How explicit instruction improves phonological awareness and perception of L2 sound contrasts in younger and older adults

Emily Felker, Esther Janse, Mirjam Ernestus, and Mirjam Broersma
Radboud University

Despite the importance of conscious awareness in second language acquisition theories, little is known about how L2 speech perception can be improved by explicit phonetic instruction. This study examined the relationship between phonological awareness and perception in Dutch younger and older adult L2 listeners, focusing on English contrasts of two types: a familiar contrast in an unfamiliar position (word-final /t/-/d/) and an unfamiliar contrast (/æ/-/ɛ/). Awareness was assessed with a task in which written minimal pairs and homophone pairs had to be judged as sounding the same or different. Perception was assessed with a two-alternative forced-choice identification task with auditorily presented words from minimal pairs. We investigated whether listeners' awareness and perception improved after a video-based explicit instruction that oriented their attention to one of these contrasts, and we tested whether including information about the phonetic cue of vowel duration increased learning. Awareness and perception of each contrast were shown to be moderately correlated at the study's outset. Furthermore, awareness and perception for each contrast generally improved more after the instruction drawing attention to that contrast. However, the effectiveness of explicit phonetic instruction varied depending on the combination of the contrast, cue information, and listener age group.

Keywords: speech perception, phonological awareness, explicit instruction, attention

https://doi.org/10.1075/lab.20113.fel | Published online: 2 November 2021
Linguistic Approaches to Bilingualism 13:3 (2023), pp. 372–408. ISSN 1879-9264 | E-ISSN 1879-9272

1. Introduction

In second language (L2) speech perception, one challenge for late bilinguals is learning to distinguish sounds that are not contrastive in their native language (L1). Previous research has shown that intensive exposure to controlled stimuli using high variability phonetic training can improve adult listeners' ability to distinguish novel L2 sound contrasts (see Sakai & Moorman's meta-analysis, 2018). These perception training paradigms, which typically involve lengthy identification tasks with corrective feedback, are theorized to bring about changes in selective attention: over the course of training, listeners shift their attention to the acoustic-phonetic cues that are relevant for a given sound contrast, as a result of the repeated exposure and feedback alone (Francis & Nusbaum, 2002). Interestingly, there is much less research about the effectiveness of bringing relevant phonetic cues to listeners' awareness through explicit instruction, despite the prominent role of explicit instruction in L2 teaching and the importance of awareness in theories of L2 acquisition (e.g., Svalberg, 2007; Tomlin & Villa, 1994). The present study investigates whether a brief explicit instruction can improve both awareness and perception. Moreover, since different types of learners may make different use of explicit instruction, we also test the effectiveness of explicit instruction for two populations, younger adults and older adults, who differ in various respects including auditory and cognitive processing abilities, L2 proficiency, and L2 usage.

Explicit instruction can improve speech perception by orienting listeners' attention to what they need to learn. Several studies have shown that perception of unfamiliar phonemic contrasts can be improved by explicitly directing listeners' attention to sounds rather than semantics, or to specific classes of sounds over others. Guion and Pederson (2007) exposed native English speakers to Hindi minimal word pairs based on Hindi stop consonant contrasts, along with the words' English translations; one participant group was told to attend to the words' sounds and the other to their meanings. For the most difficult contrast tested, the sound-attending group demonstrated greater perceptual discrimination improvement than the meaning-attending group. Similarly, Pederson and Guion-Anderson (2010) gave native English listeners identification training on Hindi words presented auditorily; listeners were instructed to attend to and identify either consonants or vowels. The consonant-attending group, but not the vowel-attending group, showed post-training improvement in consonant discrimination. The effectiveness of attention-directing has also been demonstrated for the learning of tonal contrasts. Chen and Pederson (2017) trained native Mandarin listeners on Quanzhou Southern Min words involving unfamiliar consonant and tonal contrasts, whereby listeners were instructed to attend to and identify either the consonants or the tones. At post-test, the consonant-attending group had only improved in consonant discrimination and the tone-attending group only in tone discrimination. Together, these studies show that directing attention to the target sounds facilitates perceptual learning of non-native contrasts in languages unfamiliar to the listeners.

Some evidence suggests that perceptual learning of non-native sound contrasts benefits from focusing listeners' attention even more narrowly, to the level of specific phonetic cues. Hisagi and Strange (2011) showed that native English speakers were better at discriminating unfamiliar Japanese contrasts of vowel, consonant, and syllable length if they had first received written instructions explaining that duration was what made the words different. Similarly, Porretta and Tucker (2014) found that native English speakers were better at distinguishing unfamiliar Finnish consonants differing in length if they had first received basic written instructions pointing out the difference between short and long consonants. Drawing attention to specific phonetic cues can facilitate learning new sound categories even when the sounds differ in multiple dimensions, such as Mandarin tonal contrasts that differ in both pitch height and direction. Chandrasekaran et al. (2016) showed that short written instructions telling native English listeners to focus on pitch direction, a dimension they would normally underweight, improved their categorization of Mandarin tones more than instructions to focus on pitch height or height and direction together. Thus, listeners can make use of explicit instruction about specific phonetic cues to improve their perception of non-native contrasts in unfamiliar languages.

While the previous studies' instructions provided only a few sentences about the phonetic cue, and tested listeners in an unfamiliar language, Kissling (2014) studied the effect of more intensive phonetic instruction with beginner, intermediate, and advanced learners of Spanish. Over multiple weeks, learners in the phonetic instruction group completed online modules about specific Spanish consonants that explained grapheme-phoneme correspondences, provided articulatory phonetic instructions, and included sound identification exercises. The phonetic instruction group showed greater pre-to-post-test improvement in identification and discrimination of the target sounds than a control group who completed modules with comparable sound exposure but no phonetics instruction. However, one potential confound was that the control group, unlike the phonetic instruction group, was never told which sounds were the target of the study. Thus, while exposure was controlled, the effect of the phonetic information could not be separated from the effect of simply orienting attention to the target sounds. In our study, we will separate out these factors to examine whether explaining a phonetic cue improves perception above and beyond orienting L2 listeners' attention to the critical sounds.

Second language speech perception has only rarely been studied in older adult listeners[1] whose speech processing differs from that of younger adults in various ways due to age-related hearing loss, cognitive decline, and slowed temporal processing (see reviews of Gordon-Salant, 2005; Pichora-Fuller & Souza, 2003). For instance, Sommers (1997) found that older adults were less able than younger adults to ignore phonetically irrelevant stimulus dimensions in speech, implying a breakdown of selective attention. Moreover, older adults have shown less flexibility than younger adults in lexically guided perceptual category learning (Scharenborg & Janse, 2013), but they have been shown to adapt as well as younger adults to acoustically degraded speech, provided equal baseline accuracy between groups (e.g., Peelle & Wingfield, 2005). While many studies suggest that older listeners with normal hearing are quite capable of perceptual learning of speech, the benefits of training are varied and transfer of learning may be limited (see the review of Bieber & Gordon-Salant, 2021). For instance, for time-compressed speech, older listeners have shown comparable perceptual learning, but less transfer of learning to a different speech rate, compared to younger listeners (Peelle & Wingfield, 2005). Overall, these studies show that older adults are capable of implicit perceptual learning but tend to show less selective attention, perceptual flexibility, and transfer of learning than younger adults.

To our knowledge, little research has compared older and younger adults' perceptual learning of L2 speech. The existing crosslinguistic studies with older adults use languages that are unfamiliar to the listeners. For instance, older native Japanese speakers have shown improved perception comparable to that of younger adults after training on English phonemic contrasts not present in Japanese (Kubo & Akahane-Yamada, 2006). More recently, Maddox et al. (2013) investigated the ability of older native English speakers to learn to perceive Mandarin lexical tone categories based on identification training with corrective feedback and found that older adults performed worse overall and learned more slowly than younger adults. Together, these studies demonstrate that training can improve older listeners' perception of non-native contrasts, though older listeners may learn less effectively than younger listeners. How older listeners would respond to explicit phonetic instruction, rather than implicit or feedback-based training, remains an open question.

1. The cited studies use varying age cutoffs to define "older adult," typically 55 or 65 years or older.

1.1 The present study

The present study investigates how explicit instruction improves phonological awareness and perception of L2 sound contrasts in adult listeners. As described above, most previous studies using explicit instruction have focused on teaching non-native listeners about the sounds of a language completely unfamiliar to them, a scenario unlikely to occur outside the laboratory. We study the effect of instruction about L2 contrasts for adults who are already proficient in the L2 and whose phonemic categories may therefore be entrenched after years of language use. Furthermore, we vary the range of difficulty by testing the learning of two L2 sound contrasts that differ in their relation to the L1 sound system. Finally, we extend previous research by testing not only younger adults but also older adults, whose capacity for L2 sound learning and responsiveness to explicit phonetic instruction might be more limited due to various age-related differences in (cognitive) processing.

1.1.1 *Research questions*

We have three main research questions. First, we assess the relationship between L2 listeners' prior perceptual accuracy and their phonological awareness for each of the two contrasts, operationalized as the extent to which they know that minimal pairs based on the contrasts are meant to sound different (RQ1). Then, we investigate whether explicit instruction improves listeners' phonological awareness (RQ2) and perceptual accuracy (RQ3) for the two sound contrasts. In doing so, we test the effect of attention orienting to the contrast in question, the effect of providing information about the phonetic cue of vowel duration, and whether learning differs between the two contrasts and listener age groups.

1.1.2 *Study design*

This study investigates the effect of explicit phonetic instruction on awareness and perception for Dutch younger and older adults, and it focuses on two English contrasts, word-final /t/-/d/ and /æ/-/ɛ/, which should pose differing degrees of difficulty for native Dutch listeners (see motivation below). The phonetic instruction was delivered through a short video in which a native English speaker described the contrast in question and drew attention to relevant minimal pairs. The focus on minimal pairs was inspired by the prominent role of minimal pairs in L2 teaching (Brown, 1995; Field, 2008) and in research about phonological awareness (e.g., Janssen et al., 2015; Krenca et al., 2020). With the instructional video, we aimed to test separately the effects of orienting listeners' attention to the critical sound contrast and to a specific phonetic cue. Therefore, participants were assigned to watch a video in one of four conditions: the video was either about the

/t/-/d/ or the /æ/-/ɛ/ contrast, and it either did or did not explain how the phonetic cue of duration distinguished the sounds.

We chose to focus on Dutch listeners' perception of the English word-final /t/-/d/ contrast and the /æ/-/ɛ/ contrast, and duration as a phonetic cue to distinguish both contrasts, based on previous research. In English, a salient difference between word-final /t/ and /d/ is the preceding vowel's duration: English vowels typically shorten before voiceless consonants, like /t/, and lengthen before voiced consonants, like /d/ (House, 1961). These vowel duration differences are particularly large in English, relative to other languages, and have been shown to be a sufficient cue to the word-final stop consonant voicing distinction, though various other cues, such as vowel formant transitions, closure duration, and release-burst quality can also influence the perception of voicing to varying degrees, depending on the context (Raphael, 2005). For Dutch listeners, the English word-final /t/-/d/ contrast represents a familiar contrast in an unfamiliar position: while both sounds exist in Dutch and are contrastive in word-initial and word-medial position, only /t/ occurs in word-final position (e.g., Booij, 1999). In Dutch, a word-final letter *d* is always pronounced as /t/, unlike in English where a word-final letter *t* maps onto /t/ and *d* onto /d/. While some L2 speech models incorporate phonotactics explicitly (e.g., the revised Speech Learning Model, Flege & Bohn, 2021) and others do not (e.g., the Perceptual Assimilation Model of L2 Speech Learning, Best & Tyler, 2007), they agree that L1 phonotactic constraints negatively affect perception when an L2 sound or contrast occurs in a position where it does not in the L1. Indeed, Dutch listeners' perception of the final /t/-/d/ contrast is less accurate than that of English listeners in lexical processing (Broersma & Cutler, 2008). However, as preceding vowel duration is informative for distinguishing word-*medial* voicing contrasts in Dutch (Slis & Cohen, 1969), Dutch listeners may be able to use this cue for word-*final* voicing contrasts as well. In fact, experiments with phonetically manipulated stimuli have shown that Dutch listeners are capable of exploiting preceding vowel duration as a cue for word-final obstruent voicing contrasts in English, though they do so to a lesser extent than native English listeners, who use the cue persistently even when it is made uninformative (Broersma, 2010).

The English /æ/-/ɛ/ contrast does not exist in Dutch at all, and therefore L2 speech models (e.g., PAM-L2: Best & Tyler, 2007; L2LP: van Leussen & Escudero, 2015) predict it to be even more difficult to perceive (and predict both vowels to be difficult to pronounce (SLM-r: Flege & Bohn, 2021)). Phonetically, /æ/ and /ɛ/ differ in English in both spectral frequency and duration, with /æ/ being longer than /ɛ/ (Hillenbrand et al., 1995). In this phonetic space, Dutch has only one vowel, transcribed as /ɛ/, whose phonetic realization falls between the English /æ/ and /ɛ/ (Collins & Mees, 1996). Dutch listeners have difficulty processing this contrast

(Broersma, 2012), possibly because they assimilate both the English /æ/ and /ɛ/ to their native /ɛ/ category, in accordance with the PAM-L2 (Best & Tyler, 2007). Nevertheless, Dutch listeners are capable of implicitly exploiting duration cues to categorize the English /æ/ and /ɛ/ in phonetically manipulated stimuli (Díaz et al., 2012), and they can also use duration as a cue to distinguish acoustically similar vowels in Dutch (e.g., van der Feest & Swingley, 2011). Thus, it is feasible that drawing attention to vowel duration and its relevance to the non-native /æ/-/ɛ/ contrast via explicit instruction is beneficial for Dutch listeners.

To assess the effectiveness of the instruction for both /t/-/d/ and /æ/-/ɛ/, we employed pre- and post-tests of two kinds: phonological awareness and perception. Phonological awareness is typically studied in relation to literacy development (e.g., with rhyming tasks; see Anthony & Francis, 2005), and we are not aware of studies measuring phonological awareness about specific non-native contrasts in adult L2 learners. We operationalize phonological awareness about a specific phonemic contrast as knowing that words differing minimally in that contrast are meant to sound different from each other. Our phonological awareness pre- and post-tests presented a series of written minimal pairs (e.g., *greet* and *greed*) and filler homophone pairs (e.g., *not* and *knot*). For each pair, participants had to indicate whether they thought the two words sounded the same or different. In the perception pre- and post-tests, each word from the critical minimal pairs was presented auditorily in the context of a two-alternative forced-choice listening task (e.g., hearing */bæg/* and choosing between response options *bag* or *beg*). As orthography can influence L2 phonological acquisition in various ways (Bassetti, Escudero & Hayes-Harb, 2015), the tests were designed to minimize potential reliance on orthographic strategies (see Methods for details).

1.1.3 *Hypotheses*

RQ1 concerns the relationship between phonological awareness of novel L2 sound contrasts and perceptual accuracy for those contrasts. Awareness is theorized to play an important role in L2 acquisition in general (Svalberg, 2007; Tomlin & Villa, 1994), but the correlation between awareness and perceptual accuracy for specific L2 sound contrasts has not yet been empirically demonstrated. We hypothesize that phonological awareness of each contrast (word-final /t/-/d/ or /æ/-/ɛ/) will correlate positively with perceptual accuracy for that contrast. We examine this awareness-perception relationship at pre-test in order to answer this question independently of the instructional intervention.

RQ2 is whether explicit phonetic instruction about a non-native phonemic contrast can increase L2 listeners' phonological awareness. Crucially, we expect that /t/-/d/ awareness will increase more after watching a /t/-/d/ video, and /æ/-/ɛ/ awareness will increase more after an /æ/-/ɛ/ video, as a result of attention orienting to the matching contrast. Whether providing information about the vowel duration cue (which plays a role in both contrasts) will further increase awareness (potentially even of the non-matching contrast) is uncertain. On the one hand, since the videos already very explicitly state that the sounds in question are distinctive, additional information about duration might be superfluous and therefore provide no added benefit. On the other hand, explicit duration information might reinforce phonological awareness by illustrating how the two sounds differ concretely. Additionally, awareness for both contrasts might increase from pre-test to post-test simply because the intervening perception task, which requires listeners to match each critical word they hear to one word label or another, implies that the critical minimal pair words are meant to sound different.

RQ3 is whether explicit phonetic instruction about a non-native contrast can improve L2 listeners' perception. We predict that /t/-/d/ perception will improve more after watching a /t/-/d/ video and /æ/-/ɛ/ perception will improve more after an /æ/-/ɛ/ video. Furthermore, we predict that perception of a given contrast will improve more after a video with the duration cue information for that contrast, as a result of improved phonetic awareness about the importance of the duration cue for that specific contrast.

For both RQ2 and RQ3, we expect more improvements in younger adults than older adults and more learning for word-final /t/-/d/ (a familiar contrast in an unfamiliar position) than for /æ/-/ɛ/ (an unfamiliar contrast). We also expect that, if learning transfers from one contrast to another, the perceived relevance of the phonetic cue could play a role. Specifically, the vowel duration cue in the /æ/-/ɛ/ video context may appear relevant only to vowels and therefore not be generalized to the /t/-/d/ contrast, whereas this cue in the /t/-/d/ video context may be novel enough to suggest that the vowel-pair words should *also* sound different because of duration differences.

2. Methods

2.1 Participants

The participants were 124 monolingually-raised native Dutch speakers: 64 younger adults (72% female) aged 18–31 ($M = 22.3$, $SD = 2.9$) years and 60 older adults with normal hearing[2] (62% female) aged 65–84 ($M = 69.9$, $SD = 4.1$) years, who had started learning English on average at ages 10.7 ($SD = 1.2$, range: 8–14)

2. One additional older adult was tested but excluded from analysis because of having > 35 dB of hearing loss (the threshold to qualify for hearing aids in the Netherlands).

and 12.3 ($SD = 1.1$, range: 10–16) years, respectively. The younger adults spent significantly more hours per week speaking and listening to English than the older adults: for speaking, mean 2.1 hours ($SD = 4.1$) vs. mean 0.6 hours ($SD = 1.1$), $t(72.38) = 2.91$, $p = .005$; and for listening, mean 12.4 hours ($SD = 12.5$) vs. mean 6 hours ($SD = 6.5$), $t(96.85) = 3.60$, $p < .001$). Additionally, the younger adults rated themselves higher than the older adults did on English proficiency across reading, writing, speaking, and listening skills (3.3 vs. 2.8 mean score across four scales from 0 "no ability" to 5 "perfect"; $t(120.17) = 3.88$, $p < .001$).

2.2 Materials

2.2.1 *Phonological awareness test*

The phonological awareness test comprised 100 English word pairs (40 critical pairs and 60 filler pairs) that were either homophones or phonological minimal pairs (see Appendix A). Importantly, the two words in all 100 pairs were always spelled differently from each other, so participants' decisions could not be based on spelling differences alone. The critical pairs were 20 word-final /t/-/d/ minimal pairs (e.g., *feet* and *feed*) and 20 /æ/-/ɛ/ minimal pairs (e.g., *bag* and *beg*). The critical pairs only included words that we expected participants to know; they should thus be able to use their existing phonological representations for the words, rather than having to rely on orthographic knowledge.[3] Dutch and English are highly cognate languages (Schepens, Dijkstra, & Grootjen, 2012), and the critical items varied in oral and orthographic similarity to Dutch words, though we excluded words that were both oral and orthographic cognates. Furthermore, we excluded words that sound similar to Dutch words if their word-final /d/ is interpreted as /t/ or their /æ/ as /ɛ/, as such items might exert a strong influence in this task. Due to the preceding constraints, we could not avoid including some /æ/-/ɛ/ items ending in /t/ or /d/. The variety of phonetic contexts in which the critical sounds occurred reflects the natural language variation to which learners are typically exposed.

By design, it was uncertain how participants would classify the critical pairs: though they should be classified as minimal pairs, we expected many participants to misclassify them as homophones, at least in the pre-test. To keep the proportion of homophone and non-homophone responses relatively balanced regardless of how the critical pairs were classified, the 60 filler items consisted of 30 filler

3. A post-test questionnaire that asked participants "Do you know this word?" (yes/no) for all 80 critical words confirmed that the younger and older adults knew 92.9% and 93.0% of the words, respectively; the lesser known words display regular orthography with respect to the critical sounds (e.g., the /æ/-/ɛ/ pair *radish* - *reddish*).

homophone pairs with spelling contrasts involving vowels and consonants (e.g., *son* and *sun, knight* and *night*) and 30 filler minimal pairs, involving various other vowel and consonant contrasts, in both onset and coda positions (e.g., *play* and *pray*). The four pair types (/t/-/d/, /æ/-/ɛ/, filler minimal pairs, and filler homophone pairs) were equivalent in their mean Zipf frequency in the SUBTLEX-US corpus (Brysbaert & New, 2009); $F(3,196) = 0.22$, $p = .88$.

2.2.2 *Perception test*

The perception test contained 80 individual spoken English words (for 80 trials, using the same tokens in pre-test and post-test) belonging to the same 40 critical minimal pairs (20 word-final /t/-/d/ pairs and 20 /æ/-/ɛ/ pairs) as in the awareness test. The words were recorded with careful pronunciation in random order by a female native speaker of Standard American English in a sound-attenuating booth.

2.2.3 *Phonetic instruction videos*

Four different phonetic instruction videos, two about the word-final /t/-/d/ contrast and two about the /æ/-/ɛ/ contrast, were recorded in English by the same speaker who had recorded the perception stimuli, as perceptual learning does not always generalize to new speakers (e.g., Eisner & McQueen, 2005). For each contrast, one video explained the duration cue and the other did not. Each video drew attention to the relevant contrast and encouraged the viewer to practice listening. Each critical contrast was illustrated with example words that did not occur in the pre- and post-tests and did not contain sounds from the other critical contrast. Whenever the speaker named a critical sound, a corresponding letter *T, D, A*, or *E* appeared briefly onscreen (Figure 1a), and when she pronounced an example word, it appeared briefly onscreen with the critical sound's letter darkened for emphasis (Figure 1b). The speaker maintained a friendly tone of voice and body language throughout.

All videos were approximately two minutes long, providing a short instruction comparable in length to the written explicit instructions used in previous studies (Chandrasekaran et al., 2016; Hisagi & Strange, 2011; Porretta & Tucker, 2014). The four videos' scripts (see Appendix B) were as similar as possible to each other in content, length, and structure and contained the same number of example words.

Each video comprised eight stages. First, the speaker introduced herself as a native speaker, named the critical sounds the video was about, and explained their contrastive role in English using two minimal pair examples. Second, using two non-minimal pair words, she stated the sounds' typical spellings. Third, she men-

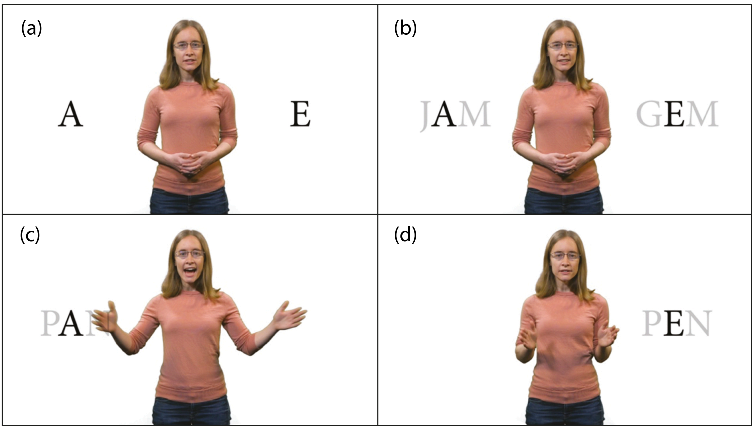

Figure 1. Stills from the /æ/-/ɛ/ duration-cue video

tioned one difference between the two sounds (aspiration for /t/-/d/ and vowel quality for /æ/-/ɛ/) but called it very subtle.

Fourth, the speaker either said that listening to how *long* the vowel was could help the viewer hear the difference (duration-cue videos) or that the sounds were easy to distinguish for native listeners but could be hard for second language learners (no-cue videos).

Fifth, the speaker said she was going to pronounce example words in an exaggerated way. She then pronounced two minimal pairs either by exaggerating short and long vowel length (duration cue videos) or by hyperarticulating the words without exaggerating the vowel length difference, likely affecting multiple acoustic cues involving vowel quality and characteristics of the closure and burst (no-cue videos). In the duration-cue videos only, the exaggerated pronunciations were accompanied by hand gestures emphasizing duration in which the palms began together and moved horizontally outward, either far beyond the body (for the longer /d/ and /æ/ words; Figure 1c) or shoulder-width apart (for the shorter /t/ and /ɛ/ words; Figure 1d).

Sixth, the speaker said she would pronounce the words more normally and invited the viewer to listen closely and try to hear the difference. In the duration cue videos only, she then described the duration cue explicitly by stating that the vowel before /d/ sounded longer than the vowel before /t/ (/t/-/d/ duration-cue video) or that /æ/ sounded longer than /ɛ/ (/æ/-/ɛ/ duration-cue video).

Seventh, she pronounced two minimal pairs, each repeated three times; this time, the words did not appear onscreen until the third pronunciation to allow listeners to test their comprehension without visual support.

Finally, the videos concluded by reiterating the duration difference for the relevant contrast (duration-cue videos; e.g., "just remember: the /æ/ sounds longer while the /ɛ/ sounds shorter") or by stating that the difference would become easier to hear with practice (no-cue videos).

2.3 Procedures

Participants were tested individually in a sound-attenuating booth. The older adults were first screened for hearing acuity with an Oscilla audiometer using an automated Hughson-Westlake procedure to obtain a pure-tone average threshold for each ear at 500, 1000, and 2000 Hz (air conduction only). All participants then completed the main tasks in the fixed order shown in Figure 2. To preclude raising participants' awareness of the contrasts prior to the first awareness test, the awareness pre-test preceded the perception pre-test. As perceptual learning is susceptible to unlearning with intervening tasks, the perception post-test immediately followed the instruction.

To minimize interaction with the experimenter, thereby limiting extraneous speech exposure, participants received written onscreen English-language instructions for each task. They wore Sennheiser over-ear headphones for the perception tests and instruction video. Responses were made with a button box, and all test trials were self-paced.

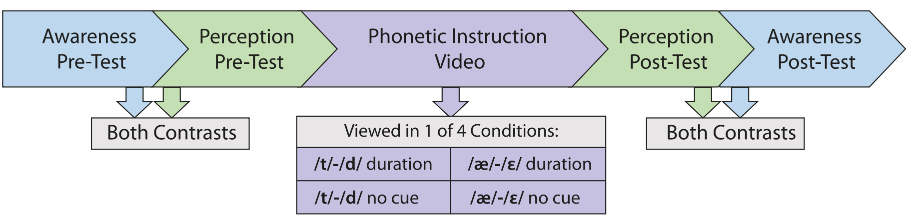

Figure 2. Order of the main tasks within each experimental session

2.3.1 *Pre-Tests and post-tests*

The pre-tests and post-tests each consisted of the same trials (100 phonological awareness trials and 80 perception trials) presented in a different randomized order for each participant and test time. In each trial, a word pair was displayed with one word on each side of the screen. The words' left-right positioning was counterbalanced across participants, and for each participant, words with the same critical sound always appeared on the same side to reduce the need to re-analyze spelling on each trial. In the awareness tests, the instruction was to answer the question "Do these words sound the same or different?" by pressing the button labeled "same" or "different." In the perception tests, the audio recording of the target word played automatically after a one-second delay; the instruction was to answer the question "Which word did you hear?" by pressing either the "left" or "right" labeled button.

2.3.2 *Phonetic instruction video*

Each participant watched one of the four phonetic instruction videos: the /t/-/d/ duration-cue video, the /t/-/d/ no-cue video, the /æ/-/ɛ/ duration-cue video, or the /æ/-/ɛ/ no-cue video. Videos were assigned to participants on a rotating basis, resulting in 16 younger adults and 14–16 older adults per condition. During the video, participants' only task was to pay attention. After the video, onscreen text instructed participants to try to apply what they had learned in the post-tests.

3. Results

3.1 Relationship between awareness and perception (RQ1)

To assess the relationship between participants' phonological awareness and perception of the /t/-/d/ and /æ/-/ɛ/ contrasts, we analyzed their performance in the awareness and perception tasks at pre-test. For each task, we calculated each participant's overall accuracy for the /t/-/d/ and /æ/-/ɛ/ items separately. Then, we computed correlations between the awareness and perception of each contrast type, using Spearman's rank correlations since the scores were not normally distributed. Figure 3 presents the results visually. As predicted, the pre-test data revealed significant (small-to-moderate) positive correlations between phonological awareness and perceptual accuracy for both /t/-/d/ ($\rho = 0.54$, $p < .001$) and /æ/-/ɛ/ ($\rho = 0.34$, $p < .001$).

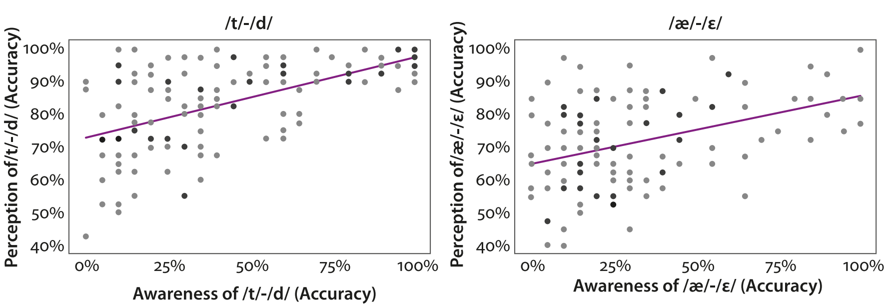

Figure 3. Pre-test correlations between phonological awareness and perceptual accuracy for /t/-/d/ (left) and /æ/-/ɛ/ (right)

3.2 Effect of phonetic instruction on awareness (RQ2)

The phonological awareness results are presented graphically in Figure 4, with pre-to-post-test change scores in Appendix C (Table C1). To examine the effect of instruction on phonological awareness, we analyzed response accuracy for the critical minimal pairs in the awareness pre-tests and post-tests using generalized linear mixed effects models with the logit link function from the *lme4* package in R (Bates et al., 2015). The binary dependent variable was accuracy (correct vs. incorrect). The random effects were item and participant (with random intercepts only, since random slopes prevented convergence). The fixed effects were age group (younger vs. older adults), contrast for the item in question (/t/-/d/ vs. /æ/-/ɛ/), test time (pre-test vs. post-test), video contrast (/t/-/d/ vs. /æ/-/ɛ/), video duration cue information (duration vs. no cue), and all possible interactions between these factors.

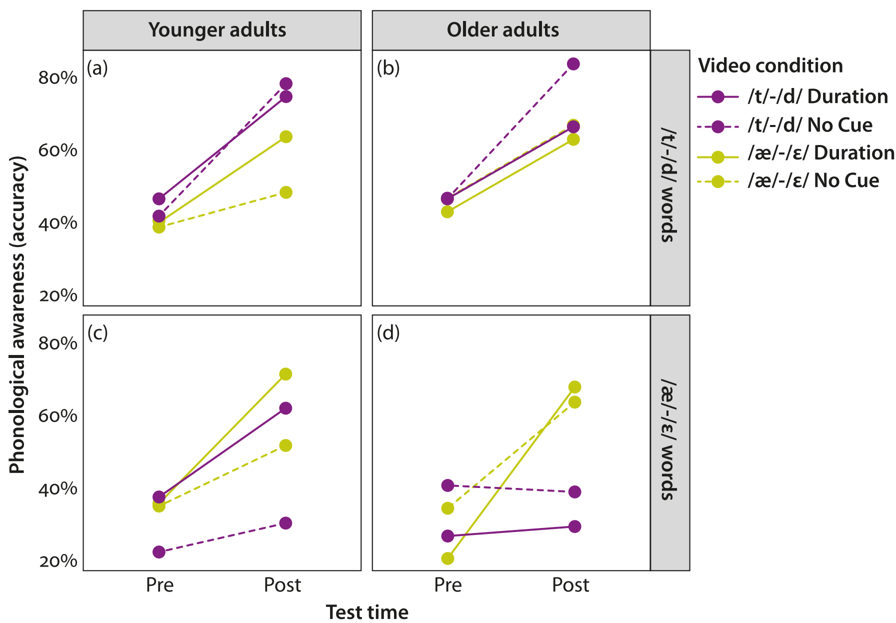

Figure 4. Changes in awareness from pre-test to post-test for the /t/-/d/ words (top row) and /æ/-/ɛ/ words (bottom row) for each combination of listener group, video contrast, and cue information

The ANOVA table for the full statistical model is shown in Appendix C (Table C2). Since this model contained three significant four-way interactions and numerous significant lower-level interactions, we split the data by age group (which factored into all of the significant four-way interactions) and calculated separate models for the younger and older adults (Table C3). Again, these models contained many significant interaction effects, so to facilitate their interpretation, we split the data for each group by item contrast (as this factored into all three significant three-way interactions for younger adults and into two of the three significant three-way interactions for older adults).

The following subsections present the results in detail for each age group and item contrast. In all tables, the beta coefficients represent logits, which are equivalent to the logarithm of the odds ratio (*OR*). We are primarily interested in the how the odds of making a correct (vs. incorrect) response may differ as a function of test time, indicating whether participants are more likely to answer correctly at post-test than at pre-test. Therefore, we also report the odds ratios for all significant effects of test time, which are calculated by exponentiating their beta coefficients (or, for significant two- or three-way interaction effects involving test time, by exponentiating the sum of the beta coefficients of the interaction effect, the simple effect of test time, and any lower-level interaction effects also involving test time).

3.2.1 *Awareness in younger adults*

For younger adults (see Figure 4a and 4c), the separate models for /t/-/d/ and /æ/-/ɛ/ awareness, with treatment coding, are presented in Table 1. For the /t/-/d/ items, younger adults showed a significant effect of test time: their accuracy was greater in the post-test than the pre-test ($OR = e^{0.66} = 1.93$, meaning that participants were nearly twice as likely to answer correctly in the post-test as in the pre-test); this effect held for all four videos. As expected, they showed a significant interaction effect between test time and video contrast indicating that the post-test increase in accuracy was greater for the /t/-/d/ videos than for the /æ/-/ɛ/ videos ($OR = e^{(0.66+1.62)} = 9.78$, meaning that participants were nearly ten times more likely to be accurate at post-test than at pre-test after watching the /t/-/d/ video). Additionally, a significant interaction between test time and cue information indicated that the pre-to-post-test improvement was greater for the duration-cue videos than the no-cue videos ($OR = e^{(0.66+0.57)} = 3.42$). The larger odds ratio for the former interaction effect suggests that the matching video contrast was more beneficial than the presence of duration cue information.

For the /æ/-/ɛ/ items, younger adults again showed a significant main effect of test time indicating that awareness increased from pre-test to post-test ($OR = e^{0.53} = 1.70$). The interaction between test time and video contrast, which would have shown the /æ/-/ɛ/ videos to lead to more improvement than the /t/-/d/ videos, did not reach significance ($p = .08$). However, there was a significant interaction between test time and cue information: the pre-to-post-test improvement was present for both types of videos but greater for the duration-cue videos than the no-cue videos ($OR = e^{(0.53+1.25)} = 5.93$).

Table 1. Models predicting awareness accuracy in younger adults (Item Contrasts separated)

Model for /t/-/d/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	−0.71	0.52	.17	[−1.73, 0.31]
Post-Test	0.66	0.21	.002*	[0.25, 1.07]
/t/-/d/ Video	0.20	0.71	.78	[−1.18, 1.59]
Duration Cue	0.21	0.70	.76	[−1.17, 1.59]
Post-Test • /t/-/d/ Video	1.62	0.31	<.001*	[1.02, 2.23]
Post-Test • Duration Cue	0.57	0.28	.04*	[0.02, 1.12]
/t/-/d/ Video • Duration Cue	−0.002	1.00	1.00	[−1.97, 1.96]
Post-Test • /t/-/d/ Video • Duration Cue	−0.54	0.44	.22	[−1.40, 0.32]
Random Effects	**Variance**			
Participant	3.63			
Item	0.40			
Model for /æ/-/ɛ/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	−1.70	0.47	<.001*	[−2.62, −0.78]
Post-Test	0.53	0.20	.01*	[0.13, 0.92]
/æ/-/ɛ/ Video	0.74	0.64	.25	[−0.52, 2.00]
Duration Cue	0.60	0.65	.35	[−0.67, 1.87]
Post-Test • /æ/-/ɛ/ Video	0.50	0.28	.08	[−0.05, 1.06]
Post-Test • Duration Cue	1.25	0.30	<.001*	[0.65, 1.84]
/æ/-/ɛ/ Video • Duration Cue	−0.49	0.91	.59	[−2.27, 1.30]
Post-Test • /æ/-/ɛ/ Video • Duration Cue	−0.24	0.42	.57	[−1.06, 0.58]
Random Effects	**Variance**			
Participant	2.91			
Item	0.25			

Note. *SE* = standard error, CI = confidence interval,
* significant.

3.2.2 *Awareness in older adults*

For older adults (see Figure 4b and 4d), the separate models for /t/-/d/ and /æ/-/ɛ/ awareness are presented in Table 2. For the /t/-/d/ items, the older listeners showed a significant simple effect of test time indicating higher accuracy in the post-test than the pre-test ($OR = e^{1.41} = 4.10$). The significant two-way interaction between test time and video contrast, indicating greater pre-to-post-test improvement for the /t/-/d/ video condition ($OR = e^{(1.41+1.63)} = 20.91$), was only significant for the no-cue videos (mapped onto the intercept), as revealed by the significa-tion three-way interaction between test time, video contrast, and cue information ($OR = e^{(1.41+1.63+0.09-1.41)} = 5.58$ for the duration-cue /t/-/d/ videos).

For the /æ/-/ɛ/ items, older listeners showed no significant simple effect of test time, but they did show a significant interaction between test time and video contrast, indicating that there was a significant pre-test to post-test improvement within the /æ/-/ɛ/ video condition ($OR = e^{(-0.10+1.89)} = 5.99$) but not the /t/-/d/ video condition.

The fact that the three-way interaction between test time, video contrast, and cue information appears significant for /t/-/d/ but not for /æ/-/ɛ/ should be interpreted with caution, given that the four-way interaction including item contrast was not significant in the parent model (Table C3). Thus, it cannot be firmly concluded that duration cue information affects awareness gains differently for the /t/-/d/ items in the /t/-/d/ video condition than it does for the /æ/-/ɛ/ items in the /æ/-/ɛ/ video condition. Table 2 just suggests that the three-way interaction effect in the parent model indicating a negative effect of duration cue information is driven by the /t/-/d/ video condition for the /t/-/d/ items.

3.2.3 *Summary of awareness results*

For younger adults, /t/-/d/ and /æ/-/ɛ/ awareness increased from pre-test to post-test in all conditions, and it increased more after duration-cue videos than no-cue videos. For /t/-/d/, but not /æ/-/ɛ/, awareness also increased more after watching a video about the matching contrast.

For the older adults, /t/-/d/ awareness increased in all conditions but more after the /t/-/d/ videos than the /æ/-/ɛ/ videos; moreover, it was specifically the /t/-/d/ no-cue video that raised /t/-/d/ awareness more than the other three videos. Older adults' /æ/-/ɛ/ awareness improved at post-test only after the /æ/-/ɛ/ videos.

Table 2. Models predicting awareness accuracy in older adults (Item Contrasts separated)

Model for /t/-/d/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	−0.13	0.66	.83	[−1.42, 1.15]
Post-Test	1.41	0.23	<.001*	[0.97, 1.86]
/t/-/d/ Video	0.08	0.93	.93	[−1.73, 1.90]
Duration Cue	−0.29	0.89	.74	[−2.04, 1.46]
Post-Test • /t/-/d/ Video	1.63	0.37	<.001*	[0.90, 2.36]
Post-Test • Duration Cue	0.09	0.32	.77	[−0.54, 0.72]
/t/-/d/ Video • Duration Cue	0.11	1.29	.93	[−2.43, 2.64]
Post-Test • /t/-/d/ Video • Duration Cue	−1.41	0.51	.01*	[−2.40, −0.42]
Random Effects	**Variance**			
Participant	5.73			
Item	0.43			
Model for /æ/-/ɛ/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	−0.45	0.47	.34	[−1.38, 0.48]
Post-Test	−0.10	0.23	.64	[−0.55, 0.34]
/æ/-/ɛ/ Video	−0.47	0.63	.45	[−1.71, 0.76]
Duration Cue	−0.97	0.63	.13	[−2.21, 0.27]
Post-Test • /æ/-/ɛ/ Video	1.89	0.31	<.001*	[1.28, 2.49]
Post-Test • Duration Cue	0.29	0.31	.35	[−0.32, 0.90]
/æ/-/ɛ/ Video • Duration Cue	−0.03	0.88	.98	[−1.74, 1.70]
Post-Test • /æ/-/ɛ/ Video • Duration Cue	0.74	0.44	.09	[−0.12, 1.59]
Random Effects	**Variance**			
Participant	2.47			
Item	0.34			

Note. SE = standard error, CI = confidence interval,
* significant.

3.3 Effect of phonetic instruction on perception (RQ3)

The perception results are presented graphically in Figure 5, with pre-to-post-test change scores in Appendix D (Table D1). To analyze the effect of phonetic instruction on perceptual accuracy, we computed generalized linear mixed effects models using the same model structures as described above for the awareness data analysis. The ANOVA table for the full statistical model is presented in Appendix D (Table D2). As the five-way interaction between all factors was significant, we again split the data by age group. The ANOVA tables for each age group's separate model are shown in Appendix D (Table D3). For both age groups, the item contrast factored into the highest-level significant interaction effect (a three-way interaction for younger adults and a four-way interaction for older adults), so we split the data further by item contrast. The following subsections describe in detail the results for each age group and item contrast, with odds ratios illustrating the significant effects of test time the same way as in Section 3.2.

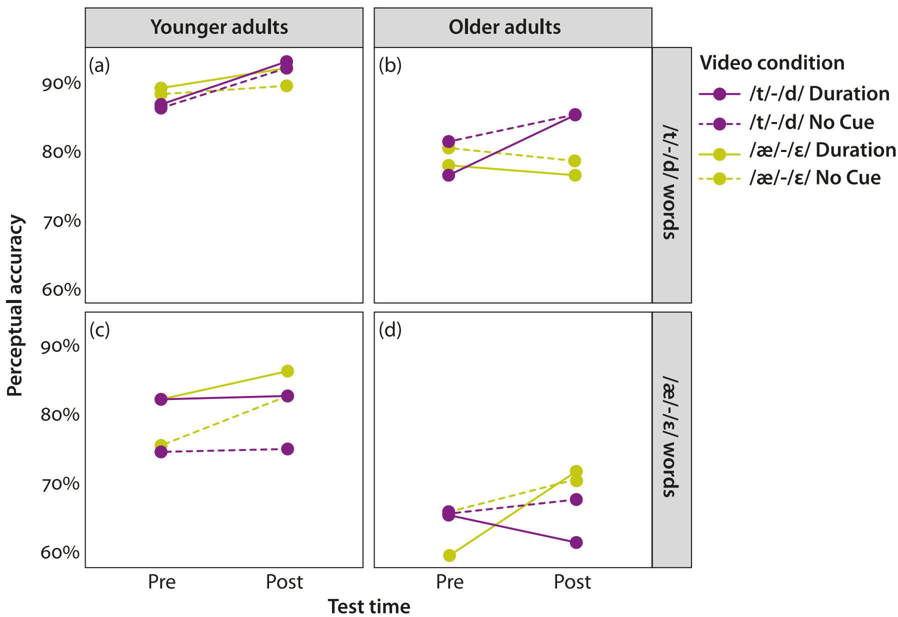

Figure 5. Changes in perceptual accuracy from pre-test to post-test for the /t/-/d/ words (top row) and /æ/-/ɛ/ words (bottom row) for each combination of age group, video contrast, and cue information

3.3.1 *Perception in younger adults*

For younger adults (see Figure 5a and 5c), the separate models for perceptual accuracy for /t/-/d/ and /æ/-/ɛ/ items are presented in Table 3. The only significant effect within the /t/-/d/ model is a significant two-way interaction between test time and video contrast indicating that listeners in the /t/-/d/ video condition, but not in the /æ/-/ɛ/ video condition, improved in t/-/d/ accuracy from the pre-test to the post-test ($OR = e^{(0.16+0.60)} = 2.14$). Similarly, within the /æ/-/ɛ/ model, the only significant effect was the interaction between test time and video contrast, indicating that listeners in the /æ/-/ɛ/ video condition, but not in the /t/-/d/ condition, improved in /æ/-/ɛ/ accuracy from the pre-test to the post-test ($OR = e^{(0.01+0.49)} = 1.65$). Neither model showed any significant effects involving cue information.

Table 3. Models predicting perception accuracy in younger adults (Item Contrasts separated)

Model for /t/-/d/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	2.72	0.34	<.001*	[2.07, 3.39]
Post-Test	0.16	0.20	.41	[−0.22, 0.55]
/t/-/d/ Video	−0.30	0.41	.47	[−1.11, 0.51]
Duration Cue	−0.05	0.42	.91	[−0.86, 0.77]
Post-Test • /t/-/d/ Video	0.60	0.28	.03*	[0.05, 1.15]
Post-Test • Duration Cue	0.25	0.28	.39	[−0.31, 0.80]
/t/-/d/ Video • Duration Cue	0.26	0.59	.66	[−0.89, 1.41]
Post-Test • /t/-/d/ Video • Duration Cue	−0.12	0.41	.77	[−0.91, 0.68]
Random Effects	**Variance**			
Participant	1.05			
Item	1.01			
Model for /æ/-/ɛ/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	1.27	0.23	<.001*	[0.83, 1.72]
Post-Test	0.01	0.14	.94	[−0.26, 0.28]
/æ/-/ɛ/ Video	0.08	0.29	.77	[−0.48, 0.64]
Duration Cue	0.56	0.29	.05	[−0.01, 1.13]
Post-Test • /æ/-/ɛ/ Video	0.49	.20	.01*	[0.10, 0.89]
Post-Test • Duration Cue	0.03	0.21	.90	[−0.38, 0.44]
/æ/-/ɛ/ Video • Duration Cue	−0.14	0.41	.73	[−0.95, 0.66]
Post-Test • /æ/-/ɛ/ Video • Duration Cue	−0.17	0.30	.56	[−0.77, 0.42]
Random Effects	**Variance**			
Participant	0.50			
Item	0.41			

Note. *SE* = standard error, CI = confidence interval,
* significant.

3.3.2 *Perception in older adults*

For older adults (see Figure 5b and 5d), the separate models for perceptual accuracy for /t/-/d/ and /æ/-/ɛ/ items are presented in Table 4. Within the /t/-/d/ model, the only significant effect was the two-way interaction between test time and video contrast, indicating that listeners in the /t/-/d/ video condition, but not the /æ/-/ɛ/ video condition, improved in /t/-/d/ accuracy from pre-test to post-test ($OR = e^{(-0.14+0.49)} = 1.42$). Within the /æ/-/ɛ/ model, the only significant effect was the three-way interaction between test time, video contrast, and duration cue, which shows that only listeners who watched the /æ/-/ɛ/ duration-cue video improved in /æ/-/ɛ/ accuracy from pre-test to post-test ($OR = e^{(0.09+0.16-0.30+0.65)} = 1.82$).

Table 4. Models predicting perception accuracy in older adults (Item Contrasts separated)

Model for /t/-/d/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	1.91	0.37	<.001*	[1.17, 2.64]
Post-Test	−0.14	0.16	.37	[−0.45, 0.17]
/t/-/d/ Video	0.25	0.50	.62	[−0.73, 1.24]
Duration Cue	−0.16	0.48	.74	[−0.10, 0.79]
Post-Test • /t/-/d/ Video	0.49	.24	.04*	[−0.02, 0.96]
Post-Test • Duration Cue	0.04	0.22	.86	[−0.39, 0.47]
/t/-/d/ Video • Duration Cue	−0.23	0.70	.74	[−1.60, 1.13]
Post-Test • /t/-/d/ Video • Duration Cue	0.39	0.33	.24	[−0.26, 1.04]
Random Effects	**Variance**			
Participant	1.58			
Item	0.77			
Model for /æ/-/ɛ/ Items				
Fixed effects	**β**	***SE***	***p*-value**	**95% CI**
(Intercept)	0.78	0.20	<.001*	[0.39, 1.18]
Post-Test	0.09	0.14	.49	[−0.17, 0.36]
/æ/-/ɛ/ Video	−0.05	0.24	.84	[−0.53, 0.43]
Duration Cue	−0.03	0.24	.91	[−0.51, 0.45]
Post-Test • /æ/-/ɛ/ Video	0.16	0.19	.41	[−0.21, 0.53]
Post-Test • Duration Cue	−0.30	0.19	.11	[−0.67, 0.07]
/æ/-/ɛ/ Video • Duration Cue	−0.25	0.34	.45	[−0.92, 0.41]
Post-Test • /æ/-/ɛ/ Video • Duration Cue	0.65	0.26	.01*	[0.14, 1.16]

Table 4. *(continued)*

Random Effects	**Variance**
Participant	0.30
Item	0.41

Note. SE = standard error, CI = confidence interval,
* significant.

3.3.3 *Summary of perception results*

The perception models show that both the /t/-/d/ and /æ/-/ɛ/ videos improved younger adults' perception of the featured contrast, regardless of whether the duration cue was mentioned, whereas their perception did not improve at post-test for the contrast not featured in the video. The older adults performed similarly to the younger adults for /t/-/d/, demonstrating improved post-test perception after either /t/-/d/ video. However, their perceptual learning for /æ/-/ɛ/ was more limited, as they only improved at post-test after the /æ/-/ɛ/ duration-cue video.

4. Discussion

The purpose of this study was to examine the relationship between phonological awareness and perception of non-native contrasts and to investigate whether explicit phonetic instruction improves awareness and perception in younger and older adult L2 listeners. To this end, we analyzed Dutch listeners' awareness and perception of two difficult English contrasts (word-final /t/-/d/ and /æ/-/ɛ/) before and after they watched a short video that used minimal pairs to explain one contrast, either /t/-/d/ or /æ/-/ɛ/, and either did or did not explain the phonetic cue of duration.

First, we assessed the relationship between phonological awareness and perception by determining their correlation for each L2 sound contrast at the outset of the experiment. In our study, awareness was operationalized as the proportion of minimal pairs with the contrast (presented visually) for which the two words were correctly judged as sounding different. As hypothesized, we found positive (small-to-moderate) correlations between /t/-/d/ awareness and /t/-/d/ perceptual accuracy and between /æ/-/ɛ/ awareness and /æ/-/ɛ/ perceptual accuracy.[4] While L2 acquisition researchers have theorized that L2 learning is closely linked

4. Recall that pre-test awareness was always measured before pre-test perception, and we do not directly compare the awareness and perception absolute scores.

to conscious awareness of specific L2 forms (Svalberg, 2007; Tomlin & Villa, 1994), to our knowledge this is the first time that the correlation between awareness and perception of specific L2 sound contrasts has been established. This correlation might arise because being able to perceive the difference between two L2 sounds makes people more likely to label them as different, and conversely, being aware that two L2 sounds are meant to differ may be a crucial step on the path of learning to perceive that difference. Our methodology for testing awareness discouraged orthography-based response strategies, but as previous research has shown that phonological learning and perception are related to orthography (Bassetti et al., 2015), orthographic awareness could explain part of the correlation observed here.

Second, we examined whether explicit phonetic instruction increased younger and older adults' phonological awareness of the two contrasts. The main question was whether phonological awareness would increase more for the contrast that was featured in the instructional video. This beneficial effect of attention orienting was indeed borne out for /t/-/d/ in younger adults and for both /t/-/d/ and /æ/-/ɛ/ in older adults: in these cases, awareness for the given contrast improved more at post-test among those who had received instruction about the matching contrast. The fact that younger adults did not quite show a significant effect of video contrast for /æ/-/ɛ/ ($p = 0.08$) seems to arise from the strength of the duration-cue transfer effect: specifically, the fact that their /æ/-/ɛ/ awareness, which improved to some degree in all four conditions, also improved remarkably from the /t/-/d/ duration-cue video. Thus, hearing about vowel duration differences, even in the context of a consonant contrast, may have been enough to trigger awareness that words with the critical vowels also ought to sound different.

The direct effect of the duration cue information on awareness-raising differed by age group. Younger adults gained more awareness about both /t/-/d/ and /æ/-/ɛ/ from the duration-cue videos than from the no-cue videos. As these duration-cue effects for both contrasts did not interact significantly with video contrast, it appears that learning about the vowel duration cue increased awareness for both contrasts regardless of the context in which the cue was presented. The placement of the perception post-test between the video instruction and awareness post-test could have supported this generalization of learning by making the younger adults more likely to notice vowel length in both the /t/-/d/ and /æ/-/ɛ/ words they heard, which subsequently made them more likely to classify them as sounding different. The older adults, in contrast to the younger adults, did not benefit from the duration cue for either contrast. In fact, while the effect of the cue information on /æ/-/ɛ/ is unclear, the older adults' awareness of the /t/-/d/ contrast seemed to improve significantly more after watching the /t/-/d/ *no-cue* video than after the /t/-/d/ duration-cue video. Thus, not only did the vowel duration information provide no added benefit, it may have even been confusing for older adults, at least when presented in the consonant context where it may have had less perceived relevance.

In addition to the aforementioned age-group differences in how phonological awareness is affected by the duration cue information, there was one more age effect in the awareness gains: while younger adults' awareness of both contrasts improved at post-test in all four conditions, older adults' awareness gains were more limited. Specifically, their /æ/-/ɛ/ awareness did not improve at all after watching a /t/-/d/ video. Thus, while older adults gained awareness from the /æ/-/ɛ/ attention-orienting instruction as expected, they did not gain /æ/-/ɛ/ awareness simply through completing the intervening perception tests nor through transferring vowel length information from the /t/-d/ duration-cue video. This matches our expectation that awareness for the unfamiliar contrast (/æ/-/ɛ/), especially in the absence of an explicit instruction orienting attention to that contrast, would be less likely to increase for the age group that tends to show limited transfer of perceptual learning to untrained stimuli (Bieber & Gordon-Salant, 2021).

Finally, we examined whether explicit phonetic instruction improved younger and older adults' perception of the two contrasts. The most important question was whether perception of the L2 contrast would improve after instruction about that contrast, which would support the benefit of attention orienting on perception. This effect was clearly borne out for younger adults: their /t/-/d/ perception improved after /t/-/d/ (but not /æ/-/ɛ/) instruction, and their /æ/-/ɛ/ perception improved after /æ/-/ɛ/ (but not /t/-/d/) instruction. These results align with previous findings that perception of sounds in an unfamiliar language improves following instruction to focus on those sounds specifically during training (Pederson & Guion-Anderson, 2010; Chen & Pederson, 2017). We have shown that this effect also holds for highly proficient L2 listeners. Our older adults' /t/-/d/ perception also improved after /t/-/d/ (but not /æ/-/ɛ/) instruction. In contrast, their /æ/-/ɛ/ perception did not improve more from /æ/-/ɛ/ instruction compared to /t/-/d/ instruction overall, and it only improved significantly after the /æ/-/ɛ/ duration-cue instruction. This more limited learning for the /æ/-/ɛ/ contrast in older listeners aligns with our general expectation of finding fewer learning effects for the unfamiliar contrast – as predicted by the PAM-L2 (Best & Tyler, 2007) and the L2LP (van Leussen & Escudero, 2015) – and for the older age group.

We had hypothesized that providing information about the vowel duration cue for a given contrast would generally improve perception of that contrast. However, the only significant effect of duration information on perceptual improvement was the aforementioned benefit of the /æ/-/ɛ/ duration-cue instruction over the /æ/-/ɛ/ no-cue instruction for older listeners, and there were no transfer effects showing that duration cue information from one contrast improved perception of another contrast. The fact that we only found a benefit of duration-cue instruction for /æ/-/ɛ/ and not for /t/-/d/ suggests that the more intuitive pairing of the vowel-duration cue with a vowel contrast may be more effective than the pairing of the vowel-duration cue with a consonant contrast, even if the consonant contrast itself is more familiar to the listeners. A possible explanation for the relatively limited benefit of the duration cue information is that duration is just one of multiple cues that distinguishes the contrasts in this study, whereas previous studies that showed perceptual improvement following explicit instruction about duration used contrasts that differed *only* in duration (Hisagi & Strange, 2011; Porretta & Tucker, 2014). Another possible explanation is that our listeners already had relatively high exposure to and proficiency in the language containing the sound contrasts, unlike in previous studies that used unfamiliar languages, for which any information would be new and therefore likely to be impactful (Hisagi & Strange, 2011; Porretta & Tucker, 2014). Because of their substantial L2 listening experience, our listeners might have already developed and entrenched personal strategies for distinguishing these contrasts that could be difficult to override. A third possible explanation is that our study's vowel duration cue is already prominent in our listeners' L1 (Booij, 1995), in contrast to Chandrasekaran et al. (2016)'s study that employed instruction about a cue that was absent in their listeners' L1. It might be that, compared to learning to use a novel cue in a novel language, the task of learning to apply an existing L1 cue in new contexts is more complicated due to interference from how that cue is already used in the native phonological system.

The fact that listeners generally showed equivalent perceptual improvement for the instructed contrast with or without the duration-cue information (except for /æ/-/ɛ/ among older adults) suggests that the repeated exposure and attention-orienting to minimal pairs is what likely drove the improvement. This study was not designed to assess which perceptual cues listeners actually relied on. While our conditions manipulated the presence of instruction about the duration cue, listeners in all conditions could conceivably have unconsciously improved their use of various cues, including vowel duration but also vowel quality (for /æ/-/ɛ/) and vowel formant transitions, closure duration, and release burst (for /t/-/d/), as a result of the attention-orienting instruction.

As mentioned above, the only age-related difference in perceptual learning was that older adults showed more limited improvement for /æ/-/ɛ/ than younger adults by failing to improve from the /æ/-/ɛ/ no-cue instruction. This aligns with our expectation that less perceptual learning would take place for the unfamiliar contrast in the older listener group. Given previous research attesting to older listeners' breakdown of selective attention for phonetically relevant stimulus dimensions in speech (Sommers, 1997), older adults might have required more or different instruction than younger adults in order to show equivalent perceptual improvement. Despite the age-group differences in perceptual learning for /æ/-/ɛ/, both groups showed the same pattern of results for /t/-/d/, consistent with Kubo and Akahane-Yamada's (2006) findings of equivalent L2 perceptual learning between older and younger adults after perceptual training. Interestingly, older listeners in our study were the only ones who benefited from the duration-cue phonetic instruction over the no-cue instruction for /æ/-/ɛ/. Thus, despite the negative effect of the vowel duration information on older adults' *awareness*, at least for the /t/-/d/ contrast, such information was apparently helpful for their *perceptual* learning of the /æ/-/ɛ/ contrast. This suggests that when information about a phonetic cue has high perceived relevance, as vowel duration does for a vowel contrast, older listeners are quite capable of using it to improve their perception of an L2 contrast.

Overall, this study's explicit phonetic instruction combined multiple components that each potentially contributed to awareness gains and perceptual category learning: the presentation of minimal word pairs involving the critical L2 sounds, the description of the sounds' contrastive role, the phonetic cue information, the listening practice with exaggerated pronunciations, and even the exposure to the native speaker's voice. This study varied the presence of the duration cue information, following up on previous research about the benefits of instruction about non-native phonetic cues (Chandrasekaran et al., 2016; Hisagi & Strange, 2011; Porretta & Tucker, 2014). Further work is needed to determine which of the instruction's other elements also impacts phonetic learning. Moreover, future research could determine whether the learning effects observed here will generalize to listeners' perception of other speakers, how long the awareness and perceptual gains will persist, and whether there is a connection between awareness-raising and perception in more naturalistic settings.

One goal of this study was to investigate the effectiveness of explicit instruction for two different populations and to explore whether they would respond to the instruction differently. We did find various differences in learning between younger and older adults. These learning differences could potentially be explained by various characteristics of these particular groups, such as generational differences in their L2 acquisition context, exposure, and usage, as well as age-related differences in auditory or cognitive processing, including those related to perception, information processing, and learning. Moreover, the groups differed in L2 proficiency, which could have affected their comprehension of the instruction itself. Future research would be needed to disentangle these explanations by testing listeners that differ minimally except on the dimension of interest. It is likely that learner characteristics will always impact the effect of experimental manipulations, and therefore research in second language acquisition would benefit from including more diverse populations. As for practical implications, our research supports the common practice in educational settings of using a range of different activities and instructional materials, as their benefits may be different for different types of learners.

In conclusion, we have shown that a brief explicit phonetic instruction can improve phonological awareness and perception of L2 sound contrasts in younger and older adult listeners. In doing so, we tested two L2 contrasts that varied in their relation to the L1 phonemic system and investigated the effect of including information about the phonetic cue of vowel duration in the instruction. First, we established the correlation between listeners' initial awareness and perception of specific L2 contrasts. Second, we demonstrated that phonological awareness generally increased more for the contrast that was featured in the instruction, thereby showing that attention-orienting enhances awareness. Moreover, while younger adults generalized the phonetic cue information to also increase their awareness of a non-attended contrast, older adults showed few transfer effects in awareness. Finally, we showed that for younger adults, explicit phonetic instruction for a given contrast improved perception of that contrast, regardless of the inclusion of the duration cue. For older adults, instruction improved perception of the familiar contrast, regardless of cue information, whereas instruction improved perception of the unfamiliar contrast only when the duration cue was provided. Altogether, these findings shed new light on the conditions under which explicit instruction can orient L2 listeners' attention and improve their speech perception, revealing several important interactions between the specific L2 contrasts in question, the content of the instruction, and the listener group.

Funding

This research was supported by a Vidi grant from the Dutch Research Council (NWO), awarded to the fourth author.

References

doi Anthony, J. L. & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259.

doi Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, *36*(1), 1–6.

doi Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Best, C.T., & Tyler, M.D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. Munro. (Eds.), *Language experience in second-language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Bieber, R.E., & Gordon-Salant, S. (2021). Improving older adults' understanding of challenging speech: Auditory training, rapid adaptation and perceptual learning. *Hearing Research*, *402*, 108054.

Broersma, M. (2010). Perception of final fricative voicing: native and nonnative listeners' use of vowel duration. *The Journal of the Acoustical Society of America*, *127*(3), 1636–1644.

Broersma, M. (2012). Increased lexical activation and reduced competition in second-language listening. *Language and Cognitive Processes*, *27*(7–8), 1205–1224.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System*, *36*(1), 22–34.

Brown, A. (1995). Minimal pairs: Minimal importance? *ELT Journal*, *49*(2), 169–175.

Booij, G. (1999). *The phonology of Dutch*. Oxford University Press.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.

Chandrasekaran, B., Yi, H., Smayda, K.E., & Maddox, W.T. (2016). Effect of explicit dimensional instruction on speech category learning. *Attention, Perception, & Psychophysics*, *78*, 566–582.

Chen, Y., & Pederson, E. (2017). Directing attention during perceptual training: A preliminary study of phonetic learning in Southern Min by Mandarin speakers. In *Proceedings of Interspeech 2017* (pp. 1770–1774).

Collins, B., & Mees, I.M. (1996). *The phonetics of English and Dutch* (3rd ed.). E. J. Brill.

Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, *22*(6), 680–689.

Eisner, F. & McQueen, J.M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, *67*(2), 224–238.

Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.

Flege, J.E. & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland. (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.

Francis, A.L., & Nusbaum, H.C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 349–366.

Guion, S.G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.S. Bohn & M. Munro. (Eds.), *Language experience in second language speech learning* (pp. 57–77). John Benjamins.

Gordon-Salant, S. (2005). Hearing loss and aging: New research findings and clinical implications. *Journal of Rehabilitation Research & Development*, *42*(4), 9–24.

Hisagi, M., & Strange, W. (2011). Perception of Japanese temporally-cued contrasts by American English listeners. *Language and Speech*, *54*(2), 241–264.

House, A.S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America*, *33*(9), 1174–1178.

Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*(5), 3099–3111.

Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2015). Lexical specificity training effects in second language learners. *Language Learning*, *65*(2), 358–389.

Kissling, E. M. (2014). Phonetics instruction improves learners' perception of L2 sounds. *Language Teaching Research*, *19*(3), 254–275.

Krenca, K., Segers, E., Chen, X., Shakory, S., Steele, J., & Verhoeven, L. (2020). Phonological specificity relates to phonological awareness and reading ability in English-French bilingual children. *Reading and Writing*, *33*, 267–291.

Kubo, R., & Akahane-Yamada, R. (2006). Influence of aging on perceptual learning of English phonetic contrasts by native speakers of Japanese. *Acoustical Science and Technology*, *27*(1), 59–61.

Maddox, W. T., Chandrasekaran, B., Smayda, K., & Yi, H. (2013). Dual systems of speech category learning across the lifespan. *Psychology and Aging*, *28*(4), 1042–1056.

Pederson, E., & Guion-Anderson, S. (2010). Orienting attention during phonetic training facilitates learning. *The Journal of the Acoustical Society of America*, *127*(2), EL54–59.

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1315–1330.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology*, *42*(Suppl2), 2S11–2S16.

Porretta, V. J., & Tucker, B. V. (2014). Perception of non-native consonant length contrast: The role of attention in phonetic processing. *Second Language Research*, *31*(2), 239–265.

Raphael, L. J. (2005). Acoustic cues to the perception of segmental phonemes. In D. B. Pisoni & R. E. Remez. (Eds.), *The handbook of speech perception* (pp. 182–206). Blackwell Publishing.

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, *39*, 187–224.

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157–166.

Scharenbourg, O., & Janse, E. (2013). Comparing lexically guided perceptual learning in younger and older listeners. *Attention, Perception, & Psychophysics*, *75*, 525–536.

Slis, I. H., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction I. *Language and Speech*, *12*(2), 80–102.

Sommers, M. S. (1997). Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment. *The Journal of the Acoustical Society of America*, *101*(4), 2278–2288.

Svalberg, A. M. L. (2007). Language awareness and language learning. *Language Teaching*, *40*(4), 287–308.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183–203.

van der Feest, S. V. H., & Swingley, D. (2011). Dutch and English listeners' interpretation of vowel duration. *The Journal of the Acoustical Society of America*, *129*(3), EL57–EL63.

van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*, 1000.

Appendix A

Table A1. Phonological awareness stimuli

Critical /t/-/d/ minimal pairs		**Critical /æ/-/ɛ/ minimal pairs**		**Filler minimal pairs**		**Filler homophone pairs**	
beat	bead*	and	end	better	bitter	air	heir
bright	bride	bad*	bed	bike	bake	allowed	aloud
built	build	bag	beg	boat	both	bare	bear
cart	card	bat	bet	came	game	blew	blue
fate	fade	cattle	kettle	chase	chess	find	fined
feet	feed	dad	dead	cloud	crowd	flew	flu
float	flowed	expanse	expense	desk	disc	flour	flower
got	god	flash	flesh	file	fail	hair	hare
great	grade	gas	guess	forgot	forget	higher	hire
greet	greed	had*	head	fork	fort	him	hymn
height	hide	lag	leg	fry	fly	hour	our
hurt	heard	land	lend	glue	clue	knight	night
right	ride	man	men	left	lift	knot	not
seat	seed	mansion	mention	lesson	listen	knows	nose
wrote	road	mantle	mental	like	look	made	maid
sight	side	radish	reddish	loose	less	mind	mined
slight	slide	sad	said	medal	middle	none	nun
spent	spend	sand	send	note	net	peace	piece
threat	thread	than	then	path	bath	rays	raise
white	wide	track	trek	pile	pale	sail	sale
				play	pray	seas	sees
				pride	proud	sole	soul
				rest	wrist	some	sum
				rice	race	son	sun
				run	pun	tale	tail
				save	shave	there	their
				taste	test	waist	waste
				trade	train	wait	weight
				true	through	way	weigh
				warn	warm	wood	would

* These words could sound similar to Dutch words if /d/ is interpreted as /t/ (*bead* resembles Dutch *biet* /bit/) or if, in addition to that, /æ/ is interpreted as /ɛ/ (*bad* resembles Dutch *bed* /bɛt/ and *had* resembles Dutch *het* /hɛt/).

Appendix B

The scripts for each of the four phonetic instruction are provided below. The text that differs between the duration-cue and no-cue versions of the /æ/-/ɛ/ and /t/-/d/ videos is underlined.

/æ/-/ɛ/ Video with duration cue

Hi! I'm Emily, and I'm a native speaker of English. In this video, I'm going to teach you about the difference between two sounds in English: the /æ/ sound and the /ɛ/ sound. They may sound similar, but these two sounds make an important distinction in English. For example, the difference between /æ/ and /ɛ/ distinguishes words like *pan* and *pen*, and *jam* and *gem*. Do you think it's hard to hear? The /æ/ sound is usually spelled with the letter A as in *map*, while the /ɛ/ sound is usually spelled with the letter E as in *desk*.

The sounds /æ/ and /ɛ/ differ in the color, or quality, of their sound, but that's very subtle. What really helps to hear the difference is paying attention to how long the sound is.

Listen closely to these examples, in which I exaggerate the difference: *Pen. Paaan. Gem. Jaaam.* Now I'm going to pronounce the words more normally. If you listen carefully, you'll hear that the /æ/ sound is longer than the /ɛ/ sound. Try to hear the difference between *pan, pen, pan, pen, pan, pen*. Can you hear the difference in another word pair? Listen to *jam, gem, jam, gem, jam, gem*.

In short, just remember: the /æ/ sounds longer while the /ɛ/ sounds shorter. I hope that helps you!

/æ/-/ɛ/ Video with no cue

Hi! I'm Emily, and I'm a native speaker of English. In this video, I'm going to teach you about the difference between two sounds in English: the /æ/ sound and the /ɛ/ sound. They may sound similar, but these two sounds make an important distinction in English. For example, the difference between /æ/ and /ɛ/ distinguishes words like "pan" and "pen", and "jam" and "gem." Do you think it's hard to hear? The /æ/ sound is usually spelled with the letter A as in "map", while the /ɛ/ sound is usually spelled with the letter E as in "desk."

The sounds /æ/ and /ɛ/ differ in the color, or quality, of their sound, but that's very subtle. Native speakers can hear the difference between the /æ/ sound and the /ɛ/ sound very easily, but for people who speak English as a second language, it can be difficult.

Listen closely to these examples, in which I exaggerate the difference: *Pen. Pan. Gem. Jam.* Now I'm going to pronounce the words more normally. Try to hear the difference between *pan, pen, pan, pen, pan, pen*. Can you hear the difference in another word pair? Listen to *jam, gem, jam, gem, jam, gem*.

It will become easier to hear the difference between /æ/ and /ɛ/ the more you practice listening. I hope that helps you!

/t/-/d/ Video with duration cue

Hi! I'm Emily, and I'm a native speaker of English. In this video, I'm going to teach you about the difference between two sounds in English: the /t/ sound and /d/ sound at the end of a word. They may sound similar, but these two sounds make an important distinction at the end of a word in English. For example, the difference between /t/ and /d/ distinguishes words like *not* and *nod*, and *bit* and *bid*. Do you think it's hard to hear? If a word ends with T or T-E, the sound is always /t/ as in *sit*. If a word ends with D or D-E, the sound is nearly always /d/ as in *did*.

The /t/ sound comes with a little puff of air, while the /d/ sound does not, but that's very subtle. What really helps to hear the difference is paying attention to how long the vowel before it sounds.

Listen closely to these examples, in which I exaggerate the difference: *Not. Noood. Bit. Bii-iiid*. Now I'm going to pronounce the words more normally. If you listen carefully, you'll hear that the vowel before the /d/ sound is longer than the vowel before the /t/ sound. Try to hear the difference between *bit, bid, bit, bid, bit, bid*. Can you hear the difference in another word pair? Listen to *not, nod, not, nod, not, nod*.

In short, just remember: if the vowel is longer, you're usually hearing a /d/; if the vowel is shorter, you're usually hearing a /t/. I hope that helps you!

/t/-/d/ Video with no cue

Hi! I'm Emily, and I'm a native speaker of English. In this video, I'm going to teach you about the difference between two sounds in English: the /t/ sound and /d/ sound at the end of a word. They may sound similar, but these two sounds make an important distinction at the end of a word in English. For example, the difference between /t/ and /d/ distinguishes words like *not* and *nod*, and *bit* and *bid*. Do you think it's hard to hear? If a word ends with T or T-E, the sound is always /t/ as in *sit*. If a word ends with D or D-E, the sound is nearly always /d/ as in *did*.

The /t/ sound comes with a little puff of air, while the /d/ sound does not, but that's very subtle. Native speakers can hear the difference between the /t/ sound and the /d/ sound at the end of a word very easily, but for people who speak English as a second language, it can be difficult.

Listen closely to these examples, in which I exaggerate the difference: *Not. Nod. Bit. Bid*. Now I'm going to pronounce the words more normally. Try to hear the difference between *bit, bid, bit, bid, bit, bid*. Can you hear the difference in another word pair? Listen to *not, nod, not, nod, not, nod*.

It will become easier to hear the difference between /t/ and /d/ at the end of a word, the more you practice listening. I hope that helps you!

Appendix C

Table C1. Mean participant-level change in phonological awareness accuracy score (%) from pre-test to post-test for different combinations of age group, item contrast, video contrast, and duration cue information

Younger adults						
	/t/-/d/ Items			/æ/-/ɛ/ Items		
Video	**Duration cue**	**No cue**	**Overall**	**Duration cue**	**No cue**	**Overall**
/t/-/d/	+28.4	+36.3	+32.3	+24.7	+8.1	+16.4
/æ/-/ɛ/	+23.8	+9.4	+16.6	+35.6	+16.6	+26.1
Overall	+26.1	+22.8	+24.5	+30.2	+12.3	+21.3
Older adults						
	/t/-/d/ Items			/æ/-/ɛ/ Items		
Video	**Duration cue**	**No cue**	**Overall**	**Duration cue**	**No cue**	**Overall**
/t/-/d/	+20.0	+37.9	+28.6	+2.7	−1.4	+0.7
/æ/-/ɛ/	+20.0	+20.0	+20.0	+47.2	+30.3	+39.0
Overall	+20.0	+28.6	+24.2	+25.6	+15.0	+20.5

Table C2. ANOVA for awareness data (Full Model)

	χ^2	*p*-value
Item Contrast	21.84	<.001*
Test Time	614.03	<.001*
Video Contrast	0.01	.91
Cue Information	0.02	.88
Age Group	0.08	.78
Test Time • Item Contrast	2.70	.10
Video Contrast • Item Contrast	150.22	<.001*
Test Time • Video Contrast	1.49	.22
Item Contrast • Cue Information	3.24	.07
Test Time • Cue Information	14.27	<.001*
Video Contrast • Cue Information	0.07	.79
Item Contrast • Age Group	17.06	<.001*
Test Time • Age Group	0.05	.83
Video Contrast • Age Group	0.04	.83
Cue Information • Age Group	4.05	.04*
Item Contrast • Test Time • Video Contrast	118.12	<.001*
Item Contrast • Test Time • Cue Information	21.94	<.001*

	χ^2	p-value
Item Contrast • Video Contrast • Cue Information	5.61	.02*
Test Time • Video Contrast • Cue Information	7.10	.01*
Item Contrast • Test Time • Age Group	0.00	.97
Item Contrast • Video Contrast • Age Group	1.09	.30
Test Time • Video Contrast • Age Group	33.63	<.001*
Item Contrast • Cue Information • Age Group	9.53	.002*
Test Time • Cue Information • Age Group	9.10	.003*
Video Contrast • Cue Information • Age Group	0.22	.64
Item Contrast • Test Time • Video Contrast • Cue Information	3.09	.08
Item Contrast • Test Time • Video Contrast • Age Group	11.11	.001*
Item Contrast • Test Time • Cue Information • Age Group	3.21	.07
Item Contrast • Video Contrast • Cue Information • Age Group	4.08	.04*
Test Time • Video Contrast • Cue Information • Age Group	5.65	.02*
Item Contrast • Test Time • Video Contrast • Cue Information • Age Group	0.26	.61

* significant.

Table C3. Separate models predicting younger and older adults' awareness accuracy

	Younger adults		Older adults	
	χ^2	p-value	χ^2	p-value
Test Time	334.13	<.001*	218.39	<.001*
Video Contrast	0.01	.94	0.07	.78
Cue Information	2.01	.16	1.93	.16
Item Contrast	11.09	.001*	29.36	<.001*
Test Time • Video Contrast	9.49	.002*	24.16	<.001*
Test Time • Cue Information	24.84	<.001*	0.09	.77
Test Time • Item Contrast	1.77	.18	1.00	.31
Video Contrast • Cue Information	0.01	.90	0.31	.58
Video Contrast • Item Contrast	93.70	<.001*	58.79	<.001*
Cue Information • Item Contrast	11.40	.001*	1.01	.31
Test Time • Video Contrast • Cue Information	0.05	.83	12.90	<.001*
Test Time • Video Contrast • Item Contrast	31.89	<.001*	97.78	<.001*
Test Time • Cue Information • Item Contrast	4.56	.03*	20.59	<.001*
Video Contrast • Cue Information • Item Contrast	9.52	.002*	0.03	.87
Test Time • Video Contrast • Cue Information • Item Contrast	2.68	.10	0.69	.41

* significant.

Appendix D

Table D1. Mean participant-level change in perception accuracy score (%) from pre-test to post-test for different combinations of age group, item contrast, video contrast, and duration cue information

	Younger adults					
	/t/-/d/ Items			/æ/-/ɛ/ Items		
Video	**Duration cue**	**No cue**	**Overall**	**Duration cue**	**No cue**	**Overall**
/t/-/d/	+6.3	+5.9	+6.1	+0.5	+0.2	+0.3
/æ/-/ɛ/	+3.0	+1.3	+2.1	+4.2	+7.2	+5.7
Overall	+4.6	+3.6	+4.1	+2.3	+3.7	+3.0
	Older adults					
	/t/-/d/ Items			**/æ/-/ɛ/ Items**		
Video	**Duration cue**	**No cue**	**Overall**	**Duration cue**	**No cue**	**Overall**
/t/-/d/	+8.8	+3.8	+6.4	−4.2	+1.8	−1.3
/æ/-/ɛ/	−1.4	−1.8	−1.6	+11.9	+4.8	+8.5
Overall	+3.5	+0.9	+2.3	+4.1	+3.4	+3.8

Table D2. ANOVA for perception data (Full Model)

	χ^2	*p*-value
Item Contrast	30.38	<.001*
Test Time	38.52	<.001*
Video Contrast	0.01	.94
Cue Information	0.52	.47
Age Group	40.90	<.001*
Test Time • Item Contrast	1.35	.24
Video Contrast • Item Contrast	14.83	<.001*
Test Time • Video Contrast	0.57	.45
Item Contrast • Cue Information	3.37	.07
Test Time • Cue Information	0.76	.38
Video Contrast • Cue Information	0.08	.78
Item Contrast • Age Group	0.01	.93
Test Time • Age Group	3.62	.06
Video Contrast • Age Group	0.66	.42
Cue Information • Age Group	3.68	.06
Item Contrast • Test Time • Video Contrast	43.33	<.001*
Item Contrast • Test Time • Cue Information	1.28	.26

	χ^2	*p*-value
Item Contrast • Video Contrast • Cue Information	0.17	.68
Test Time • Video Contrast • Cue Information	0.61	.43
Item Contrast • Test Time • Age Group	2.96	.09
Item Contrast • Video Contrast • Age Group	0.72	.40
Test Time • Video Contrast • Age Group	0.00	.96
Item Contrast • Cue Information • Age Group	2.59	.11
Test Time • Cue Information • Age Group	0.13	.72
Video Contrast • Cue Information • Age Group	0.29	.59
Item Contrast • Test Time • Video Contrast • Cue Information	2.09	.15
Item Contrast • Test Time • Video Contrast • Age Group	0.34	.56
Item Contrast • Test Time • Cue Information • Age Group	0.06	.81
Item Contrast • Video Contrast • Cue Information • Age Group	0.53	.47
Test Time • Video Contrast • Cue Information • Age Group	1.18	.28
Item Contrast • Test Time • Video Contrast • Cue Information • Age Group	4.21	.04*

* significant

Table D3. Separate models predicting younger and older adults' perception accuracy

	Younger adults		Older adults	
	χ^2	*p*-value	χ^2	*p*-value
Test Time	29.57	<.001*	12.94	<.001*
Item Contrast	25.64	<.001*	26.58	<.001*
Cue Information	3.37	.07	0.75	.39
Video Contrast	0.44	.51	0.34	.56
Test Time • Video Contrast	0.47	.50	0.16	.69
Test Time • Cue Information	0.05	.82	0.91	.34
Test Time • Item Contrast	5.22	.02*	0.14	.70
Video Contrast • Cue Information	0.33	.56	0.04	.85
Video Contrast • Item Contrast	2.74	.10	13.37	<.001*
Cue Information • Item Contrast	6.17	.01*	0.09	.77
Test Time • Video Contrast • Cue Information	0.08	.77	1.60	.21
Test Time • Video Contrast • Item Contrast	13.93	<.001*	30.29	<.001*
Test Time • Cue Information • Item Contrast	0.95	.33	0.43	.51
Video Contrast • Cue Information • Item Contrast	0.79	.37	0.03	.85
Test Time • Video Contrast • Cue Information • Item Contrast	0.44	.51	5.93	.01*

* significant.

Address for correspondence

Emily Felker
Centre for Language Studies
Radboud University
P.O. Box 9103
6500 HD Nijmegen
The Netherlands

emily.felker@ru.nl
https://orcid.org/0000-0002-6184-5563

Co-author information

Esther Janse
Centre for Language Studies
Radboud University

esther.janse@ru.nl

Mirjam Ernestus
Centre for Language Studies
Radboud University

mirjam.ernestus@ru.nl

Mirjam Broersma
Centre for Language Studies
Radboud University

mirjam.broersma@ru.nl

Publication history

Date received: 17 November 2020
Date accepted: 22 September 2021
Published online: 2 November 2021