Language Learning & Technology
ISSN 1094-3501

February 2020, Volume 24, Issue 1
pp. 86–106

ARTICLE

Pronunciation development and instruction in distance language learning

Ines A. Martin, U.S. Naval Academy

Abstract

The goal of this study was to explore how distance language learners' pronunciation skills develop with and without targeted pronunciation training during their first semester of university language instruction. To this end, a new computer-assisted method of pronunciation instruction was designed, and its effectiveness, as well as learners' experiences with the method, were assessed. The study was conducted over the course of one semester with 67 distance learners of German. Perception and oral production skills on the word and sentence level were assessed at the beginning and end of the semester for a treatment group that received targeted pronunciation instruction and a control group that did not receive pronunciation instruction, but otherwise followed the same curriculum. The results of the study showed that distance learners' pronunciation skills did not improve significantly over the course of the semester in the absence of targeted pronunciation training. Results further indicated that learners who received targeted pronunciation training improved significantly from pre- to posttest and significantly outperformed learners in the control group on measures of perception and production accuracy at the end of the semester. These findings suggest that distance language instruction can benefit from including targeted pronunciation training.

Keywords: *Distance Learning, Online Teaching & Learning, Pronunciation, Computer-Assisted Language Learning*

Language(s) Learned in This Study: *German*

APA Citation: Martin, I. A. (2020). Pronunciation development and instruction in distance language learning. *Language Learning & Technology*, *24*(1), 86–106. https://doi.org/10125/44711

Introduction

While some educators remain in doubt about the value of distance[1] language learning courses or simply fear that technological developments in online learning will replace their role in the classroom (Cheng, 2015), Seaman, Allen and Seaman's (2018) report of the state of online learning in U.S. higher education reflects undeniable growth in distance education enrollment with a 17.2% increase between 2012 and 2016. The report shows that nearly one third (31.6%) of all students in U.S. higher education have online learning experience by the time they graduate. This trend comes as no surprise considering the many advantages of online learning, including bringing education to people with geographic limitations or allowing flexible access to instruction for adult learners with professional commitments and family responsibilities (Baralt & Morcillo Gómez, 2017).

The trend for a rise in enrollment in such online courses described in Seaman, Allen, and Seaman's (2018) report is further supported by the results of the US Department of Education's meta-analysis of more than 500 online learning studies, which found that students in distance learning environments even performed modestly better than those learning the same materials in face-to-face (F2F) learning environments (Means, Toyama, Murphy, Bakia, & Jones, 2010). Such a meta-analysis suggests the lasting value and viability for online learning in general; however, it has not yet been published specifically for the context of online language learning. In fact, many educators and researchers have voiced concerns about the value of

language learning at a distance because second language (L2) learning occurs through interaction; and yet, in distance learning environments, the instructor—maybe the largest source of native or native-like L2 input and feedback on learner errors—is removed in space and time, and thus the nature of interaction is changed and limited (Doughty & Long, 2003; Nielson, Gonzalez-Lloret, & Pinckney, 2009).

Nevertheless, results from studies comparing the success of online and F2F language instruction have painted a largely optimistic picture for the value of online language learning (Chenoweth, Jones, & Tucker, 2006; De Paepe, Zhu, & Depryck, 2018; Isenberg, 2010). While research overall suggests that students can successfully learn languages online, studies rarely reliably compare proficiency in all four skills. As Deutschmann, Panichi, and Molka-Danielsen (2009) point out, the skill that suffers the most from the limited amount of natural interaction in the L2 in online environments is oral proficiency.[2] R. Blake, Wilson, Cetto, and Pardo-Ballester (2008) investigated the effectiveness of an online Spanish course in improving learners' oral proficiency. Their study compared the level of oral proficiency of a F2F, online, and hybrid class at the end of a first-semester Spanish course. Results from their comparison showed that learners in all three groups performed at a comparable level of oral proficiency at the end of the semester. Thus, the authors concluded that the online learners' oral proficiency developed on par with the F2F and hybrid learners' oral proficiency (see also C. Blake, 2009; Isenberg, 2010; and Lin, 2014).

It should be pointed out, however, that most studies on oral proficiency assess either fluency and accuracy or lexical and syntactic skills, rather than pronunciation. Thus, there is a dearth of research that addresses pronunciation development in online language learning (Inceoglu, 2017). To the best of my knowledge, there is currently no published study that explores the development of L2 pronunciation skills in online language learning environments or investigates the inclusion of particular methods to facilitate pronunciation instruction in this learning environment. The aim of the present study is to address this gap in the literature. That is, the present study investigates whether distance language learners' pronunciation skills improve without targeted pronunciation training and whether a new method of computer-assisted pronunciation training is beneficial for distance learners in improving their L2 pronunciation skills.

Literature Review

The Goal of Pronunciation Instruction

While in the second half of the 20th century, pronunciation instruction seems to have received less attention than other fields in Second Language Acquisition (SLA; Derwing & Munro, 2005), the past twenty years have seen an increase in research and interest in pronunciation instruction and, thus, many fundamental questions in the field have been answered (Derwing & Munro, 2015). Lee, Jang, and Plonsky's (2015) meta-analysis of 75 pronunciation studies, for example, showed that pronunciation instruction results in medium-to-large and statistically significant effects. Comparing their results to meta-analytic findings in other areas of instructed SLA, the authors conclude that these results show "that instruction on pronunciation can be just as (or more) effective as vocabulary, grammar, and pragmatics" (Lee et al., 2015, p. 357; see also Thomson & Derwing, 2015 for a narrative review that finds that 82% of studies report a significant improvement of learners' pronunciation through instruction). Moreover, recent research has shown that learners believe it is important to improve their L2 pronunciation and like to see pronunciation instruction included in their lessons (Couper, 2003; Huensch & Thompson, 2017). Taken together, this indicates that learners should be given the opportunity to improve their pronunciation through training in instructed SLA.

When defining improvement in pronunciation or referring to good pronunciation, however, it is important to consider the true goal of pronunciation instruction. Munro and Derwing's (1995) seminal study introduced the tripartite distinction between accentedness (a perception of the relative strength of one's foreign accent), intelligibility (the degree to which one's utterance is actually understood), and comprehensibility (a listener's assessment of ease or difficulty in understanding the utterance). While these three dimensions of speech are related, they are also independent; for example, heavily accented L2 speech

can still be intelligible and comprehensible (Derwing & Munro, 2015; Gordon & Darcy, 2016). These considerations—in addition to findings that native-like pronunciation is unlikely to be attained by most L2 learners due to maturational constraints on the acquisition of phonology (Scovel, 2000)—have shaped an ongoing shift among pronunciation researchers and teachers from goals of accent reduction (the nativeness principle) to developing intelligible and comprehensible speech in L2 pronunciation instruction (the intelligibility principle; Huensch, 2018; Levis, 2005). In fact, the focus on comprehensibility and intelligibility has become so strong in the field that some researchers have made it a point to warn teachers not to forget about the dimension of accentedness altogether, particularly because of social aspects of accent (i.e., social stigma, feelings of belonging, etc.; Derwing & Munro, 2015; Levis, 2016a).

Another important component in goal-setting for pronunciation training is target feature selection by the instructor. One factor is to select targeted sounds according to their *functional load*: that is to say, sounds are ranked based on the minimal pairs they differentiate in the L2 (Huensch, 2018; Levis, 2016b; Munro & Derwing, 2006). Another factor to consider is the overall impact of segmental vs. suprasegmental errors on the three speech dimensions. Training on the segmental level targets individual speech sounds (e.g., consonants, vowels, and their language-specific combinations), whereas suprasegmental training focuses on prosodic elements (e.g., stress, rhythm, intonation, pauses). Ultimately, current research encourages a combination of instruction in segmental and suprasegmental training (Derwing, 2013; Saito, 2012). In summary, the goal of pronunciation instruction should always be kept in mind when designing pronunciation training materials.

Perception and Production Training

While the perception-precedes-production hypothesis has remained controversial—especially with more and more studies showing that learners can produce certain phonemes quite well without being able to hear the difference between them, for example, when production is achieved through training of articulatory parameters or visual feedback (Olson, 2014; Smith, 2001)—researchers agree that perception and production are inextricably linked (Derwing & Munro, 2015). Seeing that learners can rarely produce new L2 sounds that they cannot distinguish from other perceptually similar sounds in their L1, perceptual training is then an important element of pronunciation training since learners usually have to build new perceptual categories in the L2 phonological space (Flege, 2007; Levis, 2016b). Generally, perceptual training thus still precedes production training in teaching practice, which goes back to the fact that some research has found that perceptual training alone can lead to improvements in production (Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005; Wang, Jongman, & Sereno, 2003).

Perceptual training has been shown to be particularly effective when carried out using input that fully represents the range of variation in normal speech. That is, training learners to perceive sounds that are produced by multiple speakers and in varying phonetic contexts results in significantly greater improvement than perceptual training that simply relies on a single talker or single words (Thomson, 2012). This type of training is referred to as high variability phonetic training (HVPT) and is based on the principle that exposing L2 learners to varied phonetic input allows them to establish a more robust L2 sound system, which then allows them to generalize to sounds in new words and words uttered by unfamiliar speakers (see Thomson, 2011, 2012 for a review of HVPT research; and Levis, 2016b for a summary of the role of HVPT in practical pronunciation instruction). Huensch (2016) showed that HVPT allowed learners to generalize production improvements to continuous speech, making the training viable for real life situations and showing that HVPT is worth including in L2 classrooms. In summary, pronunciation training materials should include a component of perceptual training, ideally preceding the component of production training. If possible, perceptual training should include input from various speakers.

Using Technology to Teach Pronunciation

The fast-paced development of technology over the past two decades has changed the landscape of language learning and teaching (Hsu, 2016). The wide range of technological applications and learners' highly developed digital literacy skills allow for a new approach to teaching foreign languages in general, and

teaching pronunciation in particular. While Derwing and Munro (2015) caution not to expect miracles from Computer Assisted Pronunciation Training (CAPT) and recommend only using it as one tool among others to teach pronunciation, CAPT has some distinct advantages. Since it allows for training outside of the classroom, CAPT offers a solution to the constant time restriction in traditional classrooms, permits more individualized training as learners can choose to only complete the training that addresses their individual pronunciation problems, and makes pronunciation training in distance language learning possible. In the following, a new method of CAPT, which was designed for use in F2F and distance language learning environments alike, will be introduced.

The Current Study

iCPR Design

The method of pronunciation training employed in the current study is called innovative Cued Pronunciation Readings (iCPR) and is based on Tanner and Landon's (2009) Cued Pronunciation Readings. Their original method was designed for intermediate level ESL learners, targeted suprasegmental features, and was completed by learners for extra-credit in a language lab. The new iCPRs were designed for novice learners of German—since research has shown that beginning learners benefit the most from pronunciation training (Roccamo, 2014)—and includes training on both segmental and suprasegmental features to target errors that impede L2 German learners' intelligibility, comprehensibility, and accentedness. Moreover, iCPRs are delivered through the use of easily accessible technology (i.e., Microsoft PowerPoint®) so that every learner can access the pronunciation training on a home or library computer. As such, the method was designed to be implemented in the standard curriculum as weekly homework exercises.

Informed by the research outlined above, the method of iCPR includes both perception and production exercises with perception units usually preceding production units (see Materials section for further details on perception and production units). Findings in HVPT were considered in the design of the materials: The stimuli in perception and production units were recorded by nine native speakers, thus offering input from multiple voices. iCPRs use the method of learners modelling their productions after native-speaker input, a common practice in pronunciation training (Kissling, 2013; Weinberg & Knoerr, 2003). Furthermore, iCPR units were designed so that learners can always see the word while they listen to the native-speaker recording and practice their own production of the word, based on Thomson and Isaacs's (2009) finding that learners' pronunciation of individual tokens was best when they produced them after simultaneously hearing the word and seeing its written form.

Another feature of iCPR units is the inclusion of explicit instruction on the production of individual sounds and word stress, as this approach has been shown to aid learners in improving their pronunciation (for a review, see Thomson & Derwing, 2015). Finally, it should be pointed out that the learners only received corrective feedback in the perception units—in the form of automated feedback—but not in the production units. While some research shows benefits of corrective feedback in pronunciation training (Saito & Lyster, 2012), other research suggests that corrective feedback is not always necessary for improving pronunciation skills, particularly at the lower levels of language proficiency (Saito, 2015). This is further confirmed by the results of Lee et al.'s (2015) meta-analysis in which the authors found only a small difference in effect sizes based on whether feedback was provided or not (for instance, $d_{feedback} = 0.92$ vs. $d_{no\ feedback} = 0.89$ for the within-group contrast). To minimize the workload for instructors, iCPR units were therefore designed without instructor feedback. (For a more detailed discussion of the design of iCPR units and the underlying research principles, see also Martin, 2017 and Martin, 2018).

Pronunciation Targets

Since iCPR units were designed for L1 English learners of L2 German, ten pronunciation targets were selected that all present difficulties for American learners of German and, when mispronounced, impede learners' intelligibility and readily give the L2 speaker away as being a native speaker of American English (Hall, 2003). Both segmental and suprasegmental features were addressed. The pronunciation targets were:

(a) allophones of /r/: consonantal [ʁ] and vocalic [ɐ],

(b) [ç] and [x],
(c) [e:],
(d) [o:],
(e) [y:],
(f) [ø:],
(g) the orthographic-phoneme-correspondence <z> → /ts/,
(h) the orthographic-phoneme-correspondences <ie> → /i:/ and <ei> → /aɪ/,
(i) the orthographic-phoneme-correspondence <v> → /f/, and
(j) lexical stress in German-English cognates.

Research Questions

The following research questions were addressed:

1. Do novice learners' L2 perception and production skills improve over the course of one semester of distance language instruction without targeted pronunciation training?
2. Is pronunciation training delivered through iCPR units effective in improving novice learners' L2 perception and production skills in distance language instruction?
3. How are iCPR units received among distance language learners?

Methods

Participants

All participants were enrolled in one of four sections of an undergraduate, first-semester German course taught in a distance learning environment in the US. The course did not include any F2F interaction and was asynchronous except for a spoken live chat component for a total of four hours over the course of the semester.[3] The four sections were taught by two different instructors: Each section was assigned to the treatment or control condition, so that each instructor taught one control and one treatment group. All parts of the data collection of this study were a mandatory component of students' coursework; however, students were only included as participants in the study if they (a) gave consent to participate, (b) had English as their L1, (c) submitted both a pre- and posttest of either the perception or production assessment or both, and (d) submitted at least 80% of the pronunciation homework assignments, if they were in the treatment group. The initial participant pool consisted of 67 participants, but the final pool was reduced to 34 participants due to the exclusion criteria above, with 20 participants in the treatment group (six female, fourteen male; mean age: 28.9 years; *SD*: 7.4; range: 18–44) and 14 participants in the control group (four female, ten male; mean age: 28.5 years; *SD* 4.8; range: 22–38). Learners were included in the analyses if they had completed at least one pair of the pre- and posttest: either the perception or the production assessment. This led to the following distribution: Treatment Group: $N = 20$; $n_{\text{perception}} = 17$; $n_{\text{production}} = 20$ and Control Group: $N = 14$; $n_{\text{perception}} = 13$; $n_{\text{production}} = 13$. All participants completed a language background questionnaire including a self-rated proficiency scale for German reading, writing, speaking, and listening. Independent sample *t*-tests (equal variance not assumed) revealed no significant difference in age or any of the self-rated proficiency measures between the two groups (Speaking: $t(32) = 1.08$, $p = .288$; Listening: $t(32) = 1.90$, $p = .067$; all other $t < 1$).

Treatment Materials

Pronunciation Treatment

There were 30 iCPR units, split into 10 sets of three units—each set addressing one of the 10 pronunciation targets introduced above. Each set was further split into one perception and two production units, with the perception unit appearing first. Each training item in the perception and production units was recorded by one of nine native speakers of German. A comprehensive example of the pronunciation treatment can be accessed on the IRIS database (www.iris-database.org).

Perception Units

The iCPR perception units consisted of two types of exercises: an Accentedness-Detection-Task (ADT) and a Sound-Discrimination-Task (SDT). The ADT was modelled after classroom-based perception treatments as in Botero (2011) and Roccamo (2014), with the goal of training learners to discern native from accented productions of problematic L2 speech sounds in real words and, thus, to encourage learners to self-monitor their own productions for these accent markers. Each perception unit contained 10 ADT practice items (see Figure 1; for a more detailed description of iCPR units, see Martin, 2017 and Martin, 2018).

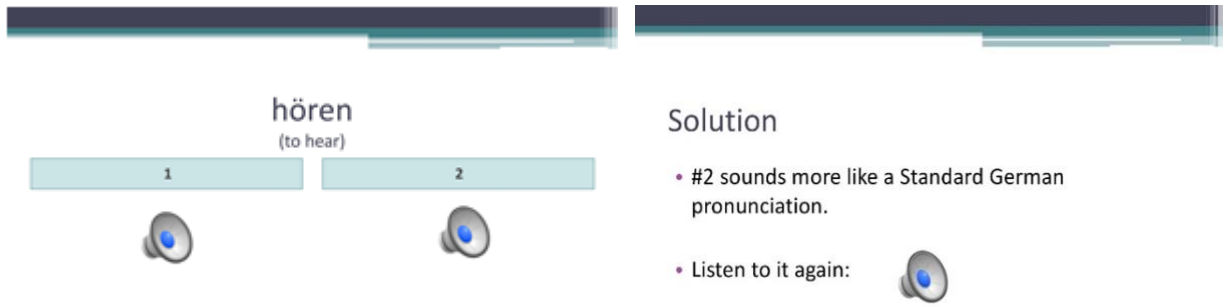

Figure 1. Sample slides for Accentedness-Detection-Task in perception iCPR units.

The SDT was modelled after lab-based perception treatments (Guion & Pederson, 2007; Thomson, 2011). It was used to train participants on the perception of problematic German sounds in contrast to similar sounds in American English. For this task, the L2 sounds were embedded in nonsense words, that is, words that follow phonotactic rules of German but do not carry meaning (see Figure 2). Each perception unit contained 10 SDT practice items.

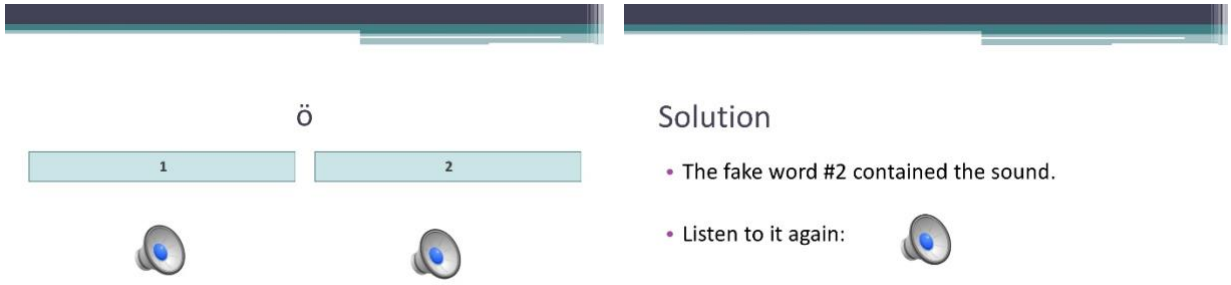

Figure 2. Sample slides for Sound-Discrimination-Task in perception iCPR units.

Production Units

The iCPR production units always started with explicit information about the week's pronunciation focus, with embedded native speaker recordings as examples, followed by the practice phase during which participants were prompted to imitate native speaker recordings of 12 German words containing the targeted sound. They were instructed to repeat the words out loud until they were satisfied with their own performance. At the end of each unit, participants were prompted to record themselves saying all of the practice items out loud (see Figure 3).

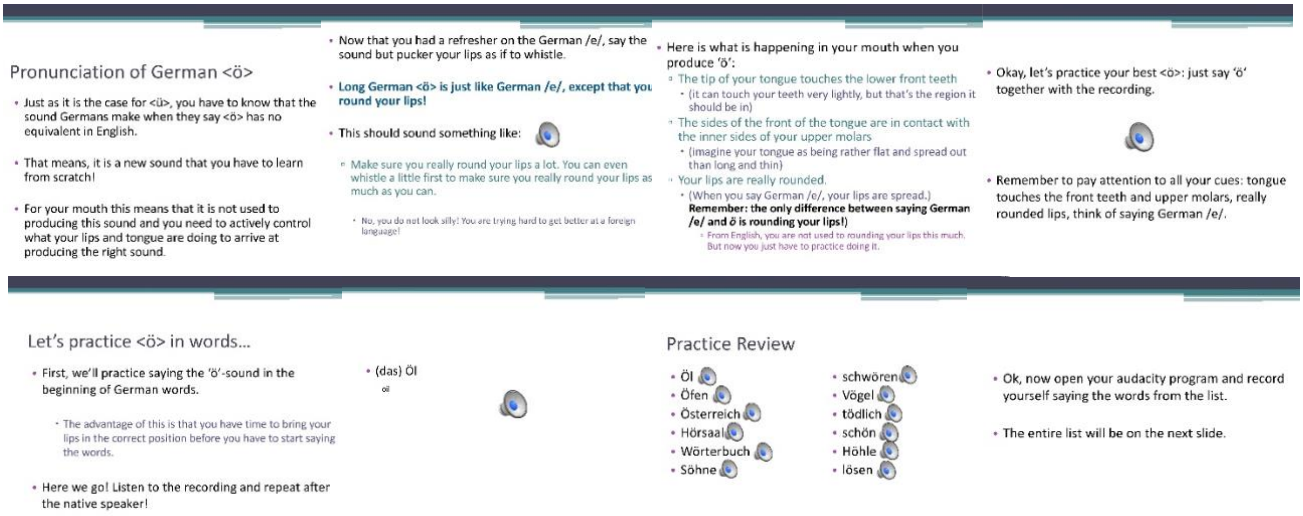

Figure 3. Sample slides for production iCPR units.

Testing Materials

Perception Test: The perception test was designed exactly as the perception treatment, only leaving out the automated feedback slide and adding a paper-and-pencil answer sheet on which participants were asked to record their responses. All items included in the perception test were unknown to the learners at the time of the pre- and posttest.[4]

Production Test: The production test was comprised of two tasks: A word-reading and a paragraph-reading task. The word-reading task consisted of 75 words, split into three blocks of 25 words each, plus two practice items. Participants were prompted to produce the words embedded in a carrier phrase. The paragraph-reading task consisted of six paragraphs of approximately 45 words each of coherent text.

Exit Questionnaire: The exit questionnaire was designed to elicit feedback from participants in the treatment group. It was comprised of three parts: (1) seven statements about the perceived usefulness of the iCPR pronunciation training that participants rated on a scale from 1 to 7, with one corresponding to *strongly disagree* and seven corresponding to *strongly agree*, (2) seven questions with limited response options targeting participants' behavior in working through the pronunciation activities, and (3) five open-ended questions (for a list of all questions, see Appendix A; for the entire questionnaire including all answer options, please see the original document on the IRIS database: www.iris-database.org).

Procedure

Two of the four intact sections of first-semester online German classes were pseudo-randomly selected to serve as a treatment group and the other two sections served as a control group. Participants in the treatment group received pronunciation instruction via iCPR units. Participants in the control group did not receive targeted pronunciation instruction, but followed the same curriculum and completed additional vocabulary, grammar, and culture exercises to ensure that both groups received the same amount of total German instruction. The pre- and posttests were administered during the second and 14th week of the semester in both groups. The perception tests were identical in both groups, but items were randomized within each task between testing times. For the production tests, the three blocks in which the items were presented were counterbalanced between groups, instructors, and testing time.

Between the pre- and the posttest, participants in the treatment group received 10 weeks of pronunciation training via iCPR units integrated in the course modules. Each unit was designed to take about 10 minutes of work, totaling approximately 30 minutes of pronunciation instruction per week for 10 weeks.

Scoring

Perception

The perception task was a binary forced-choice task. Participants received one point for choosing the correct answer and zero points for choosing the incorrect answer.

Production

Participants' pre- and posttest audio files were spliced into shorter parts using Audacity 2.0. For the word-level production, words were spliced from the carrier phrase. For the paragraph-level productions, the first 20 seconds of each paragraph-reading were extracted and initial disfluencies were removed (following O'Brien, 2014). All final audio files were normalized by scaling them to a peak intensity of 70db and by inserting 500ms of silence, using a Praat script (Boersma & Weenink, 2017). Forty-eight words and six paragraphs (matching between pre- and posttest) per person were randomly chosen to be rated by native speakers. An additional 48 words and six paragraphs recorded by eight native or near-native speakers were included in the ratings as a control measure. Following common rating procedures (Crowther, Trofimovich, & Isaacs, 2016; Derwing & Munro, 2013), four native-speaker raters were recruited to rate the participants' productions. None of the raters were teachers of German. Each rater rated 1,776 word productions and 222 paragraph productions, randomly distributed into 16 blocks of word ratings and 10 blocks of paragraph ratings that were presented counterbalanced between raters and rating time. Raters came to the lab for multiple rating sessions, each of which began with a brief training. They were then asked to judge the comprehensibility and accentedness of each word or paragraph on a 9-point Likert scale.[5] The raters heard each item only once before entering their rating (O'Brien, 2016). PsychoPy software (Peirce, 2016) was used to randomize stimuli order within each block. Inter-rater reliability between the four raters was calculated using intraclass correlation coefficients (ICCs). As shown in Table 1, inter-rater consistency was high (Cronbach's α =.89–.92), in all cases exceeding the recommended .70–.80 benchmark (Larson-Hall, 2016). This allowed for the calculation of mean comprehensibility and accentedness scores for each participant, averaging across the four raters. These average values for each participant then served as data points for the analyses.

Table 1. *Rater Consistency (Cronbach's α, [95% CI]) for Rated Continua by Task*

	Rater Consistency Cronbach's α, [95% CI]
Words	
Comprehensibility	.90 [.88, .92]
Accentedness	.91 [.89, .93]
Paragraphs	
Comprehensibility	.92 [.91, .93]
Accentedness	.89 [.88, .90]

Data Analysis

As not all of the data were normally distributed (according to Shapiro-Wilk's tests as recommended by Larson-Hall, 2016; ADT treatment group dataset p = .001) and variances were not equal, results of nonparametric tests and standardized estimates of the effect (Cohen's d) are presented.[6] Analyses were carried out separately for the Accentedness-Detection-Task, the Sound-Discrimination-Task, and for comprehensibility and accentedness ratings (further split into word- and paragraph-level productions). Pre-test scores between the treatment and the control group were compared using Mann-Whitney tests to confirm that there were no significant differences between the groups prior to the treatment. Posttest scores were then compared in the same manner to determine whether the treatment had led to differences in

perception and production skills between the groups. To further determine whether each group had made gains over time, pre- and posttest scores for each group were compared using Wilcoxon signed-rank tests. Following Plonsky and Oswald (2014), Cohen's *d* field-specific benchmarks were used for interpretation, specifically, for between-group comparison: *d* = 0.40 (small), 0.70 (medium), 1.0 (large) and for within-group comparison: *d* = 0.60 (small), 1.00 (medium), 1.4 (large).

The analyses for accentedness ratings showed the same pattern of significance and effect sizes on the word- and paragraph-level as the comprehensibility ratings. For the sake of brevity, only the results for comprehensibility ratings will thus be presented in the following.

Results

Perception

Accentedness-Detection-Task

For a visual representation of the ADT results, see Figure 4 (see Appendix B for descriptive statistics). Mann-Whitney tests did not show a significant difference between the treatment and control groups at the time of the pretest ($U = 86.00$, $z = -1.04$, $p = .320$, $d = 0.02$ [negligible effect]), but they revealed a significant difference at the time of the posttest ($U = 58.00$, $z = -2.24$, $p = .028$, $d = 0.94$ [medium effect]). Wilcoxon signed-rank tests revealed a significant difference between mean pretest and posttest scores for the treatment group ($z = -2.34$; $p = .019$, $d = 0.89$ [small effect]), but not for the control group ($z = -0.87$, $p = .385$, $d = 0.33$ [negligible effect]).

Sound-Discrimination-Task

For a visual representation of the SDT results, see Figure 5 (see Appendix B for descriptive statistics). Mann-Whitney tests did not show a significant difference between the treatment and control groups at the time of the pretest ($U = 101.50$, $z = -0.38$, $p = .711$, $d = 0.07$ [negligible effect]), but they revealed a significant difference at the time of the posttest ($U = 58.00$, $z = -2.23$, $p = .028$, $d = 1.00$ [large effect]). Wilcoxon signed-rank tests revealed a significant difference between mean pretest and posttest scores for the treatment group ($z = -2.27$; $p = .023$, $d = 1.03$ [medium effect]), but not for the control group ($z = -1.02$, $p = .309$, $d = 0.20$ [negligible effect]).

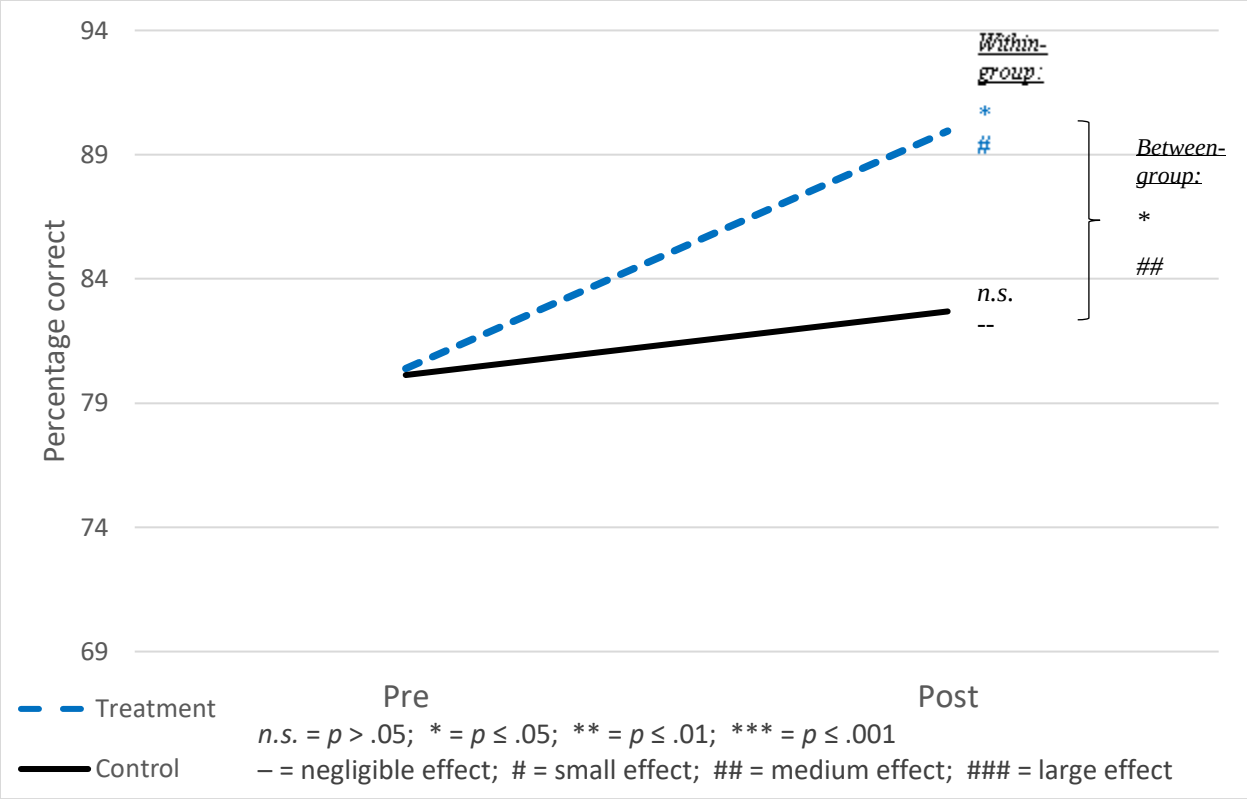

Figure 4. Results of Accentedness-Detection-Task.

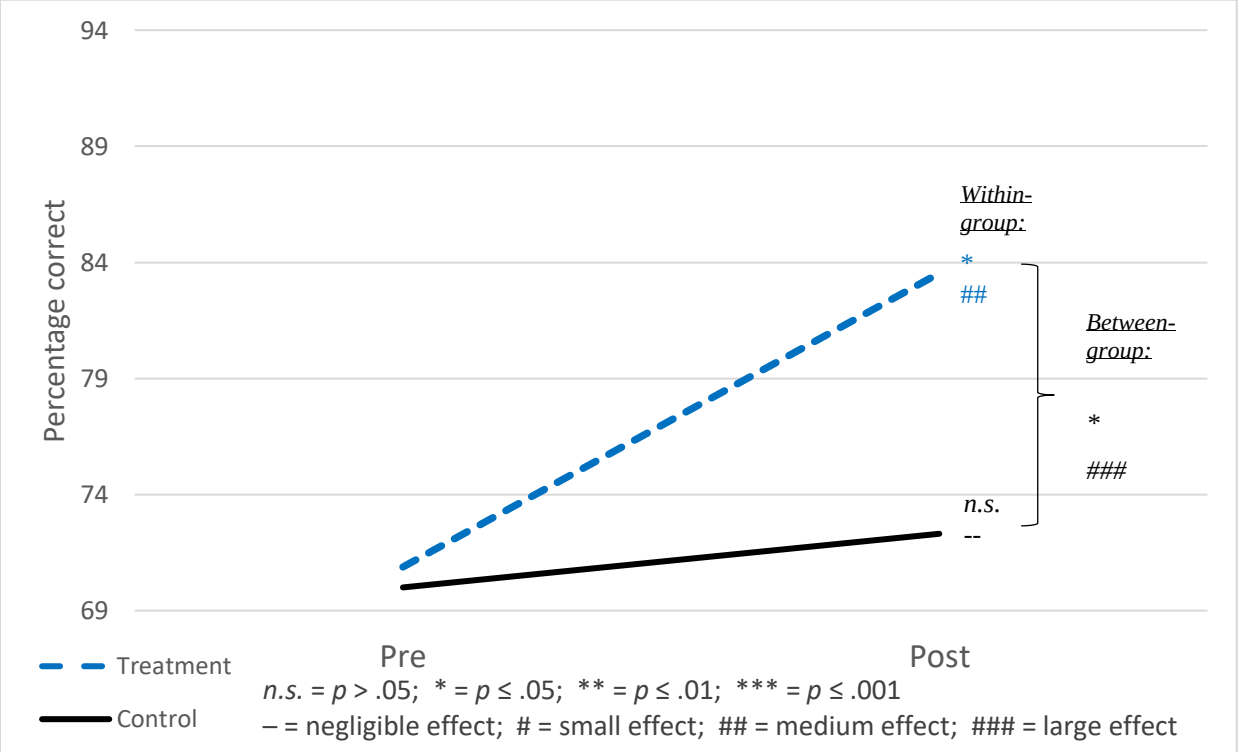

Figure 5. Results of Sound-Discrimination-Task.

Production

Word-Level Productions

For a visual representation of the word-level production results for comprehensibility, see Figure 6 (see Appendix B for descriptive statistics). Mann-Whitney tests did not show a significant difference between the treatment and control groups at the time of the pretest (U = 127.00, z = -0.11, p = .928, d = 0.04 [negligible effect]), but they revealed a significant difference at the time of the posttest (U = 73.50, z = -2.08, p = .036, d = 0.80 [medium effect]). Wilcoxon signed-rank tests revealed a significant difference between mean pretest and posttest scores for the treatment group (z = -3.70; p < .001, d = 1.02 [medium effect]), but not for the control group (z = -1.43, p = .152, d = 0.23 [negligible effect]).

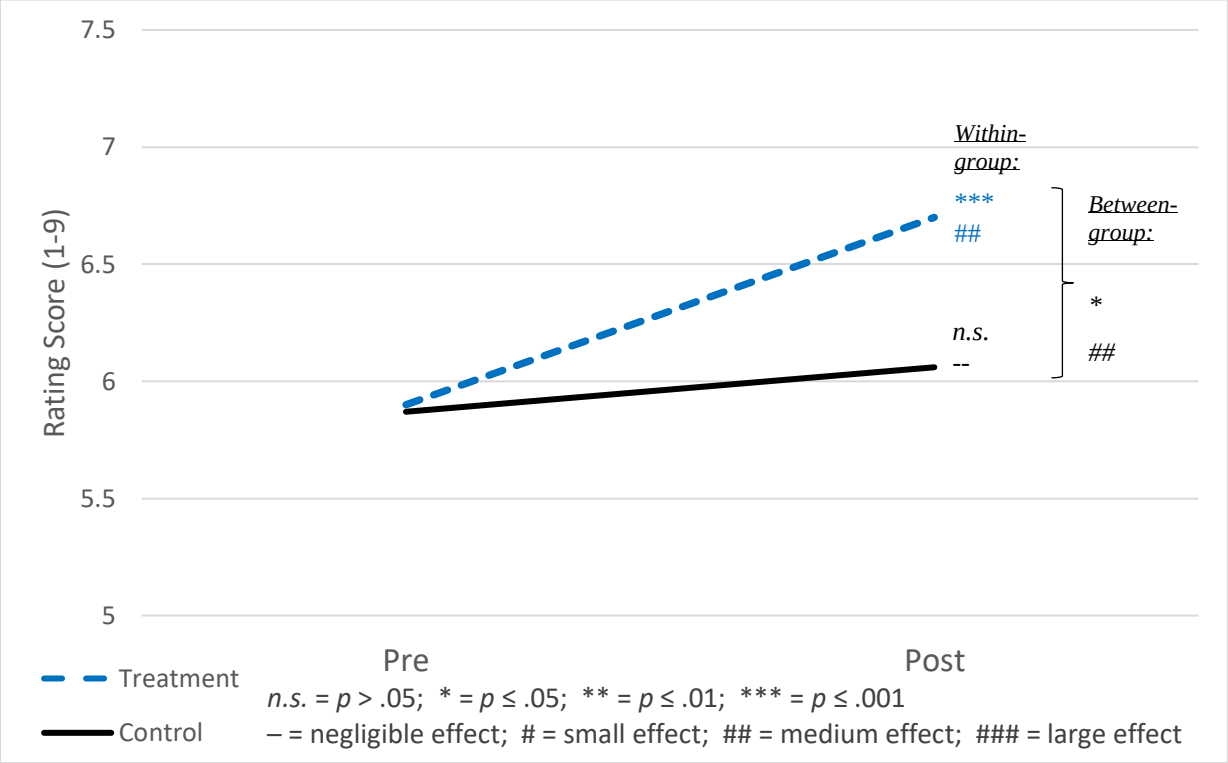

Figure 6. Results of word-level productions.

Paragraph-Level Productions

For a visual representation of the paragraph-level production results for comprehensibility, see Figure 7 (see Appendix B for descriptive statistics). Mann-Whitney tests did not show a significant difference between the treatment and control groups at the time of the pretest (U = 122.50, z = -0.28, p = .785, d = 0.01 [negligible effect]), but they revealed a significant difference at the time of the posttest (U = 69.00, z = -2.25, p = .024, d = .75 [medium effect]). Wilcoxon signed-rank tests revealed a significant difference between mean pretest and posttest scores for the treatment group (z = -3.92; p < .001, d = 1.09 [medium effect]), but not for the control group (z = -1.29, p = .196, d = 0.49 [negligible effect]).

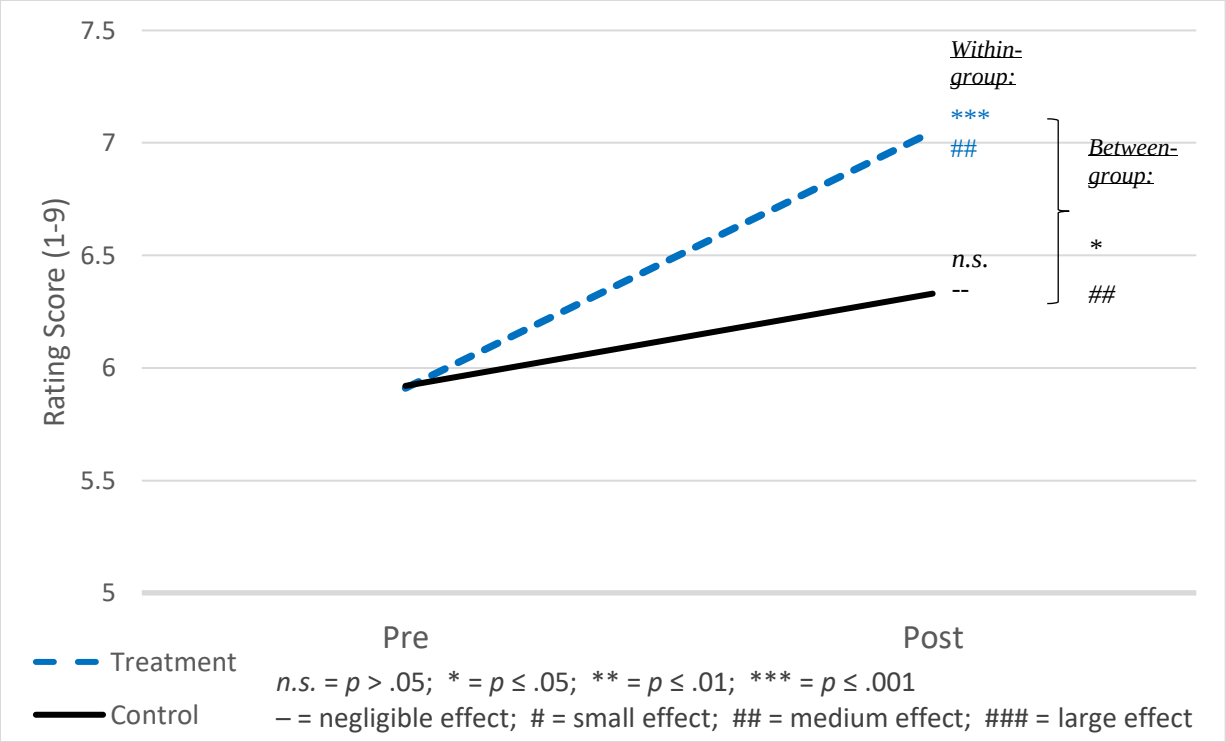

Figure 7. Results of paragraph-level productions.

Exit Questionnaire Responses

Seventeen of the 20 learners in the treatment group completed the Exit Questionnaire. Table 2 summarizes the mean scores for the six statements targeting feedback on the perceived usefulness of the iCPR pronunciation training that made up the first part of the questionnaire. Overall, results showed that participants liked the pronunciation training units and that they felt their pronunciation improved through the training.

Table 2. *Mean Scores on Exit Questionnaire Responses*

Question No.	Statement To Be Rated By Participants	Mean Response	Range
1.	I enjoyed working on the pronunciation homework.	5.2	3–7
2.	I feel like I learned something from the pronunciation homework.	5.7	2–7
3.	I feel like my pronunciation of German improved through the homework exercises.	5.9	3–7
4.	Any technological problems that I encountered were worth the overall benefit.	6.5	5–7
5.	Not only did the pronunciation homework improve my pronunciation but it also helped review some vocabulary.	5.5	3–7
6.	I would have taken the pronunciation homework more seriously and would have benefited more from it if I had received teacher feedback throughout the semester.	4.5	1–7

Note. Statements were rated on scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

The responses on the second part of the questionnaire revealed that the majority of the learners (10 of 17) estimated having spent 15–20 minutes on each iCPR unit. The largest number of participants (i.e., the mode for each response in the data set) further reported:

- Listening to each recording 2–3 times (8 of 17 participants),
- Practicing saying each word 2–5 times (8 of 17 participants),
- Practicing saying each word until they were completely happy with their own pronunciation before recording themselves (12 of 17 participants),
- Liking the speaking exercises more than the listening exercises (8 of 17 participants), and
- Thinking three units a week was a good amount of practice (14 of 17 participants).

Since there was no control of whether the participants completed the listening exercises (as opposed to the speaking exercises, students did not have to submit a recording for the perception units), they were also asked how many of the 10 listening units they had actually completed. Twelve of 17 participants reported having completed all 10 and another two reported completing 9 out of 10 units. Three learners reported having completed fewer than five units.

The third part of the questionnaire consisted of five open-ended questions. In response to the question "What did you like about the pronunciation homework exercises?", participants named the metalinguistic explanations as being helpful, reported that they simply liked receiving pronunciation practice at all, or that they enjoyed the additional vocabulary review embedded in the units. In response to the question "What did you not like about the pronunciation exercises?," eight of 17 participants wrote that they could not think of anything they did not like about the iCPR units. However, some learners mentioned that the technological component could be improved and two learners perceived the lack of instructor feedback on their pronunciation homework as negative. In response to the question "Do you have any suggestions for future improvements of the pronunciation exercises? If so, what are they?," answers varied widely. Participants gave suggestions for technological improvements, course structure, and the practice of vocabulary items. In response to the question "What, if any, are the areas of pronunciation not treated in the pronunciation exercises that you would have liked to work on and improve?," ten learners responded that they did not miss anything, others suggested more review and more words in sentences as opposed to words in isolation.

Discussion

The present study investigated how distance language learners' pronunciation skills develop with and without targeted pronunciation training and explored a new method of teaching L2 pronunciation in this unique learning environment. It further elicited and evaluated learners' opinions and feedback on this new method of pronunciation training.

Regarding the first research question, data revealed that participants in the control group—that is, learners that just followed the standard curriculum without targeted pronunciation training—did not show a significant improvement in their perception and production skills from pre- to posttest, and effect sizes for these gains were negligible in all conditions. This finding suggests that simply following the standard curriculum did not suffice to improve the participants' pronunciation skills over the course of the semester. This is a discouraging and somewhat surprising finding, since the participants improved their overall German proficiency—if they had not done so, they would not have passed the class.[7]

Unfortunately, the lack of other studies on the development of pronunciation training in distance language learning does not allow for a direct comparison of this finding to previous research. The scarce previous literature only addresses oral proficiency development as a whole, but not pronunciation development specifically (Lin, 2014). One exception is Inceoglu's (2017) unpublished pilot study presented at the *9th Pronunciation in Second Language Learning and Teaching Conference*, which investigated the effect of pronunciation training on the development of learners' pronunciation in an asynchronous online French class at the lower intermediate level. The pronunciation training employed in Inceoglu's study was

comprised of different computer-assisted exercises, such as discrimination activities, IPA transcriptions, and oral recordings. Inceoglu found significant improvement in segmental productions, suggesting that learners in the online French class benefited from the pronunciation training. Yet, since the study did not include a control group, again we have no measure to gauge how learners' pronunciation skills develop without pronunciation intervention. However, considering that previous research has shown that distance language learners' oral proficiency develops on par with F2F learners' oral proficiency (Blake et al., 2008; Lin, 2014) and that novice learners' pronunciation skills improve significantly over the course of a semester even without targeted pronunciation training when instructed in a F2F environment (Martin, 2018), we would have expected distance learners' pronunciation skills to improve significantly as well. The most striking difference between instruction in the distance and the F2F environment is, of course, the absence of oral in-class interaction in the distance learning environment—both between the instructor and the students, but also between peers. It is then possible that distance learners' pronunciation skills suffer in the absence of oral interaction because learners receive less and less varied aural input and are thus limited in their ability to establish a robust L2 sound system (Thomson, 2011, 2012).

Moreover, research has shown that learners acquire language more quickly when they are pushed to produce it (Izumi, 2002). With the lack of classroom interaction, learners are pushed less often to produce spoken language and receive less feedback on whether their productions are felicitous for communication, which in turn might explain why their pronunciation skills are slow to improve (Nielson et al., 2009). This situation could possibly be remedied by introducing more live oral chats in the online curriculum, thus giving learners additional input as well as encouraging them to speak in the L2 more often. However, doing so affects the nature of distance classes: that is, for many learners, the asynchronous nature of distance learning is a main advantage of this form of instruction, and synchronous live chats present an obstacle for learners that take distance courses particularly because their professional or family responsibilities do not allow them to take synchronous classes. Thus, it is important to find another way to help learners improve their pronunciation skills in a distance learning environment.

To this end, the second part of the present study investigated the effectiveness of iCPR pronunciation training in improving distance learners' L2 perception and production skills (Research Question 2). Data revealed that participants in the treatment group—that is, learners who received 10 weeks of pronunciation training—showed a significant improvement from pre- to posttest on all measures of perception and production skills with medium effect sizes for these gains. Data further revealed that participants in the treatment group significantly outperformed participants in the control group on all measures of pronunciation skills at the end of the semester with medium-to-large effect sizes for this comparison.

These findings suggest that the iCPR units helped the participants improve their pronunciation skills over the course of the semester and thus that iCPRs offer a viable form of pronunciation intervention in distance language classes. Given that the ultimate goal of pronunciation instruction is, of course, improvement on all levels of speech, it is particularly promising that participants in the treatment group improved significantly in their comprehensibility and accentedness of productions at the word level and at the paragraph level. This is especially meaningful considering that the training only targeted the word level. Previous research has often found less robust improvement and smaller effect sizes for pronunciation assessment on longer productions, possibly due to the fact that uttering longer productions consumes more mental resources, involving lexical, morphological, and syntactic processing, than productions at the word level, thereby possibly limiting mental resources to focus on pronunciation (Lee et al., 2015). Furthermore, with an average effect size of Cohen's $d = 1.01$ for the iCPR within-group comparison and an average effect size of $d = 0.87$ for the comparison between the treatment and control group at the posttest, iCPR training exhibited much larger effects than other CAPT, as reported in Lee et al.'s meta-analysis (average of CAPT studies in the meta-analysis: Within-group $d = 0.75$, between-group $d = 0.24$). According to Coe's (2002) effect-size-to-percentile-interpretation table, this between-group effect size of $d = 0.87$ means that the average participant in the treatment group scored higher than 79% of participants in the control group. This result, taken together with the fact that iCPR units are mostly course-independent and can simply be added-on to any existing curriculum, makes them a promising possibility to add a pronunciation component to any

distance learning curriculum that lacks targeted pronunciation instruction.

The quantitative findings showing that learners benefited from the instruction through iCPR units were further reflected in the participants' qualitative responses to the exit questionnaire (Research Question 3). Learners' responses showed that they enjoyed receiving pronunciation training in general, but also that they particularly liked working with the iCPR units and felt that iCPRs helped them improve their pronunciation. This suggests a high degree of face validity for the iCPR pronunciation treatment, which is a desirable educational goal in foreign language instruction and can lead to improved learner motivation (Brown & Abeywickrama, 2010). With an average time of 15–20 minutes that the participants self-reported working on each iCPR unit, they appear to have spent more time than anticipated on the activities (the units were designed to take 10–15 minutes), and also spent more time than learners in a F2F environment report spending on the same activities (Martin, 2018). However, this is in line with the participants' other responses, which suggest that they took the pronunciation training very seriously: They reported listening to each recording 2–3 times and practicing each word out loud 2–5 times before recording it. Considering that the participants did not receive a grade for these activities, the finding that they spent extra time on practice further shows that improving their pronunciation is important to them and that they valued the opportunity for targeted pronunciation training.

Conclusions

The present study investigated whether distance language learners' pronunciation skills improve in a first-semester German class when simply following the standard curriculum. Furthermore, a new method of pronunciation instruction in distance language classes was introduced, and its effectiveness as well as learners' perception of the method were assessed. The results suggest that simply following the standard curriculum did not suffice to significantly improve learners' pronunciation skills. Results further showed that targeted pronunciation training through the iCPR method allowed learners to significantly improve their perception and production skills in a distance learning environment and that learners enjoyed working with the iCPR pronunciation training. Since iCPR units are designed course-independently, they offer a promising possibility to add a pronunciation component to any existing curriculum.

The pedagogical implications of this study are substantial. Considering the importance of pronunciation for successful communication in the L2 (Derwing & Munro, 2015), it is a discouraging finding that novice distance learners' L2 pronunciation skills did not improve significantly over the course of the semester in the absence of targeted pronunciation training. This stresses the necessity of providing a pronunciation intervention in distance language classes. iCPRs offer one such method.

The study leaves several avenues for future research. For instance, this study did not explore if iCPR units also help learners improve their pronunciation in free speech contexts, when processing costs put even more strains on learners' productions. Moreover, testing did not assess whether the significant gains seen in the treatment group were sustained over time, that is, during the second semester of German or longer. A delayed posttest could help shed light on this question. Finally, future research should explore distance language learners' pronunciation development at different levels of L2 proficiency. In order to assess when a pronunciation intervention is most helpful to the learners, it is important to learn more about the development of distance learners' pronunciation skills more generally.

Acknowledgements

This research was supported in part by a Language Learning Dissertation Grant, a RGSO Dissertation Support Grant, a Gil Watz Dissertation Grant, and a TOEFL Small Grant for Doctoral Research in Second or Foreign Language Assessment. I would also like to thank my dissertation committee at The Pennsylvania State University for their support in the design and execution of this research project. Finally, I am grateful to Mary Grantham O'Brien and Liese Sippel as well as to the three anonymous reviewers for their helpful

comments on this manuscript.

Notes

1. The terms *distance* and *online* courses are synonymous and refer to fully virtual courses in which no face-to-face interaction takes place—but rather interaction occurs at a distance through various technological tools (Nielson, Gonzalez-Lloret, & Pinckney, 2009).
2. Oral proficiency entails speaking ability and proficiency more generally—addressing parameters such as fluency, accuracy, lexical and syntactic skills, or holistic performance in oral production, and not merely focusing on pronunciation skills (Lin, 2014).
3. The learners used the coursebook *Sag mal* (Anton, Barske, Grabowski, & McKinstry, 2014) that comes with website access and activities that include about ten minutes of listening and speaking activities a week. These activities were auto-graded and learners did not receive additional pronunciation feedback in either group.
4. An exception is the cognate stress condition: The instructional goal here was not a transferrable skill but specific training for placing correct stress in high frequency German-English cognates.
5. Since intelligibility and comprehensibility are highly correlated (Derwing & Munro, 2015) and intelligibility ratings were not feasible with this number of participants, this study followed the methodology of Gordon and Darcy (2016) and Bergeron and Trofimovich (2017), and only collected comprehensibility ratings as a general metric of understanding.
6. For parity with other studies, parametric tests (independent sample *t*-tests and paired *t*-tests) were run as well. They showed the same patterns of findings as the analyses carried out with nonparametric tests.
7. See also Isenberg, 2010 who compared distance and F2F learners' German proficiency at the same university as the present study and found comparable development in vocabulary and grammar learning, as well as oral proficiency between the two learning environments.

References

Anton, C., Barske, T., Grabowski, J. & McKinstry, M. (2014). *Sag mal. An introduction to German language and culture*. Boston, MA: Vista Higher Learning.

Baralt, M., & Morcillo Gómez, J. (2017). Task-based language teaching online: A guide for teachers. *Language Learning & Technology, 21*(3), 28–43. http://dx.doi.org/10125/44630

Bergeron, A., & Trofimovich, P. (2017). Linguistic dimensions of accentedness and comprehensibility: Exploring task and listener effects in second language French. *Foreign Language Annals, 50*(3), 547–566.

Blake, C. (2009). Potential of text-based internet chats for improving oral fluency in second language. *The Modern Language Journal, 93*(2), 227–240.

Blake, R., Wilson, N. L., Cetto, M., & Pardo-Ballester, C. (2008). Measuring oral proficiency in distance, face-to-face, and blended classrooms. *Language Learning & Technology, 12*(3), 114–127. http://dx.doi.org/10125/44158

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.09. Retrieved from http://www.praat.org

Botero, C. G. (2011). *Factors underlying the improvement of L2 phonological production using technology outside the classroom.* (Unpublished doctoral dissertation). College of the Liberal Arts, The Pennsylvania State University, University Park, PA.

Brown, H. D., & Abeywickrama, P. (2010). *Language Assessment: Principles and classroom practices* (2nd ed.). White Plains, NY: Pearson.

Cheng, T.-J. (2015). *The reality behind the hype – Online world language teaching and instructional design.* (Unpublished doctoral dissertation). Teachers College, Columbia University, New York City, NY.

Chenoweth, N. A., Jones, C. M., & Tucker, G. R. (2006). Language online: Principles of design and methods of assessment. In R.P. Donaldson & M.A. Haggstrom (Eds.), *Changing language education through CALL*. (pp. 1–21). Abingdon, UK: Routledge.

Coe, R. (2002). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at The 2002 Annual Conference of the British Educational Research Association, Exeter, UK.

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect, 18*(3), 53–70.

Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility. Nonnative listeners' perspectives. *Journal of Second Language Pronunciation, 2*(2), 160–182.

De Paepe, L., Zhu, C., & Depryck, K. (2018). Learner characteristics, learner achievement and time investment in online courses for Dutch L2 in adult education. *The Turkish Online Journal of Educational Technology*, *17*(1), 101–112.

Derwing, T. M. (2013). Pronunciation instruction. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–9). Chichester, UK: Wiley-Blackwell.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research based approach. *TESOL Quarterly*, *39*(3), 379–397.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Philadelphia, PA: John Benjamins.

Deutschmann, M., Panichi, L., & Molka-Danielsen, J. (2009). Designing oral participation in Second Life – a comparative study of two language proficiency courses. *ReCALL*, *21*(2), 206–226.

Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, *7*(3), 50–80. http://dx.doi.org/10125/25214

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J.I. Hualde (Eds.), *Laboratory Phonology 9*, (pp. 353–382). New York, NY: Mouton de Gruyter.

Gordon, J., & Darcy, I. (2016). The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *Journal of Second Language Pronunciation, 2*(1). 56–92.

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M.J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 57–77). Amsterdam, The Netherlands: John Benjamins.

Hall, C. (2003). *Modern German pronunciation: An introduction for speakers of English.* Manchester, UK: Manchester University Press.

Hsu, L. (2016). An empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning, 29*(5), 881–900.

Huensch, A. (2016). Perceptual phonetic training improves production in larger discourse contexts. *Journal of Second Language Pronunciation, 2*(2). 183–207.

Huensch, A. (2018). Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Language Teaching Research*. Advance online publication.

Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes toward pronunciation: Foreign Language Learners in the United States. *Foreign Language Annals, 50*(1), 410–432.

Inceoglu, S. (2017, September). *The development of L2 French learners' pronunciation, fluency, and comprehensibility: an online classroom study*. Paper presented at the 9th Annual Conference for Pronunciation in Second Language Learning & Teaching, Salt Lake City, Utah.

Isenberg, N. A. (2010). *A comparative study of developmental outcomes in web-based and classroom-based German language education at the post-secondary level: Vocabulary, grammar, language processing, and oral proficiency development*. (Unpublished doctoral dissertation). College of the Liberal Arts, The Pennsylvania State University, University Park, PA.

Izumi, S. (2002). Output, input enhancement, and the Noticing Hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541-577.

Kissling, E. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal, 97*(3), 720–744.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*(2), 227–247.

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R.* (2nd ed.). London and New York: Routledge.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345–366.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369–377.

Levis, J. M. (2016a). Accent in second language pronunciation research and teaching. *Journal of Second Language Pronunciation, 2*(2). 153–159.

Levis, J. M. (2016b). Thinking allowed. Research into practice: How research appears in pronunciation teaching materials. *Language Teaching, 49*(3), 423–437.

Lin, H. (2014). Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis. *ReCALL, 27*(3), 261–287.

Martin, I. A. (2017). Teaching pronunciation through homework assignments: The method of iCPRs. In M. O'Brien & J. Levis (Eds.), *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference.* (pp. 216–223). Ames, IA: Iowa State University.

Martin, I. A. (2018). *Bridging the gap between L2 pronunciation research and teaching: Using iCPRs to improve novice German learners' pronunciation in distance and face-to-face classrooms.* (Unpublished doctoral dissertation). College of the Liberal Arts, The Pennsylvania State University, University Park, PA.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 285–310.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*(4), 520–531.

Nielson, K. B., Gonzalez-Lloret, M., & Pinckney, K. (2009). *Learning foreign languages at a distance: Characteristics of effective online courses*. College Park, MD: University of Maryland Center for Advanced Study of Language.

O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning, 64*(4), 715–748.

O'Brien, M. G. (2016). Methodological choices in rating speech samples. *Studies in Second Language Acquisition, 38*(1), 587–605.

Olson, D. J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning & Technology, 18*(3), 173–192. http://dx.doi.org/10125/44389

Peirce, J. W. (2016). PsychoPy. Psychology software in Python [Computer program]. Version 1.84.1. Retrieved from http://www.psychopy.org

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912.

Roccamo, A. (2014). *Comparing the success of pronunciation instruction in elementary and intermediate German language classrooms*. (Unpublished doctoral dissertation). College of the Liberal Arts, The Pennsylvania State University, University Park, PA.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: a synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly, 46*(4), 842–854.

Saito, K. (2015). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Applied Psycholinguistics, 36*(2), 377–409.

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.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics, 20,* 213–223.

Seaman, J. E., Allen, I. E., & Seaman, J. (2018). *Grade increase: Tracking distance education in the United States.* Babson Park, MA: Babson Survey Research Group.

Smith, L. C. (2001). L2 acquisition of English liquids: Evidence for production independent from perception. In X. Bonch-Bruevich, W.J. Crawford, J. Hellermann, C. Higgins, & H. Nguyen (Eds.), *The Past, Present, and Future of Second Language Research: Selected Proceedings of the 2000 Second Language Research Forum* (pp. 3–22). Sommerville, MA: Cascadilla Press.

Tanner, M. W., & Landon, M. M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology, 13*(3), 51–65. http://dx.doi.org/10125/44191

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal, 28*(3), 744–765.

Thomson, R. I. (2012). Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Language Learning, 62*(4), 1231–1258.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics 36*(3), 326–244.

Thomson, R. I., & Isaacs, T. (2009). Within-category variation in L2 English vowel learning: Proceedings of the annual conference on the Canadian Acoustics Association. *Canadian Acoustics, 37*(3), 138–139.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation on Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*(2), 1033–1043.

Weinberg, A., & Knoerr, H. (2003). Learning French pronunciation: Audiocassettes or multimedia? *CALICO Journal, 20*(2), 315–336.

Appendix A. List of Questions on Exit Questionnaire

The entire exit questionnaire with all answer options can be accessed on the IRIS database: www.iris-database.org.

Questions to rate on a Likert scale from 1 to 7:

1. I enjoyed working on the pronunciation homework.

2. I feel like I learned something from the pronunciation homework.

3. I feel like my pronunciation of German improved through the homework exercises.

4. Any technological problems that I encountered (e.g. downloading the PPT slides, recording myself etc.) were worth the overall benefit.

5. Not only did the pronunciation homework improve my pronunciation but it also helped review some vocabulary.

6. I would have taken the pronunciation homework more seriously and would have benefited more from it if I had received teacher feedback throughout the semester.

Questions with limited response options:

1. How much time, on average, did you spend on each pronunciation homework?

2. How often, on average, did you listen to the native speaker recording of each word?

3. How often, on average, did you practice saying each word before you recorded yourself on all the words?

4. Did you always practice saying each word until you were completely happy with your own pronunciation before you recorded yourself?

5. There was no control of whether you completed the listening exercises (= the first pronunciation homework of the week). Honestly, how many of the 10 listening exercise homework units did you complete?

6. Did you enjoy the listening exercises or the speaking exercises more?

7. Do you think that 3 days of homework practice on each pronunciation problem was enough to help you or would you have preferred more or less practice?

Open-ended questions:

1. What did you like about the pronunciation homework exercises?

2. What did you not like about the pronunciation homework exercises?

3. Do you have any suggestions for future improvements of the pronunciation exercises? If so, what are they?

4. What, if any, are the areas of pronunciation *not* treated in the homework exercises that you would have liked to work on and improve?

5. Any final comments?

Appendix B. Perception and Production Results: Descriptive Statistics for Non-Parametric Tests

	Pretest					**Posttest**				
	M	***SD***	**95% *CI***	***Mdn***	***IQR***	***M***	***SD***	**95% *CI***	***Mdn***	***IQR***
Perception – Accentedness-Detection-Task										
Treatment (n = 17)	80	13.3	[73.54, 87.24]	83.3	8.3	89.95	7.4	[86.16, 93.74]	87.5	10.4
Control (n = 13)	80	7.45	[75.63, 84.63]	79.2	10.4	82.69	8.1	[77.78, 87.61]	83.3	12.5
Perception – Sound-Discrimination-Task										
Treatment (n = 17)	71	9.13	[62.90, 78.87]	70	30	83.53	7.9	[79.49, 87.57]	85	15
Control (n = 13)	70	8.17	[65.07, 74.93]	70	12.5	72.31	15	[63.53, 81.08]	70	22.5
Production – Comprehensibility on the word-level										
Treatment (n = 20)	5.9	0.84	[5.50, 6.29]	6.02	1.06	6.7	0.7	[6.36, 7.04]	6.74	1.1
Control (n = 13)	5.9	0.7	[5.44, 6.29]	5.83	1.15	6.06	0.9	[5.50, 6.62]	6.23	1.34
Production – Comprehensibility on the paragraph-level										
Treatment (n = 20)	5.9	1.04	[5.43, 6.40]	6	1.31	7.05	1	[6.56, 7.54]	7.33	1.38
Control (n = 13)	5.9	0.82	[5.43, 6.42]	5.83	1.29	6.33	0.8	[5.82, 6.83]	6.33	1.38

About the Author

Ines A. Martin is an Assistant Professor in the Languages and Cultures Department at the United States Naval Academy in Annapolis, MD. She teaches German, French, and Linguistics. Her research specializes on instructed second language acquisition, with a focus on second language pronunciation, CAPT, and peer corrective feedback.

E-mail: imartin@usna.edu