education sciences

Article

From Research in the Lab to Pedagogical Practices in the EFL Classroom: The Case of Task-Based Pronunciation Teaching

Joan C. Mora * and Ingrid Mora-Plaza

Department of Modern Languages and Literatures and English Studies, Universitat de Barcelona, 08007 Barcelona, Spain; imoraplaza@ub.edu
* Correspondence: mora@ub.edu

Abstract: Input and context-related factors identified by research as key success variables in L2 pronunciation development in immersion contexts play a very modest role in instructed foreign language (FL) learning environments. Scarce L2 exposure and use and L1-accented input make pronunciation learning extremely challenging. Current L2 speech learning models attribute difficulties in L2 speech acquisition to L2-to-L1 perceptual sound mappings guided by L1-based perception and poor phonological awareness and noticing of cross-language phonetic differences, which are typically not adequately addressed in instruction through pedagogic tasks. Explicit and incidental pronunciation teaching methods have been found effective at improving learners' pronunciation, but ways to integrate them into communicative approaches to language teaching are still largely unexplored. Thus, language education practices currently lack a research-informed pedagogical approach that incorporates principles of L2 speech learning and task-based language teaching (TBLT) into pronunciation instruction. This article (1) presents an outline of new avenues for research and practice in L2 pronunciation instruction and (2) reports on the findings of an empirical study that implemented a task-based pronunciation teaching (TBPT) approach to teaching a difficult L2 vowel contrast through computerized collaborative map tasks that could be easily integrated into communicative FL classrooms.

Keywords: task-based language teaching (TBLT); L2 pronunciation instruction; L2 pronunciation training; task-based pronunciation teaching (TBPT); form-focused communicative instruction; L2 vowel perception and production; map task

check for updates

Citation: Mora, J.C.; Mora-Plaza, I. From Research in the Lab to Pedagogical Practices in the EFL Classroom: The Case of Task-Based Pronunciation Teaching. *Educ. Sci.* **2023**, *13*, 1042. https://doi.org/10.3390/educsci13101042

Academic Editor: Martin Howard

Received: 31 July 2023
Revised: 14 October 2023
Accepted: 15 October 2023
Published: 17 October 2023

1. Introduction

Pronunciation appears to be one of the prevailing learning challenges for L2 learners [1] and a teaching challenge for L2 teachers [2,3], especially in instructed second language acquisition (ISLA) contexts [4]. Still, both learners and teachers find pronunciation to be a crucial component of the communicative competence they wish to develop to highly proficient levels [3,5]. Despite the dramatic growth of the field of L2 speech learning and L2 pronunciation teaching and learning in recent years [6–9] and recent research exploring alternative methods to L2 pronunciation teaching beyond explicit pronunciation instruction, such as task-based language teaching (TBLT) approaches to pronunciation-focused instruction [10–15] or computer-assisted pronunciation training [16–18], there is still a dearth of research on the acquisition of L2 pronunciation in ISLA contexts, including English as a foreign language (EFL) learning contexts in Spain.

Despite the challenging aspects of learning L2 pronunciation in an ISLA setting, several meta-analyses have shown that L2 pronunciation instruction is effective [19–22] and that phonetic training of specific L2 speech sounds improves L2 learners' perception and production of difficult L2 sounds and sound contrasts [23–25]. However, enhancing the outcomes of pronunciation instruction through methods that enhance learners' attention to phonetic form and, at the same time, can be fully integrated into communicative teaching practice remains a real challenge, as communication is primarily a meaning-oriented activity [3]. In addition, research-based approaches to L2 pronunciation instruction based on the intelligibility principle [26,27], in line with current L2 speech learning models [28,29], suggest that pronunciation instruction should focus on target pronunciation features with high functional load that are difficult to acquire due to L2-to-L1 perceptual sound mappings guided by L1-based perception. In other words, pronunciation instruction methods should consider L1-specific learning difficulties (at the segmental and suprasegmental level) and enhance phonological awareness and noticing of cross-language differences in addition to L2-specific phonetic differences between contrasting phonetic features. Implementing these teaching goals within communicative language learning requires large amounts of creativity in task design and implementation to enhance learners' attention to linguistic form (or phonetic form) while learners are performing communicative language learning tasks. A TBLT approach to pronunciation instruction, task-based pronunciation teaching (TBPT), may prove successful at achieving these task design goals. We illustrate this through the empirical study we report on in Section 4. Given the specific low-input conditions associated with ISLA and its well-attested limiting effects on the development of oral skills [30], pronunciation development is unlikely to take place without pronunciation training techniques providing individual intensive practice on specific (e.g., segmental contrasts) and global (e.g., intelligibility) dimensions of pronunciation proficiency. Such techniques (see Section 3 below) could be provided as a supplement to and in combination with TBPT pedagogic interventions.

2. L2 Speech Models and FL Pronunciation Learning

The acquisition of L2 phonetics and phonology in adulthood has been shown to be profoundly influenced by learners' L1 perception [31]. Speech learning theories have attributed learners' capacity to categorize L2 sounds accurately to their ability to discern differences between L1 and L2 sounds and their degree of phonological awareness and noticing of cross-language phonetic differences.

For example, Ref. [29]'s revised Speech Learning Model (SLM-r) argues that L2 sounds that are perceptually dissimilar to L1 sounds are easier to acquire and a separate L2 category may be created for them (e.g., L2 English /^{3}:/ having no clear match in L1 Catalan/Spanish) than L2 sounds that are indistinguishable from L1 sounds, for which a distinct category will not be formed (e.g., L2 English /f/ mapped onto L1 Catalan/Spanish /f/). These two scenarios are not predicted to pose any learning problems to L2 learners. However, for L2 sounds that are highly similar perceptually to L1 sounds, learners are unlikely to create separate L2 categories (e.g., L2 English /ʌ/ mapped onto L1 Catalan /a/). In such cases, a composite L1–L2 phonetic category may develop and lead to L1-based foreign-accented productions.

Similarly, the Perceptual Assimilation Model [28] extended to the L2 (PAM-L2) posits that success at detecting L2 phonological contrasts is dependent on how L2 phonemes are initially assimilated to the L1 phonemic inventory. In PAM-L2, when two L2 sounds are equally assimilated to a single L1 sound category (single-category assimilation), learners' ability to acquire the L2 sound contrast will be impaired as discriminating them will be very difficult (e.g., English /r/-/l/ for Japanese learners), whereas when two L2 sounds are unequally mapped onto the same L1 sound category so that one is judged as a better exemplar of the L1 category than the other (i.e., category-goodness assimilation), learners' ability to discriminate the L2 contrast will be moderate, making the acquisition of the contrast easier. Whereas SLM-r and PAM-L2 predictions for L2 sound acquisition have been applied to explaining phonological acquisition for monolingual immigrants living in an environment where L2 is spoken predominantly, neither model was designed with a foreign language (FL) instructional setting in mind, where L2 input is scarce and accented [30,32]. For example, [4] predicted that single-category and category-goodness assimilations were less likely to be acquired in the classroom than via immersion, especially if phonetic differences between the L2 phonemes were lacking in L2-accented spoken input or if L2 learners produced them on the basis of their written form. In fact, oral interactions in the FL classroom are likely to be L1-accented because teachers and students are likely to speak L2 with a shared L1-accent. In addition, unlike in naturalistic language learning, in FL instruction, vocabulary and grammar are primarily taught through written input, and little time is dedicated to communicative pronunciation teaching [4]. Still, models like SLM-r and PAM-L2 can inform L2 pronunciation instruction by explaining why certain L2 contrasts are easier to learn than others, which may be helpful in setting priorities when selecting pronunciation instruction targets and designing pedagogic tasks to learn them.

For example, Ref. [4] proposes that learners should be provided with plenty of opportunities for tuning in to the phonetic differences that distinguish L2 phonological contrasts prior to the acquisition of a large vocabulary so as to avoid homophony of minimal-pair words during L2 word recognition processes. An interactive map task where perceiving and producing a difficult L2 vowel contrast is essential for task resolution, such as the one described in the empirical study we report on in Section 4, can serve this purpose by raising learners' awareness of such phonetic differences and by providing practice in the distinction of the target L2 sounds. In addition, the introduction of written forms should be initially delayed, or orthography should provide one-to-one correspondence between phonemes and graphemes. Last but not least, pronunciation instruction should aim at orienting learners' attention to phonological contrasts in communicative settings so that L2 phonological learning can extend to spontaneous conversations beyond the L2 classroom.

3. Pronunciation Training Techniques and Teaching Methods in FL Learning Contexts

3.1. Phonetic Training

High-variability phonetic training (HVPT) is typically a perceptual training paradigm where learners are exposed to L2 sounds produced in a variety of phonetic environments by multiple speakers and need to identify or discriminate the target sounds after receiving individualized feedback [25]. HVPT has been investigated in relation to different segmental and suprasegmental features of L2 speech (e.g., vowels, consonants, syllable structure, tone) and has been found effective in developing L2 speech categories [24,25], leading to gains in L2 speech perception [33], lexical encoding [34] and production [35]. Importantly, phonetic and phonological learning from HVPT has been shown to generalize to novel talkers [36], untrained testing stimuli [37], new phonetic contexts [25], untrained sounds [33], and across perception and production modalities [38]. Additionally, learning gains tend to be retained sometime after the training has ended [18,33,38]. Training outcomes may differ in size as a function of presentation mode and stimulus type. For example, gains in production are greater if learners receive visual articulatory feedback than auditory-only feedback [35], if trained in adverse rather than silent conditions [39], or if trained with non-lexical rather than lexical stimuli [40,41].

Apart from HVPT, other explicit training paradigms, such as phonological specificity training, a paradigm that trains learners on minimal pairs to enhance the distinctiveness of phonolexical representations [42] and training in auditory processing skills [43], have also been shown to impact L2 phonological learning positively. Last, incidental multimodal phonetic training (e.g., playing with a mobile game application that exposes gamers to target sound contrasts when aliens are killed) has been found to be helpful for learners to automate the knowledge of L2 sounds that may have been learned explicitly [17]. Despite the multiple benefits of phonetic training for L2 speech development, one of the potential drawbacks of phonetic training is its limited ecological validity [6]. In other words, HVPT practice, unlike communicative types of practice (e.g., giving directions on a map where identifying street names depends on learners' accurate perception and production of a sound contrast), is often pedagogically decontextualized and can be disassociated with language learning and use in real-life contexts, and, hence, FL teachers may be hesitant to implement methods that have not been proven in real classrooms [44]. Still, empirically validated HVPT computer-assisted systems such as the Golden Speaker [45] or the English

Accent Coach [46] may provide accessible individualized pronunciation learning inside or outside the classroom.

3.2. Explicit Instruction

Explicit pronunciation instruction entails providing L2 learners with metalinguistic information about the voicing, place, and manner of articulation of L2 speech sounds and the acoustic as well as prosodic (stress, rhythm, and intonation) characteristics of L2 speech. Several meta-analyses [19,21,22] and individual studies [47,48] have shown the effectiveness of explicit instruction at improving L2 pronunciation and at making L2 speech more intelligible, comprehensible, and less strongly accented in classroom contexts. For example, Ref. [49] examined the impact of short-term explicit pronunciation teaching on both suprasegmental (stress, rhythm, reductions, and linking) and segmental features (/i/, /ɪ/, /æ/, /ɛ/) of L2 learners' speech. The intervention involved three 25 min sessions of explicit pronunciation instruction for three weeks. Compared to a non-explicit pronunciation instruction intervention, explicit phonetic instruction led to improvements in comprehensibility for learners trained on suprasegmental features but not for learners trained on segmentals. Other studies [50] have found little improvement resulting from explicit pronunciation instruction (around 5%) and no improvement for accentedness and comprehensibility.

While explicit pronunciation instruction helps learners notice L1–L2 phonological differences [6], it often involves a decontextualized focus on the accuracy of specific phonological forms, relying mainly on controlled practices. In fact, there is ongoing debate about whether gains obtained from explicit pronunciation instruction can be effectively maintained when using the L2 in real-life conversations. Given the prevailing emphasis on grammar-focused lessons and teachers' limited understanding of which aspects of pronunciation should be taught, researchers have a responsibility to inform teachers and teacher trainers about the key aspects of L2 pronunciation that should be given priority [51,52] and how to integrate them into content-based lessons [3]. One promising approach is engaging learners in interactive tasks that enhance learners' awareness of the communicative impact of pronunciation (e.g., an interaction map task; see Section 4). This way, learners can naturally focus on phonetic form while conveying meaning during communication.

3.3. Form-Focused Communicative Instruction

L2 pronunciation research [53–55] has demonstrated the superiority of explicit instruction combined with communicative form-focused instruction (FFI) over explicit instruction alone. FFI entails drawing learners' attention to form in communicative contexts, that is, practicing L2 pronunciation while being engaged in contextualized meaning-oriented communicative activities [20]. Recently, Ref. [56] compared explicit-only pronunciation instruction, consisting of listening and controlled practice of L2 suprasegmental features, to explicit instruction + FFI, which combined explicit teaching of pronunciation features repeatedly with communicative instruction following the Communicative Framework [57] and Automatization in Communicative Contexts of Essential Speech Segments framework (ACCESS) [58]. The results of this study showed that the "explicit-only" learners only improved L2 comprehensibility in controlled tasks, whereas the "explicit + FFI" learners improved both in controlled and, especially, in spontaneous tasks. These findings align with [20]'s synthesis of 15 quasi-experimental studies where FFI was found to contribute to the development of L2 speech in controlled and spontaneous speaking tasks, whereas the benefits of explicit-only instruction were only observable in controlled speech.

Given that the key to L2 phonological learning in the FL classroom is the automaticity of L2 phonological and phonetic processing and generalization from in-class to out-of-class language use, Refs. [3,53] suggest that learners can establish form-meaning mappings and develop L2 accuracy and fluency [59] by using activities that are intently repetitive yet have communicative value and by integrating a focus-on-phonetic-form into meaning-oriented tasks. Implementing a dual focus on form and meaning following the Communicative

Framework [57], ACCESS [58,59] or Strategy-based [60] frameworks has the potential to allow learners to notice and pay attention to L2 pronunciation features and to develop awareness of their own pronunciation problems (e.g., providing corrective feedback). The well-attested positive impact of explicit instruction in L2 phonology [22] can be maximized if it is extended to communicative language use contexts by gradually allowing learners to automate the procedural phonological knowledge they have acquired through form-focused activities when using L2 in contexts where they are primarily attending to meaning. This may in turn facilitate the spreading of L2-specific phonetic features (e.g., the aspiration of /p, t, k/ or discrimination of segmental contrasts /iː-ɪ/, /æ-ʌ/ in L2-English) to the entire lexicon, enhancing phonological acquisition and effectively improving L2 pronunciation while speaking the L2.

While most research has explored the benefits of integrating pronunciation in a communicative task after receiving explicit pronunciation instruction [10,61], practicing L2 pronunciation incidentally during communicative interaction following a TBLT approach [12–14,62] has been considerably less investigated.

3.4. New Avenues in L2 Pronunciation Training and Teaching

Current pedagogical practices in L2 pronunciation teaching and learning do not fully reflect the recent shift in the pedagogic target of L2 pronunciation learning from native-like speech to comprehensible speech. One way to promote the adjustment of pronunciation teaching to this paradigm shift is to make the outcomes of current research on the effectiveness of speech awareness-raising tasks to develop global dimensions of L2 pronunciation proficiency (intelligibility, comprehensibility, accentedness, and fluency) available to the pronunciation instruction community. Empirical research investigating the effectiveness of training tasks to develop L2 speech intelligibility, comprehensibility, accentedness, and fluency globally is scarce and varied in methods, in the level of learners' proficiency, and in its outcomes.

This section outlines a number of pronunciation training techniques whose effectiveness in raising awareness and developing L2 pronunciation has been experimentally proven: accent imitation, multimodal pronunciation training through captioned video, embodied pronunciation training, comprehensibility and accentedness self-assessment, and TBPT.

3.4.1. Accent Imitation

Research on foreign accent imitation training and its benefits for L2 pronunciation development is currently scarce, but findings so far [63,64] support the notion that training learners in imitating an L2 accent (e.g., an English accent) on their native language (L1) is helpful in developing awareness of L1–L2 cross-language phonetic differences and in enhancing the automatization of L2-specific articulatory gestures, leading to improvement of pronunciation accuracy at the segmental level, at least for low-proficiency learners [63]. In foreign accent imitation tasks, learners are asked to speak their L1 (or to produce word- or sentence-long utterances in their L1) with an L2 accent so that measures of phonetic features obtained from the imitated L2 accent (e.g., voice onset time—VOT) may be interpreted as a measure of implicit awareness (or implicit knowledge) of the phonetic properties of the L2 being imitated [65,66]. Learners have been found to imitate an L2 phonetic feature in their L1, such as VOT duration in voiceless oral stops, to the extent that they can produce them accurately in their L2 [67–69]. This technique has been applied to assess the production accuracy of L2-specific segmental phonetic features (mainly VOT) in a delayed accent-imitation paradigm. To the best of our knowledge, Ref. [63] is the only study that has used this technique to train L2-specific phonetic features in low-proficiency young learners. Given appropriate use of imitation training materials such as extended texts and spoken dialogues, accent imitation can be effective in training advanced adult learners' L2 phonetic features, including segmental and suprasegmental properties that may impact global dimensions of L2 speech.

3.4.2. Multimodal Pronunciation Training through Captioned Video

L2-captioned video (intralingual subtitles) can be effectively used to train L2 learners' simultaneous processing of L2 auditory input (speech), orthographic input (dynamic onscreen text), and visual input (onscreen dynamic images). Research has proven the benefits of this type of enriched input for listening comprehension, the incidental acquisition of vocabulary, and grammar. The presence of written word forms as learners process the spoken input has been shown to enhance auditory word recognition and speech segmentation skills [70,71] and can therefore offer interesting pedagogical possibilities for pronunciation instruction if learners' attention is guided to phonetic form while watching. For example, Ref. [72] enhanced learners' attention to phonetic form through pronunciation-related questions popping up occasionally on the screen while learners watched captioned video and found the treatment to improve L2 learners' speech segmentation and speech processing skills. Ref. [73] used audio-synchronized textual input enhancement in captions to promote learners' visual processing of usually mispronounced orthographic word forms immediately before they could be heard in the soundtrack and found benefits in learners' ability to recognize mispronounced forms, supporting the updating of non-target-like phonological representations of words.

3.4.3. Embodied Pronunciation Training

Embodied pronunciation training is based on the notion of multimodal enrichment, which holds that exposure to complementary information across multiple sensory modalities during learning activities can enhance learning benefits [74]. This is a type of multimodal phonetic training that takes advantage of the mutual effects of auditory perception and visual actions on one another to enhance the acquisition of segmental and suprasegmental features of speech [75]. For example, Ref. [76] showed that a group of L2-French learners assigned to an embodied pronunciation training condition (visuospatial hand gestures depicting speech rhythm and intonation during the oral repetition of CV syllables) improved their accentedness and suprasegmental features of L2 French in a dialogue read-aloud task significantly more than a comparable group assigned to a speech-only training condition. Similarly, recent research has also shown that phonetic training using visuospatial hand gestures, such as a fist-to-open-hand burst gesture to visually represent the auditory and articulatory features of Mandarin Chinese aspirated plosives [77], enhances phonetic learning, leading to more accurate production of aspirated stops.

3.4.4. Comprehensibility and Accentedness Self-Assessment

Self-assessment has mainly focused on identifying differences between L2 learners' assessments of their speech and assessments by native speakers (or L2 teachers) focusing primarily on accentedness and comprehensibility [78], but L2-speech self-assessment and peer-assessment tasks can be useful in raising learners' awareness of pronunciation features that make their speech difficult to understand or strongly accented [79,80]. Although L2 learners' ability to assess their own speaking performance is related to their actual speaking performance, the better their speaking skills, the more likely they are to accurately self-assess their performance [81]. Accurate speaking self-assessments indicate a lack of awareness and limitations in noticing the pronunciation features that affect their speech intelligibility, comprehensibility and accentedness [80] and the speech of others, which could have a negative impact on L2 speech development [82].

Research focusing specifically on the benefits of L2 pronunciation self-assessment for L2 pronunciation development [83,84] has generally found it difficult for learners to focus on specific phonological features and shows a mismatch between learners' self-assessments and assessments by native listeners [80,85] so that learners judged to perform well by native listeners tend to underestimate their performance, whereas learners judged to perform poorly tend to overestimate their performance. Overestimation and underestimation in speech-self assessment are indications of learners' difficulties in identifying the underlying pronunciation features that influence their speech, which can be improved with increased learners' experience in self-assessment [86] and with training in speech-self-assessment [87]. Methods for improving self-assessment skills include discussing learners' own performance, familiarizing learners with rating criteria, using self-testing exercises, benchmarking, and peer assessment. For example, Ref. [88] implemented a treatment based on benchmarking (asking learners to discuss speech evaluation criteria) and peer assessment (evaluating the speaking performance of peers) and found learners increased the alignment between their ratings and those of native listeners.

3.4.5. Task-Based Pronunciation Teaching (TBPT)

An approach to pronunciation teaching that can be integrated into communicative language teaching is task-based pronunciation teaching, or TBPT [11–13]. TBPT makes use of the task-design principles of TBLT to enhance learners' attention to phonetic form in pedagogic tasks that involve meaning-based interaction. TBLT adopts meaning-based communicative tasks as central to defining language learning needs, goals, classroom activities, and assessment, but has not devoted research efforts to investigating pronunciation-focused communicative tasks until recently [89]. Based on [90]'s Framework for Task-Based Learning, a commonly used TBPT task-design implementation procedure [12–14] involves three stages, namely, pre-task, task cycle, and post-task. During the pre-task, the teacher assists L2 learners in learning and recollecting the linguistic resources they will need to perform the interactive task. Learners focus both on phonetic form and meaning through comprehension activities and plan their speech before engaging in the main task. During the meaning-based interactive task, learners apply the language they have encountered in the preceding phase to carry out their interactions. Pronunciation targets (e.g., difficult vowel contrasts) are used incidentally but can be made essential for task completion, thus forcing learners to use the target sounds appropriately to perform the task. Immediately after, learners are asked to plan and report back on how they completed the task. This phase involves a transition from using spontaneous language with an emphasis on fluency to employing planned language that prioritizes fluency, accuracy, and clarity in organizing their public discourse. The last phase consists of orienting learners' attention to relevant pronunciation aspects that naturally arise during the conversations they engage in throughout the task cycle. The objective of the metalinguistic and communicative post-tasks is to help learners consolidate what they have learned and generalize their L2 knowledge to new contexts of L2 use.

TBLT research has shown that, by manipulating communicative task design variables (e.g., repetition, modality, or complexity), it is possible to enhance learners' attention to linguistic form during communicative interaction, helping learners develop their lexical, grammatical, and pragmatic L2 performance. For example, according to the Cognition Hypothesis [91], making a communicative task cognitively more complex by increasing the number of elements and reasoning demands needed to complete the task leads to the production of more lexically and grammatically complex language, thus implicitly enhancing L2 development. Recent research [89] indicates that manipulating task-design variables is effective at enhancing learners' attention to phonetic form in meaning-based tasks, resulting in improvements in the perception and production of L2 segmental [12,13,15] as well as suprasegmental features of L2 speech [92], and speech comprehensibility [10]. As learners tend to engage in such interactive tasks collaboratively, task performance often leads to the occurrence of pronunciation-focused language-related episodes (P-LRE) [93] that serve to raise learners' awareness of pronunciation issues and indicate the extent to which the task design was effective at helping learners focus on pronunciation. In addition, target pronunciation features can be made task-essential [94], forcing learners to pay attention to and use phonological targets to complete the task while focusing on meaning, such as when giving directions on a collaborative map task performed in pairs where the street names contain contrastive sounds that must be distinguished perceptually and in production for the learners to give and understand directions successfully [15,62].

Having reviewed novel methodologies of L2 pronunciation instruction in lab and classroom-based settings, we now illustrate the benefits of one of them, communicative TBPT, through an empirical lab study. TBPT allows teachers to fully integrate pronunciation instruction into communicative language teaching. In this lab-based study, we assess the efficacy of an interactive map task that could effectively be used by teachers in a classroom context to improve the perception and production of difficult L2 sound contrasts.

4. A TBPT Empirical Study

Although the benefits of task complexity for linguistic development are well attested in the domain of grammar and lexis [89], such benefits have only begun to be explored for L2 pronunciation. The few studies available to date [10,12–15] suggest that task complexity may be effective in enhancing attention to phonetic form during communicative pronunciation-focused tasks. As a means of illustrating methodological issues related to integrating a focus on phonetic form within communicative tasks, we report on an empirical study that aimed at applying TBPT principles to the design of a computerized map task. This task, which aimed at enhancing accuracy in the perception and production of a difficult L2 vowel contrast for L1-Spanish learners of English, was designed as a pronunciation-focused, meaning-oriented interactive pedagogic task. The design and preliminary outcomes of this study demonstrate the potential of TBPT for L2 pronunciation learning. Alternative interactive TBPT tasks to the ones described here can be found in [95,96].

4.1. Materials and Methods

In the current study, Ref. [90]'s TBLT task-cycle design was followed in that the experimental pedagogic intervention included a pre-task intended to familiarize learners with the phonetic targets to be improved, as well as an interactive communicative task designed around [97]'s definition of task: the primary focus of the task was on meaning; there was some kind of communicative gap; learners had to rely on their linguistic and non-linguistic resources to solve the task; and language was the means to achieve a clearly defined outcome. The task was framed as a two-way, close, convergent, giving-directions task [98] because it was performed in pairs; there was only one possible solution that students had to agree on, and students took it in turns to give and follow directions on a map. Unlike [90]'s framework, for research purposes, the current implementation of the task design lacked a post-task activity usually included in pedagogic tasks to consolidate the knowledge acquired through task performance.

This study followed a pre-test > intervention > post-test design. The pre-test and post-test included an ABX discrimination test to assess the effect of the intervention on L2 learners' accuracy and speed of response in perceptually distinguishing /iː/ from /ɪ/ (e.g., *feet-fit*), as well as delayed nonword repetition (DNWR) and delayed sentence repetition (DSR) tests to assess the effect of the intervention on L2 learners' accuracy in distinguishing /iː/ from /ɪ/ in production.

The intervention consisted of three 30 min computerized tasks in two sessions. Each task consisted of a perception pre-task, a production pre-task, and the corresponding interactive map task. The perception pre-task consisted of an identification task where learners were presented with the nonwords to be used in the corresponding map task and were asked to identify the stressed vowel in them as either English /iː/ or /ɪ/. The production pre-task consisted of an immediate repetition task where learners were presented with the nonwords to be used in the corresponding map task and were asked to repeat them as accurately as they could. Pre-tasks were aimed at familiarizing learners with the auditory and orthographic forms of the street names (English nonwords) they would practice during that map task session (18 minimal-pair nonword pairs). The map tasks were sequenced in order of increasing task complexity (simple > + complex > ++ complex). Learners needed to be able to distinguish /iː/ from /ɪ/ in perception and production to be able to follow and give instructions on the map and thus complete the task successfully. Map tasks 1 and 2 were performed in session 1, and map task 3 in session 2.

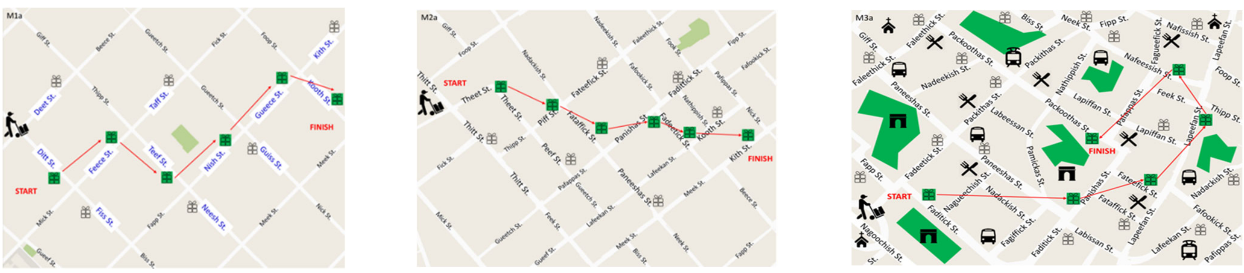
Given the exploratory nature of the current study and the relatively large inter-subject variability in L2 pronunciation proficiency of the participants, they were further randomly assigned to three task difficulty sub-groups. Task difficulty (low, medium, high) was operationalized in terms of how easy or difficult the target contrast was to perceive and produce in the street names based on whether the contrast was embedded in monosyllabic nonwords (easy: peef /pi:f/ vs. piff /pɪf/), trisyllabic nonwords (medium: lapeefan /lə'pi:fən/ vs. lapiffan /lə'pɪfən/), and a mixture of monosyllabic and trisyllabic nonwords (difficult), which might make the target contrast effectively harder for learners to attend to. Perceiving and producing the target vowels /i:/ and /ɪ/ accurately in a trisyllabic nonword was expected to pose greater difficulty and require greater attentional effort for learners than doing so in monosyllabic nonwords. In preliminary analyses of the participants' pre-test and post-test perception scores according to task stimuli difficulty [62], all sub-groups were found to improve sensitivity to the contrast after the intervention, but gains between testing times were smaller and did not reach significance for those participants assigned to the mixed stimuli difficulty condition, suggesting that stimuli variability limited the extent to which learners could benefit from the TBPT intervention and made the task more demanding for learners. As our aim here is to assess the effectiveness of the TBPT map task intervention, we focus on the overall results obtained by all the participants performing the map tasks as the experimental group.

4.1.1. Participants

The participants were 77 upper-intermediate to advanced L1-Spanish learners of English (*Mage* = 20.88, *SD* = 4.30). They had learned English through formal instruction in a school context since around the age of 6 (*SD* = 3.62), and at the time of testing, they were first-year students enrolled in a degree in English studies at a Spanish university. On a 9-point Likert scale (1 = very poor command of English, 9 = near-native command of English), participants rated themselves as fairly proficient (*M* = 6.67, *SD* = 1.17). They were randomly assigned to either an experimental group (*N* = 62) or a control group (*N* = 15) that did the pre-test and post-test but did not perform any of the familiarization pre-tasks or the intervention map tasks.

4.1.2. Map Task Design

Following the SSARC model of task sequencing [99], we designed three map tasks that differed in cognitive task complexity (simple, + complex, ++ complex) in terms of the number elements (streets and crossroads) on the map (see Figure 1) and asked learners to perform them in order of increasing complexity in two intervention sessions.



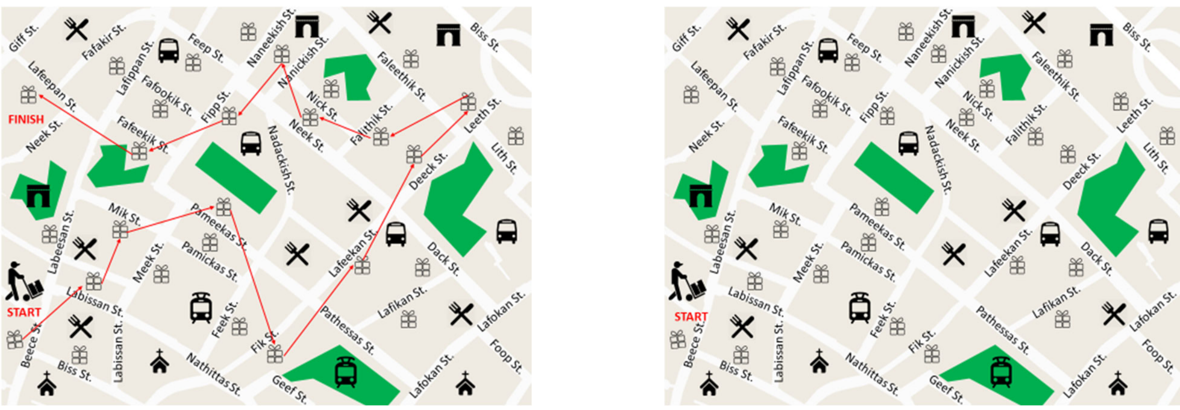
(**a**) (**b**) (**c**)

Figure 1. Simple (**a**), + Complex (**b**), and ++ Complex (**c**) map tasks.

The pronunciation target was the English vowel contrast /i:/-/ɪ/ (e.g., feet-fit), which Spanish learners of English find difficult to perceive and produce accurately because both English /i:/ and /ɪ/ are perceptually mapped onto the single Spanish front vowel category /i/ [100]. Although the L2 learners that participated in this study were relatively advanced, the target contrast was embedded in English nonwords and was deemed to pose difficulties in perception and production. The reason why we opted for using nonwords rather than words as stimuli for the street names on the map was twofold. First, using nonwords would avoid learners activating lexical representations for words they might be misrepresenting phonologically in their mental lexicon, which could possibly lead to a less effective intervention, as found in HVPT paradigms [40,41]. Secondly, using nonwords for which lexical representations cannot be activated could be helpful in helping learners concentrate on phonetic form (rather than meaning) when giving and following instructions on the map.

The learners' task was to give (Student A) and follow (Student B) directions using the non-word street names to pick 14 parcels located at streets off crossroads on the map. At each crossroad, a decision had to be made as to which street to take (e.g., lapeefan vs. lapiffan) where a parcel had to be picked. This was conducted in order to make the target vowel contrast task essential [94], forcing learners to focus on the qualitative differences between the contrastive vowels /iː/ and /ɪ/ in production and perception. The map task was collaborative in that each parcel to be picked involved producing and perceiving the contrast accurately by each learner dyad. Getting the wrong parcel would involve negotiating interactively to make the contrast clear to one another.

Student A (giving directions) and Student B (following directions) would be seated in front of each other, and each one of them would be in front of two monitors. Student A's monitor 1 would show a red line indicating the directions to be given according to the path Student B would need to follow to pick the 14 parcels (see Figure 2), whereas monitor 2 would show the same as students' B monitor 2 (the same map without the path). The red path was not visible for Student B, for whom only monitor 2 would be on. This design allowed Student A to monitor at all times what Student B's mouse pointer was doing. In order to pick up a parcel, Student B would need to double-click on it once its location on the map was reached according to the instructions. When clicking on it, the parcel would turn green if the correct street had been taken at the crossroads or red if the wrong street had been taken. The correct (and wrong) streets targeted an equal number of /iː/ and /ɪ/ nonwords. Taking the wrong street implied either Student A not pronouncing the street name correctly or Student B not perceiving the street name correctly, or both, which generated a number of P-LRE as students went back to the crossroads and tried to give and follow directions again to make up for the wrong choice of street. In order for learners to be able to monitor their production accuracy, all of the street names were clickable. When clicking on a street name, the learners would hear the pre-recorded nonword pronounced by either a male or a female native speaker of English. When getting to the end of the path and having picked the last parcel, students A and B changed roles in giving and following directions.



(**a**) (**b**)

Figure 2. Student A's monitor 1 (**a**) and monitor 2 (**b**) for the complex map task.

Having correctly picked 14 parcels implies having correctly produced and identified the contrastive vowels in 12 minimal-pair nonwords per map task (12 × 3 map tasks = 36 minimal-pair nonword pairs). The nonword street names were different in versions A and B of each map task, so students A and B had to produce and identify different sets of street names. Each map task also included two control vowel contrasts that posed no difficulty to learners (e.g., /i:/-/æ/, /ɪ/-/æ/, /i:/-/u:/, /ɪ/-/u:/). Students performed the map tasks in pairs while wearing Beyerdynamic DT 990 PRO open headphones, over which they could hear the names of the streets (whenever they clicked on them) and their partners' voices. All of the map tasks were audio-recorded (separately for each speaker) through Shure SM-58 voice microphones onto Marantz PMD660 solid-state digital recorders (44.1 kHz, 16-bit).

4.1.3. Testing and Analyses

In the ABX discrimination test, participants had to decide, as fast and accurately as possible, for every A-B-X trial, whether X contained the same vowel as A or the same vowel as B (e.g., lapeefan-lapiffan-lapiffan = B; lapeefan-lapiffan-lapeefan = A). ABX trials were presented with an inter-stimulus interval of 500 ms and 2000 ms after each response, a new trial was presented, or 2500 ms after the onset of the last item in the trial if no response was provided. In each ABX trial, A and B were spoken in the same voice (either male or female) and X in a different voice. This provided a measure of learners' ability to identify the target vowel across two different voices, an indication of having developed relatively robust, distinct sound categories for the target vowels /i:/ and /ɪ/. The test contained 96 ABX trials, corresponding to 16 test trials and 8 control trials in 4 orders (ABB, ABA, BAA, and BAB). The 96 trials contained an equal number of trained nonwords (nonwords included in the map tasks) and untrained nonwords (new nonwords produced by new voices), an equal number of monosyllabic and trisyllabic nonwords, and an equal number of nonwords produced by a male and a female voice. Accuracy rates and reaction times (RT) in correctly identifying X in ABX trials were used as a measure of learners' sensitivity to the target contrast in perception.

For production, we used DNWR and DSR tests. In the DNWR test, participants were asked to repeat, after a 1500 ms delay followed by a 250 ms beep signal, 64 test nonwords and 32 control nonwords. The 64 test nonwords corresponded to the nonwords in the 16 test trials included in the ABX test (16 /i:/ nonwords + 16 /ɪ/ nonwords = 32 nonwords) presented auditorily in fully randomized order and repeated twice, once after a male voice and once after a female voice. As in the ABX test, the test nonwords included a balanced design in terms of trained and untrained nonwords, monosyllabic and trisyllabic nonwords, and male and female voices. In the DSR test, participants were presented with a set of 8 mini-dialogues involving short prompt-response interactions between two different voices they listened to (e.g., Speaker A: Shall I put the heating on?; Speaker B: Yes, my feet are cold) and were then asked to repeat the response (Speaker A: Shall I put the heating on?; Participant: Yes, my feet are cold). The 8 responses targeted 8 common English words (4 with /i:/ and 4 with /ɪ/: sheep-ship, feet-fit, sit-seat, chips-cheap). Assessing learners' pronunciation of real words (apart from untrained nonwords) allowed us to determine whether the intervention map tasks could be effective at improving the phonological representations of lexical items.

The analysis of the production data involved computing average acoustic measures per vowel (/i:/ and /ɪ/) for each learner to assess changes in vowel quality (degree of vowel height and frontness) in the direction of the four native speakers' vowel productions used as testing stimuli in the ABX and the DNWR tests. Based on the $f0$, F1, and F2 formant frequency measures in Hertz which we extracted from the midpoint of the steady portion of each vowel, we computed a normalized Bark-converted (B) distance metric where B1-B0 represented a normalized height measure and B2-B1 a normalized frontness measure. We also assessed changes in the extent to which learners could make a qualitative distinction between the contrasting vowels in production by computing spectral distance scores (Euclidean distances) between /iː/ and /ɪ/ calculated on a two-dimensional height (B1-B0) by frontness (B2-B1) space. The larger the spectral distance score (SDS) between /iː/ and /ɪ/, the more distinct the production of the contrasting vowels was. We therefore expected SDSs to be larger at the post-test than at the pre-test.

4.2. Results

Having first checked that participants performed at much higher accuracy rates when discriminating control trials (/iː/-/æ/, /ɪ/-/æ/, /iː/-/uː/, /ɪ/-/uː/) than test trials (/iː/-/ɪ/) in the ABX test at pre-test (M = 0.87, SD = 0.332 vs. M = 0.68, SD = 0.468, respectively), we explored the effects of the map task intervention and generalization effects to untrained nonwords based on test nonwords only. The results show that accuracy scores between testing times were slightly larger for the experimental group than the control group and that participants seem to improve similarly between testing times on trained and untrained items (Figure 3). The RT data also showed slightly larger differences in speed between testing times for the experimental group than the control group (Figure 4). However, unexpectedly, the control group obtained considerable perception gains, which can only be explained by a task effect, that is, improvement associated with repeating the perception test twice. In addition, both the experimental and control groups seemed to perform on trained and untrained nonwords equally well, suggesting that improvements in perception could not be attributed to testing trials having been previously trained.

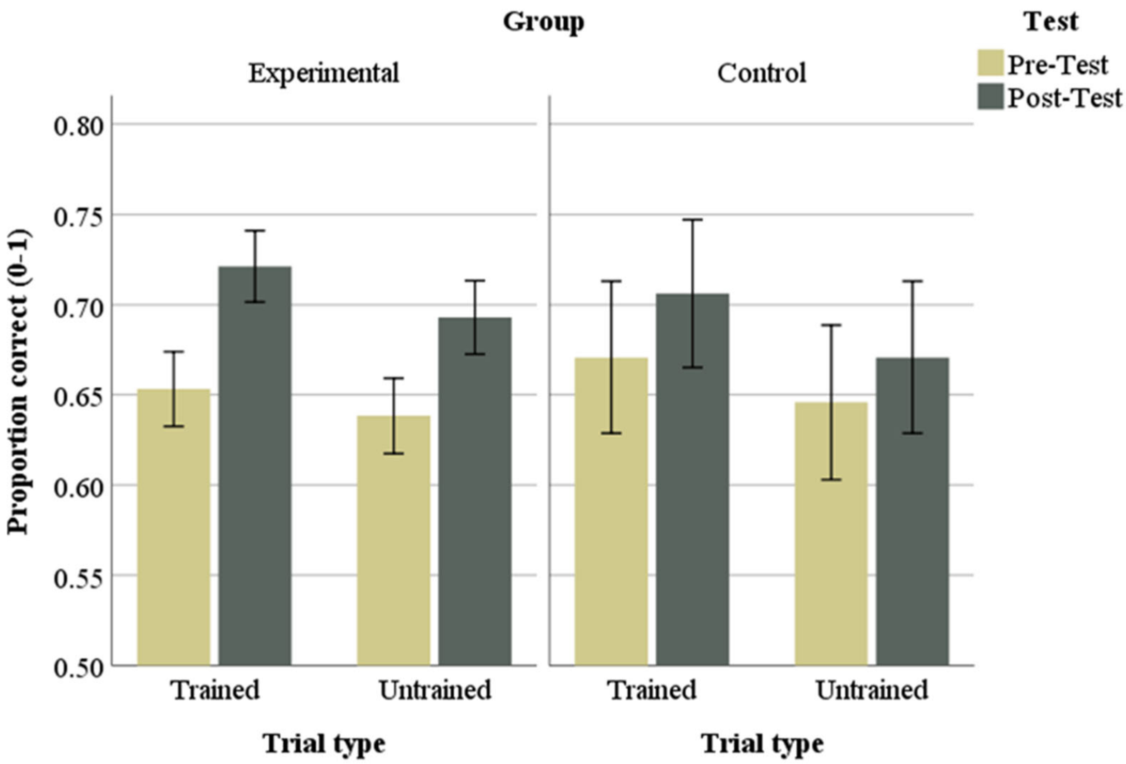


Figure 3. ABX accuracy by group, test and trial type (Error bars = 95% CI).

We first assessed whether learning gains could generalize to untrained nonwords by submitting participants' ABX responses to test items (correct or wrong) to a linear mixed-effects model with a binomial logistic regression (in SPSS 25) with test (pre-test, post-test), group (experimental, control), and trial type (trained, untrained) and their interactions as fixed effects, including a random intercept for subject. These analyses showed significant main effects of test ($F(1, 9912) = 19.75$, $p < 0.001$) and trial type ($F(1, 9912) = 4.98$, $p = 0.026$), but neither the main effect of group ($F(1, 9912) = 0.001$, $p = 0.977$) nor any of the interactions reached significance (see Appendix A for the model's parameter estimates). Tests of pairwise contrasts showed that while at pre-test the experimental group performed equally well on trained and untrained items ($t(9912) = 1.01$, $SE = 0.015$, $p = 0.316$), at post-test they discriminated the target vowels at significantly higher correct rates in trained than in untrained nonwords ($t(9912) = 1.98$, $SE = 0.014$, $p = 0.048$). This indicated, as expected, that training had a larger effect on trained than untrained nonwords, although both improved significantly between testing times. Interestingly, whereas the experimental group discriminated the target vowels significantly more accurately at post-test than at pre-test in both trained ($t(9912) = -5.01$, $SE = 0.015$, $p < 0.001$) and untrained ($t(9912) = -5.01$, $SE = 0.015$, $p < 0.001$) nonwords, the control group did not significantly improve on either trained ($t(9912) = -1.21$, $SE = 0.030$, $p = 0.229$) or untrained ($t(9912) = -0.026$, $SE = 0.031$, $p = 0.406$) nonwords. These findings indicate that for the experimental group (but not for the control group), discrimination accuracy improved for both trained and untrained nonwords, showing that training gains generalized to untrained nonwords. Consequently, in all subsequent analyses, accuracy and RT scores from test items (including trained and untrained nonwords) were used as dependent measures.

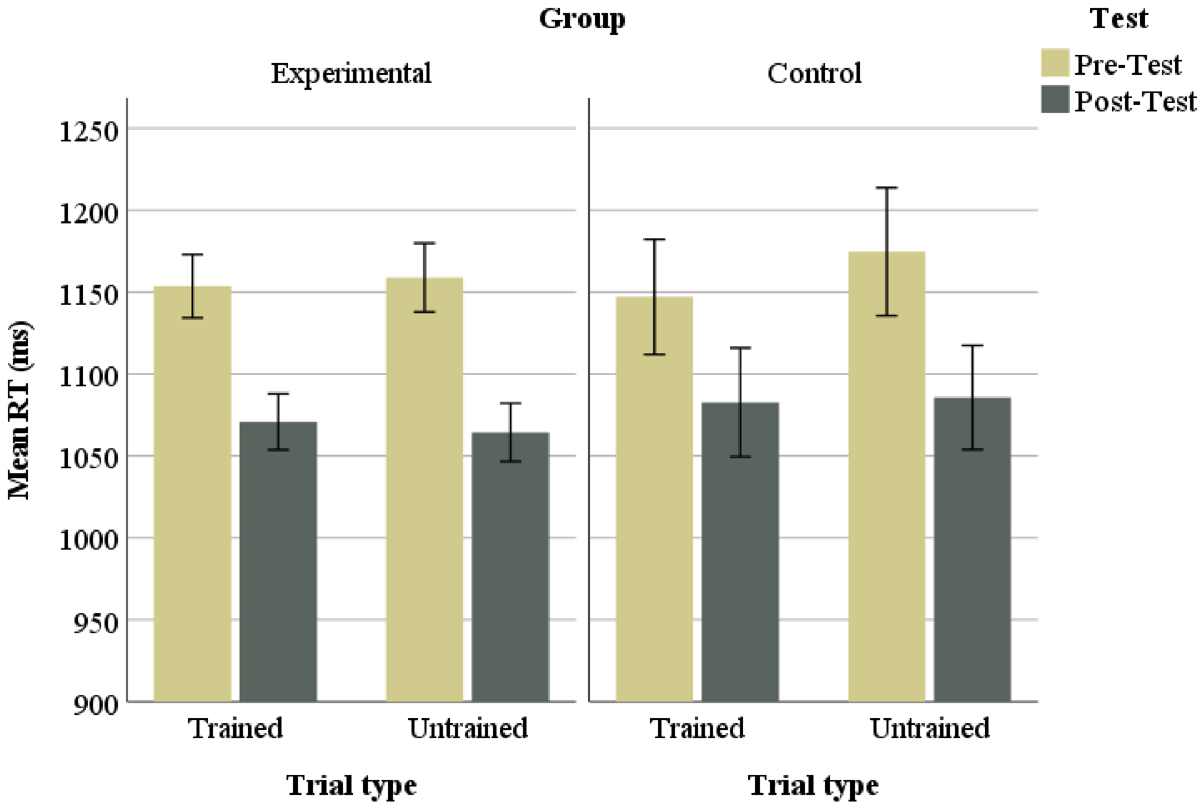


Figure 4. ABX response speed by group, test and trial type (Error bars = 95% CI).

In order to assess whether the intervention effects on perception reached significance for the experimental and control groups, we first submitted participants' ABX responses to test items (correct or wrong) to a linear mixed-effects model with a binomial logistic regression with test (pre-test, post-test) and group (experimental, control) and their interaction as fixed effects and random intercepts for subject and item. These analyses showed a significant main effect of test ($F(1, 9912) = 19.75$, $p < 0.001$), but neither the main effect of group ($F(1, 9912) = 0.001$, $p = 0.978$), nor the test $\times$ group interaction ($F(1, 9912) = 2.75$, $p = 0.097$) reached significance (see Appendix A for the model's parameter estimates). According to Bonferroni-adjusted pairwise contrasts, the main effect of the *test* was driven by the gains obtained by the experimental group, who gained a modest but significant 7.6% in accuracy ($t(9916) = -6.52$, $SE = 0.012$, $p < 0.001$), whereas the gains by the control group (3.5%) did not reach significance ($t(9916) = -1.54$, $SE = 0.023$, $p = 0.123$). We then submitted participants' ABX RT on test items, including only those for correct responses and excluding those beyond 2.5 standard deviations from each subject's mean (2.05%)

to a linear mixed-effects model with test (pre-test, post-test) and group (experimental, control) as fixed effects and random intercepts for subject and item. The outcome of these analyses was similar to those we obtained for accuracy: there was a significant main effect of test ($F(1, 6551) = 111.4$, $p < 0.001$), but neither the main effect of group ($F(1, 6551) = 0.152$, $p = 0.697$), nor the test × group interaction ($F(1, 6551) = 0.459$, $p = 0.498$) reached significance (see Appendix A). However, in this case, both the experimental ($t(6551) = 12.74$, $SE = 7.47$, $p < 0.001$) and control ($t(6551) = 5.50$, $SE = 15.22$, $p < 0.001$) groups significantly improved in RT at post-test, with the experimental group only improving 12 ms on average more than the control group (95.2 ms vs. 83.7 ms).

In terms of production accuracy in the DNWR test, improvement between testing times in vowel quality was small and mainly affected /ɪ/ (as English /i:/ is already acoustically very similar to Spanish /i/). Learners' /ɪ/ became lower in height and slightly more centralized, whereas /i:/ became slightly more target-like only in height. However, spectral distances appear to become larger at post-test, suggesting that learners had improved in making a distinction between /i:/ and /ɪ/ in production, although the spectral distance produced between vowels was much smaller than the one the native speakers produced (see Table 1). In order to assess the effectiveness of the map-task intervention in effecting improvement in learners' ability to distinguish /i:/ from /ɪ/ in production, we ran a Paired-samples T-test on the learners' pre-test and post-test spectral distance scores, which confirmed that spectral distances were of a significantly larger magnitude at post-test than at pre-test ($t(61) = -2.59$, $p = 0.012$).

Table 1. Normalized Bark distance metrics and spectral distances for natives ($N = 4$) and Learners ($N = 63$) in the DNWR and DSR tests.

Group	Test	/i:/ B1-B0 (Height)		/i:/ B2-B1 (Frontness)		/ɪ/ B1-B0 (Height)		/ɪ/ B2-B1 (Frontness)		/i/-/ɪ/ Euclidean Distance			
		M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**Min.**	**Max.**
Natives		1.96	0.24	10.87	0.55	2.98	0.42	8.65	0.85	2.45	0.79	1.85	3.52
Learners (DNWR)	Pre-test	1.76	0.53	10.50	0.84	2.48	0.44	9.32	0.64	1.58	0.60	0.59	3.37
	Post-test	1.71	0.51	10.65	0.76	2.53	0.45	9.26	0.67	1.75	0.66	0.69	3.72
Learners (DSR)	Pre-test	1.62	0.53	10.68	0.74	2.11	0.48	9.76	0.77	1.13	1.0	0.13	4.82
	Post-test	1.53	0.56	10.73	0.80	2.06	0.49	9.82	0.78	1.14	1.0	0.07	4.50

We also measured the vowel productions from the DSR test in the same way as those from the DNWR test and computed spectral distance scores between /i:/ and /ɪ/ to assess improvement in learners' ability to distinguish /i:/ from /ɪ/ in minimal-pair words they knew. Spectral distances were on average much smaller between /i:/ and /ɪ/ on these words than they were on the nonwords from the DNWR test (see Table 1), suggesting that learners made less of a distinction between /i:/ and /ɪ/ in real words embedded in sentences produced from memory than in the production of isolated nonwords. Improvement between pre-test and post-test spectral distances was very small (see Table 1) and non-significant ($t(61) = -0.117$, $p = 0.907$), suggesting that the positive effect of the map-task intervention on production did not generalize to untrained words (sheep-ship, feet-fit, sit-seat, chips-cheap) whose phonological representations were already well-established.

4.3. Discussion

Overall, the effects of the computerized map-task intervention, albeit small, suggest that a two-session TBPT communicative interaction helped learners improve their ability to distinguish a difficult L2 vowel contrast (/i:/-/ɪ/) in perception and production.

On the one hand, the perception results suggest that performing three oral map tasks of increasing cognitive complexity resulted in an improvement in the discrimination accuracy of the target vowels /i:/ and /ɪ/. These findings echo those observed in studies that carried out implicit high-variability perceptual training of segmentals [17] or those applying form-focused communicative interventions [61]. The fact that discrimination gains were lower in this study than in others obtained through explicit HVPT [33,101] might be due to the short two-session (three map tasks) intervention. Still, generalization to untrained items, in line with other studies [33,37,41] signals robust improvement in L2 vowel discrimination.

On the other hand, the production results show that learners produced larger distances (i.e., less overlap) between contrastive vowels after the TBPT intervention. In accordance with [102] and [15]'s findings, reflection on phonetic form during communicative tasks that make pronunciation targets task-essential and necessitate agreement on a single correct outcome [97] leads to more distinct realizations of L2 confusable sounds, hence, more target-like productions. Our findings for production are in accordance with production gains reported in other short communicative form-focused instruction interventions [55]. Nevertheless, learners were unable to generalize vowel distinctiveness gains to untrained word contexts, contrary to [12,13], which involved learners' performance on many more lexically-based tasks during a longer intervention. This could be attributed to the short two-session intervention as well as to the size of the small data set the DSR test generated (vowel productions from 8 /iː/-/ɪ/ minimal-pair words).

Finally, a number of important limitations need to be considered. First, the lack of L2 production data from the control group suggests the outcome of the analyses in terms of L2 production gains obtained by the experimental group resulting from the TBPT intervention cannot be ascertained and needs to be interpreted with caution. Second, a two-session intervention may not be long enough for learners to develop detectable gains in how distinctly they can produce L2 vowels. Last, the fact that the battery of perceptual and production tasks used at pre-test and post-test to gauge learning gains were very different from the map tasks used in the intervention may have made it difficult for us to observe L2 pronunciation learning gains that might otherwise have shown in an interactive testing task (a map task) more similar to the intervention tasks. Further data analyses on the frequency of P-LRE during learners' interaction and of the performance on the test map task would likely provide further insights into the effectiveness of TBPT for improving L2 pronunciation.

5. Conclusions

The current article has outlined and discussed current issues in L2 pronunciation instruction research and practice in ISLA arising from the need to overcome the limitations and challenges FL learning contexts pose to L2 learners' pronunciation development. These include, but are not limited to, learners' scarce L2 exposure and use, the difficulty in applying well-established findings within current L2 speech learning models (SLM-r, PAM-L2) to the FL classroom, the need to seek novel methods and techniques to train L2 pronunciation globally, and the difficulty of integrating a focus on phonetic form in meaning-based tasks in a primarily grammar-centered communicative approach to language teaching. We have highlighted key features of L2 speech models that can inform pronunciation training and instruction and synthesized commonly used as well as novel training and teaching pronunciation instruction methods. Finally, we have exemplified TBPT as a pronunciation instruction method by reporting on the results of an empirical study that used communicative map tasks to teach segmental pronunciation targets.

As a means to incorporate research findings from speech learning models into pronunciation instruction, we suggest that both the design of pronunciation-focused communicative tasks (e.g., TBPT) and the design of individualized pronunciation training techniques (e.g., accent imitation) consider two key aspects of L2 speech learning: namely, (a) that the difficulties in the acquisition of L2 segmental phonology (e.g., L2 sound contrasts) are based on specific L2-to-L1 perceptual sound mappings, and (b) that the target phonetic and phonological features to be acquired are likely to vary in how essential they are to L2 speech intelligibility. Therefore, HVPT paradigms should consider training learners not only on the identification and discrimination of L2 sound contrasts but also on the perception of the cross-language differences between the segmental phonologies of the L1 and the L2 that determine which L2 sound contrasts are difficult to acquire, as this would help learners overcome L1-based perception. Similarly, pronunciation-focused communicative tasks could be designed to make pronunciation targets (especially those having a larger impact on speech intelligibility) essential to task resolution to enhance attention to phonetic form during communicative interaction.

Most of the pronunciation training and teaching methods outlined in this article exemplify novel creative ways of overcoming the difficulties associated with learning pronunciation in ISLA, but their effectiveness needs further exploration and empirical support. Evaluating their impact on L2 pronunciation development requires establishing consistent measurement standards, which are currently too varied and inconsistent across studies. Research assessing the effectiveness of pronunciation training techniques for L2 pronunciation development makes use of either fine-grained acoustic analyses of target phonetic features (e.g., [12,13] as well as the HVPT and accent imitation studies we have reported on) or perceptual judgments of global dimensions of L2 speech such as intelligibility, comprehensibility, accentedness, and fluency (e.g., [54] and the embodied pronunciation training studies we have reported on). However, the multidimensional nature of L2 speech and the current trend to define the functional load of pronunciation targets in terms of speech intelligibility and comprehensibility [51,52] suggest that pronunciation assessment, especially when aiming at determining the effectiveness of pronunciation training and teaching methods, should be carried out through a combination of acoustic measurements and listeners' judgments of global dimensions of L2 pronunciation proficiency [21].

The TBPT empirical study we report on, together with a few other similar studies [10,12], illustrates a novel approach that incorporates the research findings of L2 speech learning models and TBLT into pronunciation instruction while effectively integrating pronunciation instruction into communicative classroom teaching in ISLA. Although further research is needed to confirm the pedagogic value of TBPT, current research findings already offer preliminary evidence of its effectiveness for L2 pronunciation learning. In addition to integrating pronunciation instruction into communicative language teaching, pronunciation development in instructed SLA needs to combine classroom pedagogical practice with individualized pronunciation training that can provide personalized feedback [48]. A pronunciation teaching approach that combines in-class TBPT tasks with out-of-class pronunciation training through tasks such as those described above may provide a very effective way to learn L2 pronunciation, especially if combined in a structured way, so that individualized pronunciation training tasks serve to raise awareness of challenging L2 phonological features that will be at a later stage practiced communicatively through TBPT tasks in the classroom. Assessing the effectiveness of such a combined approach to L2 pronunciation learning opens exciting research avenues in L2 pronunciation instruction.

Author Contributions: Conceptualization, J.C.M. and I.M.-P.; methodology, J.C.M.; formal analysis, J.C.M. and I.M.-P.; investigation, J.C.M.; data curation, J.C.M. and I.M.-P.; writing—original draft preparation, J.C.M. and I.M.-P.; writing—review and editing, J.C.M. and I.M.-P.; project administration, J.C.M.; funding acquisition, J.C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish Ministry of Science & Innovation, grant number PID2019-107814GB-I00, and the Catalan Department of Research and Universities (AGAUR), grant number 2021SGR00303.

Institutional Review Board Statement: The study protocol adhered to the good practices of data collection, data anonymization, data processing, and data storage of the Institutional Review Board of the University of Barcelona (IRB0003099) and in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to part of the data being still under analysis.

Acknowledgments: We would like to thank Mayya Levkina for giving us permission to use for the current study part of the data collected and analyzed in Mora, J. C. & Levkina, M. (2018). Training vowel perception through map tasks: The role of linguistic and cognitive complexity. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference* (pp. 151–162). Ames, IA: Iowa State University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of this study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Parameter estimates of linear mixed-effects models on ABX discrimination scores.

		β	SE	t	p	95% CI	
						Lower	Upper
Model 1 [1]	Intercept	0.74	0.15	5.03	<0.001	0.45	1.03
	Group	0.10	0.16	0.64	0.526	−0.22	0.43
	Test	−0.12	0.14	−0.83	0.406	−0.39	0.16
	Item type	0.17	0.14	1.20	0.228	−0.11	0.45
	Group × Test	−0.16	0.15	−1.04	0.299	−0.46	0.14
	Group × Trial type	−0.03	0.16	−0.20	0.845	−0.34	0.28
	Test × Trial type	−0.06	0.20	−0.28	0.779	−0.45	0.33
	Group × Test × Trial type	−0.02	0.22	−0.08	0.938	−0.45	0.42
Model 2 [2]	Intercept	0.93	0.18	5.04	<0.001	0.57	1.29
	Group	0.10	0.17	0.62	0.535	−0.22	0.43
	Test	−0.17	0.11	−1.55	0.121	−0.38	0.04
	Group × Test	−0.20	0.12	−1.66	0.097	−0.43	0.04
Model 3 [3]	Intercept	111.48	52.28	21.24	<0.001	1007.99	1212.98
	Group	−27.59	56.70	−0.49	0.627	−138.73	83.56
	Test	83.73	15.22	5.50	<0.001	53.90	113.56
	Group × Test	11.49	16.95	0.68	0.498	−21.74	44.72

[1] ABX accuracy (generalization effects), [2] ABX accuracy (test effects), [3] ABX RT (test effects).

References

1. Sardegna, V.G.; Lee, J.; Kusey, C. Self-efficacy, attitudes, and choice of strategies for English pronunciation learning. *Lang. Learn.* **2018**, *68*, 83–114. [CrossRef]
2. Baker, A. Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *Tesol Q.* **2014**, *48*, 136–163. [CrossRef]
3. Darcy, I. Powerful and Effective Pronunciation Instruction: How Can We Achieve It? *Catesol J.* **2018**, *30*, 13–45.
4. Tyler, M.D. PAM-L2 and phonological category acquisition in the foreign language classroom. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*; Nyvad, A.M., Hejná, M., Højen, A., Jespersen, A.B., Hjortshøj Sørensen, M., Eds.; Department of English, School of Communication & Culture, Aarhus University: Aarhus, Denmark, 2019; pp. 607–630.
5. Huensch, A. Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Lang. Teach. Res.* **2019**, *23*, 745–764. [CrossRef]
6. Derwing, T.M.; Munro, M.J. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*; John Benjamins: Amsterdam, The Netherlands, 2015.
7. Derwing, T.M.; Munro, M.J.; Thomson, R.I. (Eds.) *The Routledge Handbook of Second Language Acquisition and Speaking*; Routledge: New York, NY, USA, 2022.
8. Kang, O.; Thomson, R.; Murphy, J. *The Routledge Handbook of Contemporary English Pronunciation*; Routledge: New York, NY, USA, 2018.
9. Wayland, R. *Second Language Speech Learning*; Cambridge University Press: New York, NY, USA, 2021.
10. Gordon, J. Pronunciation and Task-Based Instruction: Effects of a Classroom Intervention. *Relc. J.* **2021**, *52*, 94–109. [CrossRef]
11. Mora, J.C.; Levkina, M. Task-based pronunciation teaching and research: Key issues and future directions. *SSLA* **2017**, *39*, 381–399. [CrossRef]
12. Mora-Plaza, I. Task-Based Pronunciation Teaching and Learning of L2 Vowels in EFL Learners: Task Complexity Effects. Ph.D. Thesis, Universitat de Barcelona, Barcelona, Spain, June 2023.

13. Mora-Plaza, I. Task-based pronunciation teaching helps improve L2 vowel production: Generalization effects. In Proceedings of the 7th English Pronunciation: Issues & Practices Conference, Grenoble, France, 18–20 May 2022; Henderson, A., Kirkova-Naskova, A., Eds.; Université Grenoble-Alpes: Grenoble, France, 2023; pp. 172–187.
14. Mora-Plaza, I.; Mora, J.C.; Gilabert, R. Focus on phonetic form and cognitive complexity in the acquisition of an L2 vowel contrast. In Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference, Salt Lake City, UT, USA, 1–2 September 2017; Levis, J.M., Ed.; Iowa State University: Ames, IA, USA, 2018; pp. 174–184.
15. Solon, M.; Long, A.Y.; Gurzynski-Weiss, L. Task Complexity, Language-Related Episodes, and Production of L2 Spanish Vowels. *SSLA* **2017**, *39*, 347–380. [CrossRef]
16. Fouz-González, J. Using apps for pronunciation training: An empirical evaluation of the English File Pronunciation app. *Lang. Learn. Technol.* **2020**, *24*, 62–85.
17. Saito, K.; Hanzawa, K.; Petrova, K.; Kachlicka, M.; Suzukida, Y.; Tierney, A. Incidental and Multimodal High Variability Phonetic Training: Potential, Limits, and Future Directions. *Lang. Learn.* **2022**, *72*, 1049–1091. [CrossRef]
18. Wang, X.; Munro, M.J. Computer-based training for learning English vowel contrasts. *System* **2004**, *32*, 539–552. [CrossRef]
19. Lee, J.; Jang, J.; Plonsky, L. The effectiveness of second language pronunciation instruction: A meta-analysis. *Appl. Linguist.* **2015**, *36*, 345–366. [CrossRef]
20. Saito, K. Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies. *Tesol Q.* **2012**, *46*, 842–854. [CrossRef]
21. Saito, K.; Plonsky, L. Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Lang. Learn.* **2019**, *69*, 652–708. [CrossRef]
22. Thomson, R.I.; Derwing, T.M. The effectiveness of L2 pronunciation instruction: A narrative review. *Appl. Linguist.* **2015**, *36*, 326–344. [CrossRef]
23. Barriuso, T.A.; Hayes-Harb, R. High Variability Phonetic Training as a Bridge from Research to Practice. *Catesol J.* **2018**, *30*, 177–194.
24. Sakai, M.; Moorman, C. Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Appl. Psycholinguist.* **2018**, *39*, 187–224. [CrossRef]
25. Thomson, R.I. High variability [pronunciation] training (HVPT): A proven technique about which every language teacher and learner ought to know. *JSLP* **2018**, *4*, 208–231. [CrossRef]
26. Levis, J.M. Changing contexts and shifting paradigms in pronunciation teaching. *Tesol Q.* **2005**, *39*, 369–377. [CrossRef]
27. Levis, J.M. Revisiting the intelligibility and nativeness principles. *JSLP* **2020**, *6*, 310–328. [CrossRef]
28. Best, C.; Tyler, M. Non-native and second language speech perception. In *Language Experience in Second Language Speech Learning*; Bohn, O.-S., Munro, M.J., Eds.; John Benjamins: Amsterdam, The Netherlands, 2007; pp. 15–34.
29. Flege, J.E.; Bohn, O.-S. The revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*; Wayland, R., Ed.; Cambridge University Press: Cambridge, UK, 2021; pp. 3–83.
30. Muñoz, C. Contrasting effects of starting age and input on the oral performance of foreign language learners. *Appl. Linguist.* **2014**, *35*, 463–482. [CrossRef]
31. Tyler, M.D.; Best, C.T.; Faber, A.; Levitt, A.G. Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica* **2014**, *71*, 4–21. [CrossRef]
32. Piske, T. Implications of James E. Flege's research for the foreign language classroom. In *Language Experience in Second Language Speech Learning*; Bohn, O.-S., Munro, M.J., Eds.; John Benjamins: Amsterdam, The Netherlands, 2007; pp. 301–314.
33. Carlet, A.; Cebrian, J. Assessing the Effect of Perceptual Training on L2 Vowel Identification, Generalization and Long-term Effects. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*; Nyvad, A.M., Hejná, M., Højen, A., Jespersen, A.B., Hjortshøj Sørensen, M., Eds.; Department of English, School of Communication & Culture, Aarhus University: Aarhus, Denmark, 2019; pp. 91–119.
34. Melnik-Leroy, G.A.; Peperkamp, S. High-variability phonetic training enhances second language lexical processing: Evidence from online training of French learners of English. *Biling. Lang. Cogn.* **2021**, *24*, 497–506. [CrossRef]
35. Kartushina, N.; Hervais-Adelman, A.; Frauenfelder, U.H.; Golestani, N. The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *JASA* **2015**, *138*, 817–832. [CrossRef] [PubMed]
36. Lee, D.Y.; Baese-Berk, M.M. Non-native English listeners' adaptation to native English speakers. *JASA Express Lett.* **2021**, *1*, 105201. [CrossRef] [PubMed]
37. Mora, J.C.; Mora-Plaza, I. Contributions of cognitive attention control to L2 speech learning. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*; Nyvad, A.M., Hejná, M., Højen, A., Jespersen, A.B., Hjortshøj Sørensen, M., Eds.; Department of English, School of Communication & Culture, Aarhus University: Aarhus, Denmark, 2019; pp. 477–499.
38. Rato, A.; Rauber, A. The effects of perceptual training on the production of English vowel contrasts by Portuguese learners. In Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, UK, 10–14 August 2015.
39. Cooke, M.; García Lecumberri, M.L. Effects of exposure to noise during perceptual training of non-native language sounds. *JASA* **2018**, *143*, 2602–2610. [CrossRef] [PubMed]
40. Mora, J.C.; Ortega, M.; Mora-Plaza, I.; Aliaga-García, C. Improving the pronunciation of L2 vowels under different training conditions: The use of non-lexical materials and masking noise. *Phonetica.* **2022**, *79*, 1–43. [CrossRef] [PubMed]

41. Ortega, M.; Mora-Plaza, I.; Mora, J.C. Differential effects of lexical and non-lexical high-variability phonetic training on the production of L2 vowels. In *English Pronunciation Instruction: Research-Based Insights*; Kirkova-Naskova, A., Henderson, A., Fouz-González, J., Eds.; John Benjamins: Amsterdam, The Netherlands, 2021; pp. 327–356.
42. Llompart, M.; Reinisch, E. Lexical representations can rapidly be updated in the early stages of second-language word learning. *J. Phon.* **2021**, *88*, 101080. [CrossRef]
43. Saito, K.; Petrova, K.; Suzukida, Y.; Kachlicka, M.; Tierney, A. Training auditory processing promotes second language speech acquisition. *JEPHPP* **2022**, *48*, 1410–1426. [CrossRef] [PubMed]
44. Martin, I.A.; Inceoglu, S. The laboratory, the classroom, and online: What works in each context. In *Second Language Pronunciation: Bridging the Gap between Research and Teaching*; Levis, J.M., Derwing, T.M., Sonsaat-Hegelheimer, S., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2022; pp. 254–272.
45. Ding, S.; Liberatore, C.; Sonsaat, S.; Lučić, I.; Silpachai, A.; Zhao, G.; Chukharev-Hudilainen, E.; Levis, J.; Gutierrez-Osuna, R. Golden speaker builder–An interactive tool for pronunciation training. *Speech Commun.* **2019**, *115*, 51–66. [CrossRef]
46. Thomson, R.I. Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Lang. Learn.* **2012**, *62*, 1231–1258. [CrossRef]
47. Gordon, J.; Darcy, I. Teaching segmentals and suprasegmentals: Effects of explicit pronunciation instruction on comprehensibility, fluency, and accentedness. *JSLP* **2022**, *8*, 168–195. [CrossRef]
48. Saito, K.; Lyster, R. Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Lang. Learn.* **2012**, *62*, 595–633. [CrossRef]
49. Gordon, J.; Darcy, I. The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *JSLP* **2016**, *2*, 56–92. [CrossRef]
50. Kissling, E.M. Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *Mod. Lang.* **2013**, *97*, 720–744. [CrossRef]
51. Munro, M.J.; Derwing, T.M. The functional load principle in ESL pronunciation instruction: An exploratory study. *System* **2006**, *34*, 520–531. [CrossRef]
52. Suzukida, Y.; Saito, K. Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. *Lang. Teach. Res.* **2021**, *25*, 431–450. [CrossRef]
53. Darcy, I.; Hancock, Z.; Lee, J.; Rocca, B. Improving pronunciation in spontaneous speech? A comparison of two instructional methods. In Proceedings of the L2 Pronunciation Research Workshop: Bridging the Gap between Research and Practice, Barcelona, Spain, 6–7 June 2019.
54. Darcy, I.; Rocca, B.; Hancock, Z. A window into the classroom: How teachers integrate pronunciation instruction. *RELC J.* **2021**, *52*, 110–127. [CrossRef]
55. Saito, K. Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Appl. Psycholinguist.* **2015**, *36*, 377–409. [CrossRef]
56. Darcy, I.; Rocca, B. Comprehensibility improvements in integrated pronunciation instruction: A comparison of instructional methods and task effects. *J. Second. Lang. Pronunciation* **2023**, *8*, 328–362. [CrossRef]
57. Celce-Murcia, M.; Brinton, D.; Goodwin, J. *Teaching Pronunciation: A Reference Book for Teachers*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2010.
58. Gatbonton, E.; Segalowitz, N. Rethinking communicative language teaching: A focus on access to fluency. *Can. Mod. Lang. Rev.* **2005**, *61*, 325–353. [CrossRef]
59. Trofimovich, P.; Gatbonton, E. Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *Mod. Lang. J.* **2006**, *90*, 519–535. [CrossRef]
60. Sardegna, V.G. Evidence in favor of a strategy-based model for English pronunciation instruction. *Lang. Teach.* **2022**, *55*, 363–378. [CrossRef]
61. Ruan, Y.; Saito, K. Communicative focus on phonetic form revisited: Less precise auditory processing limits instructed L2 speech learning. *System* **2023**, *114*, 103020. [CrossRef]
62. Mora, J.C.; Levkina, M. Training vowel perception through map tasks: The role of linguistic and cognitive complexity. In Proceedings of the 9th Pronunciation in Second Language Learning and Teaching Conference, Salt Lake City, UT, USA, 1–2 September 2017; Levis, J.M., Ed.; Iowa State University: Ames, IA, USA, 2018; pp. 151–162.
63. Everitt, C. Accent Imitation on the L1 as a Task to Improve L2 Pronunciation. Master's Thesis, Universitat de Barcelona, Barcelona, Spain, June 2015.
64. Ruellot, V. French stereotypical accent and pronunciation development of /p/,/t/, and /k/. In Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference, Ames, IA, USA, 6–8 September 2018; Levis, J.M., Nagle, C., Todey, E., Eds.; Iowa State University: Ames, IA, USA, 2019; pp. 341–353.
65. Kopečková, R.; Wrembel, M.; Gut, U.; Balas, A. Differences in phonological awareness of young L3 learners: An accent mimicry study. *Int. J. Multiling.* **2021**, *20*, 408–424. [CrossRef]
66. Mora, J.C.; Rochdi, Y.; Kivistö-de-Souza, H. Mimicking accented speech as L2 phonological awareness. *Lang. Aware.* **2014**, *23*, 57–75. [CrossRef]
67. Flege, J.E.; Hammond, R. Mimicry of non-distinctive phonetic differences between language varieties. *SSLA* **1982**, *5*, 1–17. [CrossRef]

68. Henderson, A.; Rojczyk, A. Foreign-language accent imitation: Matching production with perception. In *English Pronunciation Teaching: Theory, Practice, and Research Findings*; Sardegna, V., Jarosz, A., Eds.; Multilingual Matters: Bristol, UK, 2023; pp. 102–117.
69. Rojczyk, A. Using FL accent imitation in L1 in foreign-language speech research. In *Teaching and Researching the Pronunciation of English*; Waniek-Klimczak, E., Pawlak, M., Eds.; Springer: Cham, Switzerland, 2015; pp. 223–233.
70. Charles, T.; Trenkic, D. The effect of bi-modal input presentation on second language listening, The focus on speech segmentation. In *Subtitles and Language Learning: Principles, Strategies and Practical Experiences*; Gambier, Y., Caimi, A., Mariotti, C., Eds.; Peter Lang: Bern, Sweden, 2015; pp. 173–198.
71. Mitterer, H.; McQueen, J.M. Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE* **2009**, *4*, e7785. [CrossRef] [PubMed]
72. Wisniewska, N.; Mora, J.C. Can captioned video benefit second language pronunciation? *SSLA* **2020**, *42*, 599–624. [CrossRef]
73. Galimberti, V.; Mora, J.C.; Gilabert, R. Audio-synchronized textual enhancement in foreign language pronunciation learning from videos. *System* **2023**, *116*, 103078. [CrossRef]
74. Mathias, B.; von Kriegstein, K. Enriched learning: Behavior, brain, and computation. *Trends Cogn. Sci.* **2023**, *27*, 81–97. [CrossRef]
75. Chan, M.J. Embodied pronunciation learning: Research and practice. *Catesol J.* **2018**, *30*, 47–68.
76. Baills, F.; Alazard, C.; Prieto, P. Embodied prosodic training helps improve accentedness and suprasegmental accuracy. *Appl. Linguist.* **2022**, *43*, 776–804. [CrossRef]
77. Li, P.; Xi, X.; Baills, F.; Prieto, P. Training non-native aspirated plosives with hand gestures: Learners' gesture performance matters. *Lang. Cogn. Neurosci.* **2021**, *36*, 1313–1328. [CrossRef]
78. Isaacs, T.; Harding, L. Pronunciation assessment. *Lang. Teach.* **2017**, *50*, 347–366. [CrossRef]
79. Isbell, D.R.; Lee, J. Self-Assessment of Comprehensibility and Accentedness in Second Language Korean. *Lang. Learn.* **2022**, *72*, 806–852. [CrossRef]
80. Strachan, L.; Kennedy, S.; Trofimovich, P. Second language speakers' awareness of their own comprehensibility: Examining task repetition and self-assessment. *JSLP* **2019**, *5*, 347–373. [CrossRef]
81. Li, M.; Zhang, X. A meta-analysis of self-assessment and language performance in language testing and assessment. *Lang. Test.* **2021**, *38*, 189–218. [CrossRef]
82. O'Brien, M.G. Attending to second language lexical stress: Exploring the roles of metalinguistic awareness and self-assessment. *Lang. Aware.* **2019**, *28*, 310–328. [CrossRef]
83. Dlaska, A.; Krekeler, C. Self-assessment of pronunciation. *System* **2008**, *36*, 506–516. [CrossRef]
84. Isbell, D.R. Can the test support student learning? Validating the use of a second language pronunciation diagnostic. *Lang. Assess. Q.* **2021**, *18*, 331–356. [CrossRef]
85. Trofimovich, P.; Isaacs, T.; Kennedy, S.; Saito, K.; Crowther, D. Flawed self-assessment: Investigating self-and other-perception of second language speech. *Biling. Lang. Cogn.* **2016**, *19*, 122–140. [CrossRef]
86. Ma, W.; Winke, P. Self-assessment: How reliable is it in assessing oral proficiency over time? *Foreign Lang. Ann.* **2019**, *52*, 66–86. [CrossRef]
87. Kissling, E.M.; O'Donnell, M.E. Increasing language awareness and self-efficacy of FL students using self-assessment and the ACTFL proficiency guidelines. *Lang. Aware.* **2015**, *11*, 283–302. [CrossRef]
88. Tsunemoto, A.; Trofimovich, P.; Blanchet, J.; Bertrand, J.; Kennedy, S. Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Lang. Ann.* **2022**, *55*, 135–154. [CrossRef]
89. Gurzynski-Weiss, L.; Long, A.; Solon, M. TBLT and L2 Pronunciation: Do the Benefits of Tasks Extend Beyond Grammar and Lexis? [Special issue]. *SSLA* **2017**, *39*, 213–224. [CrossRef]
90. Willis, J. *A Framework for Task-Based Learning*; Longman: Harlow, UK, 1996.
91. Robinson, P. Second language task complexity, the cognition hypothesis, language learning, and performance. In *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*; Robinson, P., Ed.; John Benjamins: Amsterdam, The Netherlands, 2011; pp. 3–37.
92. McKinnon, S. TBLT instructional effects on tonal alignment and pitch range in L2 Spanish imperatives versus declaratives. *SSLA* **2017**, *39*, 287–317. [CrossRef]
93. Loewen, S.; Isbell, D. Pronunciation in face-to-face and audio-only synchronous computer-mediated learner interactions. *SSLA* **2017**, *39*, 225–256. [CrossRef]
94. Loschky, L.; Bley-Vroman, R. Grammar and task-based methodology. In *Tasks and Language Learning: Integrating Theory and Practice*; Crookes, G., Gass, S., Eds.; Multilingual Matters: Clevedon, UK, 1993; pp. 123–167.
95. Gurzynski-Weiss, L.; IATBLT. The TBLT Language Learning Task Bank. Available online: https://tblt.indiana.edu (accessed on 1 April 2023).
96. Mora-Plaza, I.; Saito, K.; Suzukida, Y.; Dewaele, J.-M.; Tierney, A. Tools for Second Language Speech Research and Teaching. Available online: http://sla-speech-tools.com (accessed on 15 December 2022).
97. Ellis, R. *Task-Based Language Learning and Teaching*; Oxford University Press: Oxford, UK, 2003.
98. Pica, T.; Kanagy, R.; Falodun, J. Choosing and Using Communicative Tasks for Second Language Instruction. In *Tasks and Language Learning: Integrating Theory and Practice*; Crookes, G., Gass, S., Eds.; Multilingual Matters: Clevedon, UK, 1993; pp. 9–34.

99. Robinson, P. Situation and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In *Cognitive Processing in Second Language Acquisition: Inside the Learner's Mind*; Putz, M., Sicola, L., Eds.; John Benjamins: Amsterdam, The Netherlands, 2010; pp. 243–268.
100. Cebrian, J. Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. *J. Acoust. Soc. Am.* **2019**, *145*, EL52–EL58. [CrossRef]
101. Rato, A. Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. *Copal* **2014**, *5*, 529–546.
102. Sicola, L. "No, They won't 'Just Sound Like Each Other'": NNS-NNS Negotiated Interaction and Attention to Phonological Form on Targeted L2 Pronunciation Tasks. Ph.D. Thesis, University of Pennsylvania, Philadelphia, PA, USA, January 2008.