Putting an accent on the positive: New directions for L2 pronunciation research and instruction

Tracey M. Derwing
University of Alberta; Simon Fraser University, Canada
tderwing@ualberta.ca

Abstract

The last ten years have witnessed an explosion of activity in second language (L2) pronunciation research, and many signs of increased interest on the part of L2 teachers to include pronunciation in their classes. However, most surveys of language teachers (e.g., [1], [2]) indicate that instructors feel they have received insufficient training to teach pronunciation and consequently do not know where to start; despite their feelings of inadequacy, they realize that some of their students would benefit from pronunciation instruction (PI). Moreover, many of the questions that teachers and students have about L2 pronunciation are not yet fully answered. Nonetheless, research has suggested several aspects of English that are ‘more important’ to teach than others. Identifying these requires consideration of functional load and stress, among others. A brief history of L2 pronunciation teaching is followed by definitions of key terms. Appropriate goals for the field of L2 pronunciation research and instruction will be outlined, and the findings of several studies that address instructors’ questions will be discussed, as will possible future directions for research and teaching alike.

Keywords: pronunciation instruction; pronunciation research; intelligibility; comprehensibility.

1. Introduction

To understand the present it is informative to know something about the past. In this paper, I will outline a brief history of second language pronunciation research and teaching, followed by a description of current concerns. Several important key concepts will be defined, and some successful, illustrative studies will be reviewed, followed by a discussion of speech models to be used in the classroom. Finally, recommendations will be made for both PI researchers and teachers.

Although language teaching, and therefore pronunciation teaching, have both been carried out for centuries, this review focuses on the last 60 years, beginning with one of the most influential methods of language teaching, Audiolingualism. Audiolingual instruction was very popular in many countries from the late 1950s through the early 1970s, and in fact, it has never really gone away. Audiolingual instruction involved a tremendous amount of controlled repetition to ensure that learners had accurate pronunciation right from the beginning; students were given few opportunities to learn erroneous patterns. Although this method was popular with teachers, as it was easy to use, often audiolingual classes were boring for the students; Flanders and Nuttall [3] in a comprehensive review, found that learners were most satisfied in classrooms where the teacher deviated from the text. By the 1970s, a strong interest in other approaches emerged, and as a result, several new ‘designer methods’ became popular, including Suggestopedia, Community-Counselling Learning, and The Silent Way. The latter, developed by Gattegno [4], was the only one that prioritized pronunciation. Gattegno advocated associating each phoneme with a distinct colour to serve as a mnemonic. A drawback of this approach is the very limited vocabulary provided in the first few months of a Silent Way program. The designer methods reappear from time to time, but for the most part, they are no longer used extensively.

During the period between the late 1970s and the 1990s, communicative language teaching (CLT) became extremely popular, especially in North America, but other areas were influenced as well. Although this approach to language teaching was more engaging for learners, an unfortunate aspect of CLT was the almost complete abandonment of PI. Krashen [5] introduced a theory that language acquisition was primarily a matter of input, and that development in all language skills was dependent on adequate exposure rather than formal teaching. At the same time, the critical period hypothesis (that individuals can learn a second language to a level similar to their first, using the same mechanisms until puberty) was a topic of considerable interest in applied linguistics. Much of the discussion about the critical period focused on the fact that most learners past puberty retained a foreign accent in their L2 [6]; many people concluded that since native speaker-
like productions were beyond the reach of adult language learners, it was pointless to teach pronunciation. Perhaps the final nail in the coffin for pronunciation teaching during this period was an influential study suggesting that there was no relation between instruction and accent: Purcell and Suter [7] maintained that “teachers and classrooms seem to have had remarkably little to do with how well our students pronounced English” (285). This sentiment spread, and by the 1990s it was generally accepted that “There is no convincing empirical evidence which could help us sort out the various positions on the merits of pronunciation teaching” Stern, p. 112 [8]. Stern was right in the sense that there had been very little research on PI. Most of the strongest advocates of PI were educators who followed their intuitions, but who did not carry out controlled investigations. CLT, often with a focus on form, and task-based language teaching, continue to dominate to the present day. However, teachers began to realize that some learners with access to a lot of input still have difficulties with pronunciation. Pronunciation did not seem to be directly tied to proficiency and amount of input after all. In 2010, Celce-Murcia et al. [9] decried the lack of attention to pronunciation, proclaiming that the area was the ‘Cinderella’ of second language acquisition (SLA). That statement was a bit late coming, because considerable PI research started to appear in the 1990s and early 2000s, and now, in 2018, it can be considered the ‘Belle of the Ball’ in SLA circles. Empirical investigations have proliferated, as noted in a meta-analysis of 86 studies conducted by Lee, Jang and Plonsky [10], who noted a strong significant effect for PI. Interestingly, though, most of the studies surveyed focused solely on accent reduction, or nativeness, rather than taking into consideration improved intelligibility [11]. This distinction is important; it refers to Levis’ [12] Principles of Nativeness and Intelligibility. The former maintains that the goal of PI is native-like (accent free) productions, whereas the latter states that pronunciation should be taught with the goal of increased intelligibility (retaining some aspects of accent from the L1). Levis argued that realistically, instructors and researchers should be guided by the second principle.

Although PI research has expanded greatly in the last 15 years, teaching has not yet followed suit to the same extent. Several surveys have been conducted in Australia [13, 14], Canada [15, 1] the UK [16] Europe [17] and the USA [2]. Most of these studies involved teachers of English, but Huench’s study [2] surveyed teachers of Spanish, French and German. All the studies pointed to very limited PI in language programs, and expressed a concern on the part of teachers that their students needed PI, but they didn’t know how to provide it. Moreover, some surveys indicated that even if teachers had wanted training in teaching pronunciation, very few universities actually offer it. There aren’t many other options for training either: most pronunciation-oriented conferences are geared to researchers, and the only journal dedicated to pronunciation is somewhat technical for the needs of most instructors. The last four years have witnessed the publication of several empirically-based textbooks (e.g., [18, 19, 20]; but teachers are unlikely to read a text unless it is tied to a course. One highly accessible resource is a reliable and up-to-date website designed by John Levis: pronunciationforteachers.com. It has short, readable entries from experts on key concepts and teaching techniques as well as links to vetted webinars. This website will continue to grow over the next few years; the challenge is to make language teachers aware of it.

This very short history brings us to the present; we are now in a period of strong interest in pronunciation teaching. Much more research is being conducted, but what about the teachers’ concerns: how can we help them and what does the research actually say? The primary focus of the rest of this paper is on English pronunciation, because that is the L2 in which most research has been done, but there are studies in other L2s as well. However, some concepts are universal; and we will now consider some distinctions that hold for all L2s.

**Definition of key terms**

*Intelligibility* is the most important dimension for L2 pronunciation speakers, teachers, and researchers; it is how much the listener actually understands of the speaker’s intended message. Clearly, language users’ primary linguistic goal is to communicate, and if an interlocutor doesn’t understand a speaker there is a serious problem. Sometimes the issue lies with both the listener and the speaker, but in the case of second language learners, pronunciation can be the cause. A second important dimension is *comprehensibility*: how easy or difficult it is for a listener to understand a speaker. If a listener must expend a great deal of effort to understand a given speaker, he or she may avoid interaction. *Accent* is the third dimension; that is, how different a speaker’s speech is from that of the listener. In some cases heavily accented speech is both easy to understand and fully intelligible. In others, some features of an accent may cause the speaker to be difficult to understand, and even partially unintelligible.
These three dimensions of speech can all be measured, but intelligibility is the most difficult to determine because the listener is sometimes at fault for a lack of understanding (e.g., a momentary lack of attention may interfere with understanding a speaker’s message). Although imperfect, dictations, summaries, and responses to true/false sentences will help the researcher to determine whether the listener has understood. Comprehensibility and accent are commonly measured with scalar judgments. Munro [21] has compared scales of 1-9 with 1-1000, and the results are virtually identical. It appears that anything from 9 points and up is enough for listeners to rate these dimensions (although fewer points on the scale can be too restrictive). One of the most robust findings that has emerged from every study my colleagues and I have done is that intelligibility and accentedness are partially independent. In general, comprehensibility ratings are more closely related to intelligibility than are accented ratings. Moreover, comprehensibility and accentedness are quite distinct. When we ask people to rate speech, listeners usually agree with each other quite strongly on who has a heavy accent and who does not, and who is easy to understand and who is not. This is true whether they are English as an L1 or L2 listeners. Reliability is almost always over .9. From our perspective, listeners’ judgments provide the best insights into accentedness and comprehensibility. The human ear is better than software or acoustic measures at deciding what is easy or difficult to understand. Many interesting things can be learned from acoustical analyses, but those types of measures do not correlate completely with humans’ perceptions.

2. Important Considerations for Pronunciation Instruction

Which aspects of pronunciation should be taught?

What should second language teachers teach in terms of pronunciation? Teachers cannot be expected to allocate a lot of time for pronunciation as a separate skill. The aspects of pronunciation deserving highest priority are those that have the greatest effect on intelligibility and/or comprehensibility. Does research do this? Not always. Aspects of speech that interfere with understanding matter more than aiming for nativelike performance. So for example, Abe counted instances of connected speech in a sentence reading task both before and after instruction, comparing L2 speech to NS speech. After instruction, the L2 speakers’ productions were more native-like, but whether it would have made any difference to listeners in terms of intelligibility or comprehensibility, we do not know, because counting instances of connected speech is not the same as asking listeners. Research involving human judgments is useful because it can help us test the value of pedagogical activities.

Many teachers have excellent intuitions, but others do not. In fact, research should be the foundation for most things related to pronunciation, including the cognitive processes involved, attitudes, teaching goals, the focus of instruction, and the effectiveness of activities. There is a difference between research and intuition, which is the main reason I think teachers need to have pronunciation-specific training. If they rely on intuitions, sometimes students will be at a disadvantage. For instance, a teacher in Canada reported that he had discovered a good technique for teaching Arabic students the difference between English /b/ and /p/. Rather than making the voicing distinction, he claimed to have discovered that /b/s are made by inhaling, while /p/s are made by exhaling. He had students inhaling as they repeated the word ‘baby’ 50 times. This was a case where intuition failed. Another pronunciation teacher in Canada, using intuition, also advocated ingressive /b/s. But this technique is not only misleading, it could lead to less intelligibility for students.

Teachers often advise their students to slow down: an intuition shared by many. For some individuals, this is good advice – that is, for those people who are already fast talkers. However, most L2 speakers actually speak at a slower rate than their native speaker counterparts. The reason they may seem to speak quickly to their listeners’ ears has to do with other differences in their speech, some related to pronunciation, and some to unexpected vocabulary choices, grammatical errors, and so on. These are usually the real culprits, not speech rate, and if a teacher really wants to help his or her students, it’s important to identify what the actual problems are rather than simply glossing over them by asking students to speak more slowly.

Munro and Derwing [22] examined speech rate using computer software to speed up L2 productions by 10% without changing pitch or distorting the utterances in any way. A curvilinear relationship was found between comprehensibility and speaking rate, such that faster was better for slow to medium rate speakers, but if speakers already talked quickly, increasing their rate had a negative effect. But it wasn’t until speakers were at the 5 syll/second point that comprehensibility was compromised. The majority of the learners would have benefited from speaking faster rather than slower, a finding that is counter-intuitive for many teachers. This is just one
example of how research may provide insights that are not fully clear or intuitive on the surface. Another is the relative value of individual segments. Many textbooks and curricula treat all segments or phonemes in English as though they were equally important. This brings us to the concept of functional load: the number of words in a language that a particular phonemic contrast keeps apart.

Catford [23] produced a chart with his estimates of high, mid and low functional load pairs for English. His hierarchy indicated, for instance, many words differ only in terms of whether they have a /p/ or a /b/, but very few words are separated by /v/ and /z/. In an exploratory study, Munro and Derwing [24] assessed whether low functional load errors in L2 speech matter less than high functional load errors, as Catford [23] had suggested. A large database of Cantonese-accented sentences was searched for utterances with no errors of any type; with one low functional load error; two low functional load errors; one high functional load error; and sentences with two high functional load errors. Raters heard a randomized set of these utterances and judged their comprehensibility. As hypothesized, high functional load errors caused a greater reduction in comprehensibility than low functional load errors.

Another feature of pronunciation that can affect intelligibility is stress. Laura Hahn [25] compared listeners’ understanding of sentences produced using appropriate nuclear stress, misplaced stress, and in monotone or no stress condition. She found that listeners had more difficulty understanding the misstressed and monotone conditions and that correctly stressed sentences were easier to understand. This finding suggests that stress is important. Field [26] and Zielinski [27] also found evidence that stress placement in words matters to intelligibility. Kate Hahn [28] conducted a study in which students in a pronunciation course were tested before instruction, after, and in a delayed post-test. She found a significant effect of instruction and overall the learners retained most of what they had learned. Tanner and Landon [29] conducted research on instruction word stress and found a resulting improvement of comprehensibility. These research projects suggest that a) stress matters to both intelligibility and comprehensibility; b) it can be taught; and c) students can maintain their progress. There are many other studies of individual features of pronunciation, but often without an indication of whether they affect intelligibility.

Sometimes teachers think that pronunciation is too difficult to explain or introduce with beginner level students, but that is not the case. It’s never too early, as Zielinski and Yates [30] have pointed out.

Derwing and Munro [18] followed the naturalistic progress of two groups of learners for 10 years, and noted that the greatest improvement in pronunciation was in their first six months of massive exposure, a phenomenon called the window of maximal opportunity (WMO). The WMO seems to be the ideal time to introduce PI. However, it is never too late; even people who have been speaking their L2 for years can still benefit from focused PI, as the following example demonstrates.

Derwing et al. [31] were invited to help some employees in a window factory who had difficulty with making themselves understood. Seven participants volunteered to take part in this study. They had lived in Canada for an average of 19 years and their average age was 43. All of them improved. Pre- and post-instruction tests were designed, and a course was developed that took place three times a week for 30 minutes. In total, the students had 17 hours of class-time, but they also committed to a minimum of ten minutes of homework a day, five days a week. First, two perception tasks indicated that the participants’ listening in English improved significantly. To assess speaking, 28 judges listened to randomized pre- and post-instruction samples from a picture narrative and rated them for comprehensibility on a 9-point scale. The judges also rated samples from a safety talk, an authentic speech task that most of the employees had to perform in their jobs. They talked about the dangers of working when very tired. Both tasks evidenced a significant improvement in comprehensibility. Finally, the judges listened to true/false sentences. There were significantly more correct answers in the post-instruction sentences, indicating improved intelligibility. A delayed post-test 3 months later showed that the comprehensibility and perception gains were maintained. The practical implications of all these studies are the following:

- Pronunciation instruction can work to the extent that listeners (both L1 and L2) can detect a significant improvement.
- A focus on intelligibility/comprehensibility makes more sense than focusing on accent alone.
- It is better to start early, during the WMO before certain patterns get entrenched.
- It is better to intervene late rather than never.

**Which speech model should be adopted?**

Which speech model is most appropriate to teach? This can be a political question and most often people debate about English as a Lingua Franca, or World Englishes versus Received Pronunciation or General American English; textbooks and recordings tend to fit into the last two categories. This question is often moot because in most cases, the students...
will learn whatever dialect their teachers speak. But it is important to ask, to whom will students talk in English? If, for the most part, they will interact with other speakers in the same community, then the local variety is probably the best choice. However, teachers do not always know with whom their students will interact in the future. Their jobs, studies, and travels may take them out of the local context, therefore they may benefit from exposure to a wide range of Englishes. Given that High Variability Phonetic Training (HVPT) works so well [32] in improving perception generally, and, in some cases, production also, and the fact that many students may need to understand people from several varieties of English, it makes sense to use HVPT more often in classrooms, and especially for homework. Furthermore, the notion of HVPT can be expanded beyond phonetic contexts to discourse/listening activities with many different speakers in general.

In a recent paper, Hansen Edwards, Zampini, and Cunningham [33] compared the accentenedness, comprehensibility and intelligibility of speakers from Hong Kong, Singapore, China, and the USA. A key finding in this study was that speakers with international experience received higher intelligibility scores than those who had not lived abroad. This is an indication that a broad notion of HVPT, in conjunction with other approaches, is an appropriate way to go, because those who had international exposure were likely exposed to many more voices and differences than those who were not.

In the typical sense of HVPT, the website Englishaccentcoach.com provides perception practice. It is based on empirical research in which learners’ perception performance was traced over time as they used a pilot version of this site. Another useful website for multiple exposures to pronunciation is Youglsh, a fabulous resource for listening. Students type in any word or phrase they want to hear, and immediately videos of different people saying the word or phrase come up, in context, and in natural speech. Sitcoms with scripts are readily available; students can mark the prosody and then act the scene, imitating the speakers. Interviews and talk shows from several different World Englishes can be used in a similar manner; students can transcribe short segments and analyse them.

3. Discussion

There are some basic principles for PI. First, teach perception to L2 learners. Students who have been speaking their L2 for a long time are their own most frequent source of input; they hear their own speech more than that of any other person. Years of input from their own speech patterns contributes to entrenchment.

Second, provide explicit instruction. Once a teacher has identified where students’ perceptual problems lie, he or she should give explicit instruction with examples, explaining what the differences are. So for example, if a student confuses /U/ and /n/, a high functional load minimal pair, it is advisable that the teacher point this out.

Third, provide explicit corrective feedback; when students mishear, or produce an error, this should be brought to their attention. Saito and Lyster [34] describe an approach to teaching English /a/ to Japanese learners within the context of communicative activities. Significant improvement was noted after four hours of instruction, where the emphasis was on corrective feedback.

Fourth, use authentic language. Videos from the internet can be incorporated in a range of activities, including heightening perception, serving as a catalyst for explicit explanations, and providing shadowing opportunities. The suggestion to use authentic language is directed at researchers too. In Thomson and Derwing’s [11] review of PI research, the vast majority of studies used reading stimuli for their listening experiments. In other words, the L2 speakers read a sentence or even just a word. That meant that they could put all their efforts on pronunciation, but in real life, using authentic language, L2 speakers also have to find the words and the appropriate grammar, as well as the pronunciation. For these reasons, researchers would be well advised to use more authentic stimuli when testing the effectiveness of instruction.

Fifth, make good use of technology, which offers the tremendous advantage letting learners practice on their own time. This allows for individualization of instruction; the teacher can point the learner to areas of focus that are troublesome. Many options are available on the web, including, unfortunately, many poor quality offerings, some of which may actually do more harm than good. It is therefore important that the teacher suggest what the students should work on at home.

There are many possibilities for research to support L2 learners. First, more investigations are needed of L2 phonological development, both naturalistic and instructed across multiple languages. Some aspects of L2 phonology may not need to be taught if they tend to develop naturally, whereas others may require intervention. Matters such as voice quality have barely been touched. More studies are needed investigating the relative importance of various accent features to
intelligibility. Research in languages other than English is also a priority.

As for pedagogy, preliminary studies of pronunciation teaching have given us a starting point, but there is room for more research to identify effective teaching approaches. Technology offers a great deal of promise, provided that technological applications are guided by pedagogical specialists who understand appropriate goals and priorities in teaching pronunciation. There certainly is a place for more teacher training; without preparation, many language instructors will neglect this aspect of language learning.

4. Conclusions

Ultimately, it is important to move from seeing accent as inherently a problem, to viewing accent as a normal phenomenon. Instead of focusing on accent reduction or eradication, we should be aiming for intelligibility and comprehensibility enhancement. Rather learners emulating NSs, teachers should be focused on intelligibility-based goals. A haphazard, one size fits all approach to pedagogy must be replaced by discerning pedagogy, focusing solely on problems that interfere with communication; this holds not only for the classroom but for teaching materials. Finally, instead of defining intelligibility entirely in terms of native listeners’ perceptions, we should be aiming for intelligibility for a diverse audience. This is a complex issue and needs considerably more investigation.

5. Acknowledgements

I am grateful to the conference organizers, in particular, Ian Wilson, for their wonderful hospitality. Many of the ideas and studies in this paper were developed with my longtime research colleague and friend, Murray Munro. I also want to acknowledge my colleagues Marian Rossiter, Ron Thomson, Jen Foote, and Erin Waugh. Many of the studies reported here were conducted at NorQuest College, MacEwan University and Metro: I thank the participants, language teachers, and program directors. Much of the work reported here was funded by grants from the Social Sciences and Humanities Research Council of Canada, and the University of Alberta.

6. References


