of proficiency in pronunciation. The size of the sample and, particularly, the level of proficiency of the subjects may have skewed the findings. It is likely that the results would have been different had the researcher looked at students at a lower or higher level of proficiency. Another limitation is the source of the data. Controlled speech production of the type elicited for this study is a convenient source of data, but the results would be more generalizable had they also included data from spontaneous speech production. Finally, there is the limitation inherent to any human rating method. In this case, the ratings reflected impressionistic and somewhat subjective assessments of comprehensibility done by native speakers of Spanish who may have been biased to find certain foreign accents more easily comprehensible than others.

9.7 Conclusions

As was pointed out in the introduction to this chapter, the language teaching profession is still on the fence regarding the feasibility of creating effective learning conditions within the structure of massive online courses. The study reported here

suggests that the particular affordances of online technologies can provide the necessary environment to facilitate the acquisition of the target language sound system. Although the massive nature of the course obviously creates some challenges for the instructor(s), the incorporation of frequent peer-generated feedback seems to be a facilitating factor in the improvement of comprehensibility. The increased flexibility of the format and the added level of autonomy for the student also supported more frequent access to different types of implicit and explicit feedback, which may explain the significant gains in pronunciation.

As a word of caution, the results of this study do not imply that LMOOCs are a better format than F2F courses for language teaching in general and for pronunciation instruction in particular. There are many issues that still remain unexplored, the answers to which will help further our understanding of Second Language Acquisition (SLA) in instructed massive online formats. Regarding the acquisition of the sound system specifically, future research on LMOOCs should look at how the changing roles of students and instructors in this format affect learners' development. With the increased autonomy facilitated by online learning, are advanced students more likely to benefit than those at lower levels of proficiency? A more complete picture would also emerge if future research could look at the acquisition of supra-segmentals, or the effect that increased access to explicit feedback has on the acquisition of certain segmental features that seem to fossilize with exposure to only implicit feedback. The role of peers is another issue that requires further study. As one of the key characteristics of MOOCs, we need to understand how to make the most of the access that learners have to large numbers of peers. In what ways is peer feedback similar to, different from, or complimentary to instructor-generated feedback and how can we best combine the two? Another issue worth exploring is the effect that motivation may play in LMOOCs. As has been widely reported and was the case in this course, most participants in MOOCs are not students in the institutional sense of the word and this certainly affects their motivations for learning a language. Although comparing LMOOCs and F2F language courses is a necessary first step, once the pedagogical viability of LMOOCs has been established, future research should focus on how to best serve the needs of a different audience in a different format that provides significantly different affordances.

Bibliography and Webliography

- Arteaga, D. (2000). Articulatory phonetics in the first year Spanish classroom. *Modern Language Journal*, 84, 339–354.
- Anderson, T. (2004). Towards a theory of online learning. In T. Anderson & F. Elloumni (Eds.), *Theory and Practice of Online Learning* (pp. 33-60), Athabasca: Athabasca University Press. Retrieved May 1, 2014, from http://cde.athabascau.ca/online_book/
- Carroll, S.E. (2001). *Input and Evidence: the Raw Material of Second Language Acquisition*. Benjamins, Amsterdam.

- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1999). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.
- Chung, W. L. (2007). *The effectiveness of explicit, implicit, and noticing instruction: Mandarin speakers' perceptions and production of English sentence stress*. (Unpublished doctoral dissertation). Alliant International University, San Francisco, CA.
- Clougherty, R. J., & Popova, V. (2013). Crowdsourcing to Assess MOOCs: A Position Paper. *MOOCs Forum* 1(P), 10-13.
- de Bot, K., & Mailfert, K. (1982). The teaching of intonation: fundamental research and classroom applications. *TESOL Quarterly* 16 (1), 71-77.
- Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1997). Pronunciation instruction for “fossilized” learners: Can it help? *Applied Language Learning*, 8, 217– 235.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thompson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Testing*, 54, 655–679.
- Derwing, T. M., & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39 (3), 379-397.
- Glaska, A., & Krekeler, C. (2008). Self-assessment of Pronunciation. *System* 36 (4), 506-515.
- Glaska, A., & Krekeler, C. (2013). The short-term effects of individual corrective feedback on L2 pronunciation. *System*, 41, 25-37.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies*. (pp. 339–360). Oxford: Oxford University Press.
- Fullana, N. (2006). The development of English (FL) perception and production skills: Starting age and exposure effects. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 41–64). Clevedon, UK: Multilingual Matters.
- Friedman, T. (2012, May 15). Come the Revolution. *The New York Times*. Retrieved from http://www.nytimes.com/2012/05/16/opinion/friedman-come-the-revolution.html?_r=0.
- Hattie, J., & Timperley, H., (2007). The power of feedback. *Review of Educational Research* 77 (1), 81-112.
- Hendrickson, J. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal*, 62, 387–398.
- Isaacs, T. (2009). Integrating form and meaning in L2 pronunciation instruction. *TESL Canada Journal/ Revue TESL Du Canada*, 27, 1–12.
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing Comprehensibility: Identifying the Linguistic Influences on Listeners' L2 Comprehensibility Ratings. *Studies in Second Language Acquisition*, 34, 475-505.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, Comprehensibility, and Accentedness of L2 Speech: The Role of Listener Experience and Semantic Context. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 64 (3), 459-489.
- Kennedy, S., Blanchet, J., & Trofimovich, P. (2014). Learner Pronunciation, Awareness, and Instruction in French as a Second Language. *Foreign Language Annals*, 47, 79-96.
- Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental Measures of Accentedness and Judgments of Language Learner Proficiency in Oral English. *The Modern Language Journal*, 94 (4), 554-566.
- Kissling E., M. (2013). Teaching Pronunciation: Is Explicit Phonetic Instruction Beneficial for FL Learners? *The Modern Language Journal*, 97, 3, 72-744.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Loewen, S. (2012). The role of feedback. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 24-40). New York: Routledge.

- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
- Long, M. L. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
- Mackey, A., Gass, S., & McDonough, K. (2000). Learners' perceptions about feedback. *Studies in Second Language Acquisition*, 22, 471–497.
- Martinsen, R. A., Alvord, S. M., & Tanner, J. (2014). Perceived Foreign Accent: Extended Stays Abroad, Level of Instruction, and Motivation. *Foreign Language Annals*, 47 (1), 66–78.
- Pappano, L. (2012, February 11). The Year of the MOOC. *The New York Times*. Retrieved from <http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all>.
- Online Courses: Possibilities and Pitfalls. (2013, January 8). *The New York Times*. Retrieved from http://www.nytimes.com/2013/01/29/opinion/online-courses-possibilities-and-pitfalls.html?_r=0
- Rajadurai, J. (2007). Intelligibility studies: A consideration of empirical and ideological issues. *World Englishes*, 26, 87–98.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning* 62 (2), 595–633.
- Spada, N., & Lightbown, P. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *The Modern Language Journal*, 83, 1–22.
- Trofimovich, P., Ammar, A., & Gatbonton, E. (2008). How effective are recasts? The role of attention, memory, and analytical ability. In: Mackey, A. (Ed.), *Conversational Interaction in Second Language Acquisition. A Collection of Empirical Studies* (pp. 171–195). Oxford: Oxford University Press.
- Truscott, J. (1999). What's wrong with oral grammar correction. *Canadian Modern Language Review*, 55, 437–456.
- Venkatagiri, H., & Levis, J. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness*, 16, 263–277.

Appendix

Paragraph used for the pretest and post-test

Tengo muchos amigos en la Universidad. Mi amigo Eduardo estudia ingeniería y vive en una residencia en el campus. También tengo una amiga que estudia matemáticas; se llama Reyes y vive en la misma calle que yo. Comparte su apartamento con Rosa, otra estudiante. Todos estamos muy ocupados durante la semana, pero los fines de semana, cada uno prepara algo de comida y nos juntamos para cenar en mi casa. Hoy es sábado y todos han preparado algo especial para celebrar el cumpleaños de Pedro, que es de Portugal y también estudia en la Universidad.