American Psychological Association

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/19/$12.00

2019, Vol. 45, No. 7, 1252–1270
http://dx.doi.org/10.1037/xlm0000647

Is There Adaptation of Speech Production After Speech Perception in Bilingual Interaction?

Wouter P. J. Broos, Aster Dijkgraaf, Eva Van Assche, Heleen Vander Beken, Nicolas Dirix, Evelyne Lagrou, Robert J. Hartsuiker, and Wouter Duyck
Ghent University

In dialogue, speakers tend to adapt their speech to the speech of their interlocutor. Adapting speech production to preceding speech input may be particularly relevant for second language (L2) speakers interacting with native (L1) speakers, as adaptation may facilitate L2 learning. Here we asked whether Dutch-English bilinguals adapt pronunciation of the English phonemes /æ/ and coda /b/ when reading aloud sentences after exposure to native English speech. Additionally, we tested whether social context (presence or absence of a native English confederate) and time lag between perception and production of the phoneme affected adaptation. Participants produced more English-like target words that ended in word-final /b/ after exposure to target phonemes produced by a native speaker, but the participants did not change their production of the phoneme /æ/ after exposure to native /æ/. The native English speaking confederate did not show consistent changes in speech production after exposure to target phonemes produced by L2 speakers. These findings are in line with Gambi and Pickering's simulation theory of phonetic imitation (Gambi & Pickering, 2013).

Keywords: speech alignment, phonetic adaptation, second language production, bilingualism

Speech production is highly variable. This variability is caused by between-speaker differences such as the mother tongue, age, gender, dialect, and articulatory properties of a particular speaker. In addition, within-speaker differences manifest themselves through peripheral factors such as the time of day, mood, or even just having a cold. Therefore, articulation of words or even phonemes varies considerably. As a consequence, listeners must find a way to cope with this variation. The fact that listeners mostly do not experience difficulty understanding (variable) speech suggests that they can do this very efficiently. Indeed, studies on speech perception have shown that listeners can quickly adjust their perceptual system, for instance to deal with an unusual way in which a speaker realizes a particular phoneme (e.g., Norris, McQueen, & Cutler, 2003). Such adjustment may be particularly useful in a second language (L2), given that the realization of phonemes varies across languages and that such adjustments may help L2 learning (Costa, Pickering, & Sorace, 2008), especially when interacting with native speakers who master the language better. The goal of the present study is to test whether nonnative listeners (of English) are not only sensitive to differences between their own L2 phoneme production and native production, but also whether these differences affect their L2 speech production (in other words, whether there is *alignment* between L2 speech production and perception). To gauge whether any such adaptation is automatic or strategic, we considered the effects of several further variables. First, we tested whether the physical presence of a native speaker has an additional effect on speech alignment, because previous studies suggested that social context modulates alignment (e.g., Babel, 2012). Second, we manipulated the lag (number of intervening trials) between perception and production of the critical phoneme.

Phonetic Alignment in L1 Speech Production

Previous studies have shown that L1 listeners can adjust their perception to speech that is produced by their interlocutor, including accents and other nonnative speech characteristics (Bradlow & Bent, 2008; Eisner & McQueen, 2006; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Norris et al., 2003). Norris et al. (2003) for instance, demonstrated this by using a paradigm in which participants were exposed to an ambiguous fricative [?], midway between [f] and [s]. When listeners were exposed to ambiguous [f]-final words, they categorized later ambiguous [?] more often as an [f], whereas when listeners were exposed to ambiguous [s]-final words, they categorized the ambiguous [?] more often as an [s]. Therefore, listeners can perform perceptual adaptation by using their lexical knowledge to adjust their phonemic representations, making them consistent with specific speech

This article was published Online First September 27, 2018.

Wouter P. J. Broos, Aster Dijkgraaf, Eva Van Assche, Heleen Vander Beken, Nicolas Dirix, Evelyne Lagrou, Robert J. Hartsuiker, and Wouter Duyck, Department of Experimental Psychology, Ghent University.

This article received funding from the special research fund of Ghent University (GOA - Concerted Research Action BOF13/GOA/032).

The data sets and scripts used for the analyses in this article are available online at Open Science Framework (https://osf.io/p62j4/). Wouter P. J. Broos and Aster Dijkgraaf made an equal contribution to this article.

Correspondence concerning this article should be addressed to Wouter P. J. Broos, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. E-mail: wouter.broos@ugent.be

variants. This effect also occurs when listening in L2 (Weber, Betta, & McQueen, 2014).

There is also evidence suggesting that speakers adapt speech production to speech of an interlocutor. Alignment of speech production occurs at the syntactic (e.g., Bernolet, Hartsuiker, & Pickering, 2012, 2013; Pickering & Branigan, 1999), lexical (e.g., Branigan, Pickering, Pearson, McLean, & Brown, 2011), and phonetic (e.g., Babel, 2012; Lametti, Krol, Shiller, & Ostry, 2014; Pardo, 2006) levels. The Interactive Alignment Model (Garrod & Pickering, 2004) accounts for such effects in speech production and assumes that speech alignment occurs because for communication to be successful, mental states of interlocutors should become aligned. If mental states are aligned, interlocutors come to understand the ideas under discussion in the same way. According to the interactive alignment account, alignment percolates between different levels (e.g., phonological, lexical, and syntactic levels) so that alignment on one level stimulates alignment on other levels in both perception and production. Alignment is assumed to be an automatic process in the sense that it is effortless and speakers are unaware of the process. Garrod and Pickering (2004) suggested that alignment comes about through priming of representations between speakers and listeners. In a more recent account of (phonetic) adaptation, Gambi and Pickering (2013) suggested that adaptation occurs because listeners simulate speakers' utterances by constructing forward model predictions of the speakers' utterances using their own speech production system (Pickering & Garrod, 2013). Adaptation to an interlocutor occurs because the listener's predictions mismatch the speaker's utterance and the listener will try to correct the prediction error in perception. Both Pickering and Garrod's interactive alignment model and Gambi and Pickering's simulation theory assume parity between perception and production. Therefore, an adaptation as a consequence of a prediction error in speech perception can lead to adaptations in speech production as well.

Social factors influence the occurrence of phonetic alignment. Babel (2012), for instance, focused on several social variables. Participants first produced a list of target words in a baseline block after which they performed a shadowing task where they repeated words that were presented auditorily over headphones. During the shadowing task, participants either saw a picture of the speaker on the screen or no picture at all. There was more alignment in the social condition (with a picture of the speaker on the screen) than in the auditory exposure only condition. Liking the model speaker (as measured with ratings) also increased alignment. These findings support the view that alignment can be socially driven. However, alignment did not occur to the same extent for each vowel type: There seemed to be more alignment when there was more acoustic space available for alignment. According to Gambi and Pickering (2013), social factors and context factors may influence alignment by affecting how much a listener relies on forward-models of the speaker.

A further important social variable affecting alignment may be the perceived social distance between the interlocutors. One reason for such social distance effects is that comprehension may occur through either a prediction-by-simulation route (simulating interlocutors' speech using one's own production system), or a prediction-by-association route (predicting interlocutors' speech using perceptual experience; see Pickering & Garrod, 2013 for a detailed discussion). Gambi and Pickering (2013) suggest that in some contexts—for example, when an interlocutor is perceived as very different from the listener—listeners may be more inclined to rely on the prediction-by-association route. As this route does not rely on the listener's production system, subsequent speech production is not affected by the predictions made about the interlocutor's speech. This may explain why adjustments in phoneme perception do not always lead to changes in production. For instance, Kraljic, Brennan, and Samuel (2008) exposed half of their participants to speech where /s/ was replaced with the pronunciation ~s∫ (ambiguous between /s/ and /∫/) when immediately followed by the [tr] (such as in known English dialects). The other participants were exposed to speech in which all instances of /s/ were replaced by ~s∫ (idiolectal condition). There was perceptual learning for the idiolectal variation, but not for the dialectal variation. More important, the changes found in perception did not affect subsequent production.

Phonetic Alignment in L2 Speech Production

According to Gambi and Pickering (2013), speech alignment occurs to a larger extent when interlocutors are more similar to each other or when they perceive each other as being more similar. Thus, alignment may fail when interlocutors are highly dissimilar, for example when a nonnative speaker is engaged in conversation with a native speaker. Nonnative speakers may also lack the flexibility and automaticity in speech production necessary for alignment (Costa et al., 2008), because they may have more limited or erroneous knowledge of L2 linguistic representations and because language perception and production are more effortful in L2.

In line with simulation theory (Gambi & Pickering, 2013), Kim, Horton, and Bradlow (2011) show that closer interlocutor language distance facilitates phonetic alignment. The authors studied alignment in interlocutor pairs with different dialects or with a different L1 with an AXB perceptual similarity test. In this similarity test, an independent group of listeners heard three repetitions of the same target word. The first and last production of the target word represented pronunciation of the target word in the pre- and postexposure phase (A and B). The second production of the target word (X) was produced by the first speakers' interlocutor. The listeners who judged pronunciation of the target word were asked to decide whether A or B sounded more like X. Therefore, the judgment of the listeners was used as a subjective measure of alignment. Phonetic alignment only occurred when two speakers with the same L1 and dialect were engaged in dialogue and not when the dialects differed or when one conversation partner had a different L1.

Kim et al.'s (2011) finding that alignment was strongest for interlocutor pairs that shared L1 and dialect differs from findings by Hwang, Brennan, and Huffman (2015). These authors studied phonetic alignment in nonnative dialogue and asked whether the amount of alignment depended on social affiliation and on the necessity of phoneme disambiguation in dialogue. Unbalanced Korean-English bilinguals interacted with a Korean English-speaking confederate and a monolingual American English-speaking confederate in English. Participants were asked to explain to the confederate how to rearrange a board with words so that it would match that of the participant. Acoustic measures were used to quantify alignment (formant frequencies, closure voicing

duration, and vowel duration). Participants produced more English-like phonemes when being immediately primed by a monolingual American confederate pronouncing that same phoneme and their pronunciation did not change when they were speaking to a Korean confederate. Simulation theory can still account for this finding if we assume that the bilingual participants perceived themselves as more similar to the native English confederate than to the Korean confederate. A second experiment showed that participants also produced more English-like phonemes when they needed to distinguish between two potentially ambiguous words on the board.

As in L1, social factors seem to have an influence on the amount of phonetic alignment in L2 speakers. Trofimovich and Kennedy (2014) focused on the nature and the amount of interactive alignment in L2-L2 dialogue. A pair of L2 speakers of English with different L1 backgrounds performed an information exchange task in which interlocutors were required to transmit information unknown to one of the two interlocutors to reach a common goal. In line with Kim et al. (2011), alignment was stronger when interlocutors' speech characteristics (fluency, language complexity) were initially more similar. Greater alignment also occurred when interlocutors' affective/personal qualities were initially more similar. This suggests that speakers are perceptive to social context so that similar personality traits lead to an increase in speech alignment (see below).

Kim (2012) observed phonetic alignment of an L1 speaker toward an L2 speaker. In contrast to Kim et al. (2011), who only found alignment in L1-L1 dialogue where speakers shared the same dialect, alignment occurred irrespective of whether the participant shared L1 or dialect with the other speaker. Interestingly, Kim (2012) found that phonetic alignment was larger for larger initial acoustic distances between the two speakers.

Present Study

Most previous studies on phonetic alignment in L2 speakers used subjective measures to test whether interlocutors sounded more alike after an interaction. Here, we will use objective acoustic measures to test whether L2 speakers adjust their speech production of specific phonemes, after being exposed to those phonemes in a sentence context produced by a native confederate. Pickering and Garrod (2013) argue that alignment is a rather automatic process, driven by priming. Hence, an L2 speaker may not only adapt their speech to an L1 speaker, but also vice versa (Kim, 2012). Therefore, we will also test whether a native English speaking confederate aligns her own speech to that of an L2 speaker.

Specifically, we will investigate whether L2 speakers of English adjust their nonnative realization of the English phonemes /æ/ and word-final /b/ toward a more native realization after exposure to native realizations of the phoneme. We use word-final /b/ in this study because Dutch nonnative speakers of English often replace the English phoneme /b/ with the Dutch phoneme /p/ when it is positioned at the end of a word (Collins & Mees, 1996). This phenomenon exists because Dutch has final devoicing: All voiced consonants in final position are realized as voiceless (Giegerich, 1992). For instance, the English word "mob" /mɒb/ is often mispronounced as /mɒp/. The distinction between the voiced consonants/b *d* g/and voiceless consonants/p t k/in syllable-final position in English is made mainly by vowel length; vowels that precede a word-final voiced consonant are longer than vowels that precede a word-final voiceless consonant (Luce & Charles-Luce, 1985; Raphael, 1972). If alignment occurs, the duration of vowels preceding /b/ should increase with increasing amounts of exposure to native speech. Additionally, closure duration tends to be shorter for voiced word-final stops and longer for voiceless word-final stops (Lisker, 1957; Luce & Charles-Luce, 1985).[1] Therefore, we expect closure duration of /b/ to become shorter when participants are exposed to native speech. Yet, if the confederate aligns with the participant, her vowels preceding /b/ will be shorter whereas closure duration is expected to be longer.

The vowel /æ/ does not exist in Dutch and is often substituted by /ɛ/ by Dutch speakers (Collins & Mees, 1996). To study adaptation in the realization of vowel /æ/, we determine both the first spectral peak (F1) and second spectral peak (F2) as well as the duration of /æ/. F1 correlates with the height of the tongue (vertical tongue position); if the tongue is low (as in/a:/), F1 is high and if the tongue is positioned high (as in/i:/), F1 is low. F2 correlates with the tongue being placed at the front or back of the mouth (horizontal tongue position). In the former placement, F2 is high; in the latter, F2 is low. It is hypothesized that a difference in F1 and F2 before and after exposure should be seen if speech alignment occurs. F1 of /æ/ is slightly higher (lower tongue/jaw position) than F1 of /ɛ/, and F2 of /æ/ is slightly lower (tongue position more back) than F2 of /ɛ/ (tongue position more back). Therefore, if alignment takes place, we expect nonnative speakers to adjust their F1 upward and their F2 downward when attempting to pronounce the English vowel /æ/. The opposite is expected for the confederate. Also, /æ/ is longer than /ɛ/ (e.g., Bohn & Flege, 1990; Collins & Mees, 1996) and we, therefore, expect participant to lengthen the vowel if they align with the confederate. However, we expect the confederate to shorten the vowel if she aligns with the participants.

We will also test whether the amount of alignment depends on social context, contrasting a confederate who is present during the experiment with exposure to speech over headphones. The physical presence of the confederate is expected to increase the extent to which participants feel engaged in dialogue, thereby stimulating alignment. Furthermore, we will test whether the amount of phonetic alignment depends on the time lag between perception and production. We expect alignment to be stronger when the time lag between perception and production is short (zero intervening sentences). This would be in line with accounts assuming parity between production and comprehension (Gambi & Pickering, 2013; Pickering & Garrod, 2013) and it would verify Hwang et al.'s (2015) finding that there is alignment when production of a target immediately follows perception.

[1] The duration of voicing in the closure phase of /b/ is often measured to determine voicing of /b/ (e.g., Hwang et al., 2015). However, voicing duration could not be measured reliably because of considerable noise in the recording. Please note that vowel length is the most reliable cue in distinguishing voiced and voiceless final stops (Luce & Charles-Luce, 1985), but for sake of systematicity we also measured closure duration of word-final /b/.

Method

Participants

Thirty-two female students from Ghent University (age M = 25.38, SD = 8.17, range: 19 to 57) participated in the experiment in exchange for monetary compensation. They were divided into two groups of 16 (the confederate-absent and confederate-present groups, see below) by random assignment. Because men and women differ in formant frequencies and our confederate was female, we decided to test only female participants. They were all late Dutch-English bilinguals who started learning English around the age of 12 at secondary school for approximately 3 to 4 h a week. In addition to this classroom exposure, students in Belgium are regularly exposed to English through TV, books, video/computer games, and other kinds of media. All participants were born and raised in Flanders. Proficiency in L1 and L2 was measured using the LexTALE test of vocabulary knowledge for advanced learners of English (Lemhöfer & Broersma, 2012) and a self-report questionnaire. In this questionnaire, participants rated their L1 and L2 proficiency in reading, writing, speaking, and listening on a 5-point scale ranging from 1 (*not at all*) to 5 (*perfect/mother tongue*; see Table 1 for participant characteristics). They also provided more background information on their (previous) place of residence. Besides Dutch and English, all participants also spoke French (mean rating = 3.28 on a scale from 1 to 5 where 1 (*not at all*) to 5 (*perfect/mother tongue*). Participants all reported not to have dyslexia or hearing deficiencies and eyesight was normal or corrected-to-normal.

Confederate

The confederate was female and she originated from the Pacific Northwest of the United States of America. She was 30 years old at the time of testing and had been living in Belgium for little over a year. English was her native language but she also spoke French and Dutch. The confederate also performed the LexTALE in Dutch and in English. Her score for Dutch was 67.5 and her score for English was 96.25.

Design

The experiment consisted of three blocks: a baseline block, an exposure block, and an alternating block. In the baseline block, 30 sentences, each with two target words (one for /æ/ and one for /b/) were presented to the participant to read out loud. In the exposure block 30 different sentences with the same 60 target words were read out loud by the confederate. In the postexposure (alternating) block, the participant and the confederate alternated in reading 120 sentences out loud that each contained one of the 60 target words. Over the course of the experiment, each target word occurred four times (produced twice by the confederate and twice by the participant) but it was presented in a different sentence each time.

In the alternating block, the lag between the sentence containing a target word that was produced by the confederate and the test sentence containing that same target word produced by the participant could be either zero or four. A lag of zero means that the critical sentence for the participant was presented immediately after the confederate produced a sentence containing the same target word. A lag of four indicates that four intervening sentences were presented between the critical sentences of the participant and confederate. Lag was a within-participant variable. To enable the lag manipulation, 30 fillers were added to the 120 sentences in the alternating block. These filler sentences had a similar structure and length as the critical sentences but they did not contain the target words or the specific contrast. Half of the fillers were read by the participant and half of the fillers were read by the confederate. Each phoneme was presented 15 times at lag zero and 15 times at lag four in the alternating block.

There was a condition in which the confederate was present in the same room as the participant during the experiment, and a condition in which the confederate was not present in the same room but read out loud sentences in a microphone (Røde USB 1000A) in another room (see Procedure for details). This social context (confederate present or absent) was manipulated between participants. Table 2 below summarizes the design of the experiment.

Materials

There were two target phonemes: word—final /b/ and the vowel /æ/ (see Appendix A for the full stimuli list). We selected 30 English target words for each of the two phonemes. English /æ/ (as in "map" and "trap") is affected by dark [†], giving a retracted [ä] such as in pal, shall. The mouth is not as open when pronouncing English /æ/ before velar phonemes/ŋ, k, g, ʍ, w/giving rise to [æ] (e.g., back, bag, bang; Collins & Mees, 1996). Therefore, the vowel was never followed by one of these sounds in a target word. In addition, /æ/ was never word-initial. /b/ was always preceded by

Table 1
Self-Ratings on Language Proficiency (SD) and LexTALE Scores (SD)

Language	Listening[a]	Speaking[a]	Reading[a]	Writing[a]	Overall mean[a]	LexTALE
Dutch						
Confederate present	5 (.00)	5 (.00)	5 (.00)	5 (.00)	5 (0)	92.11 (4.49)
Confederate absent	5 (.00)	5 (.00)	5 (.00)	5 (.00)	5 (0)	85.16 (14.65)
English						
Confederate present	2.56 (.50)	2.38 (.77)	2.69 (.42)	2.31 (.77)	2.48 (.80)	76.80 (12.62)
Confederate absent	2.69 (.60)	2.31 (.60)	2.50 (.63)	2.25 (.58)	2.44 (.54)	70.94 (12.49)

Note. There were no significant differences between English proficiency scores in the confederate absent and confederate present groups (all p values $>$ .1). The difference between the proficiency scores for Dutch and English was significant in each condition (all p values $<$ .0001).
[a] Ratings were given on a scale from 1 to 5 with 1 = *not at all* and 5 = *native speaker*.

Table 2
Design of the Experiment

Block	Sentences	Speaker	Lag	Social context
Baseline block	30 sentences with words including /æ/ and word-final /b/	Participant	No lag	Confederate present/absent
Exposure block	30 further sentences with same targets as baseline block	Confederate	No lag	
Alternating block	60 further sentences (targets appeared twice in this block: once for participant and once for confederate) + 30 fillers	Participant + Confederate	Lag 0 + lag 4	

a vowel in a target word (as in "tub" and "job").[2] The target words never occurred at the end of a sentence, or before /f/ and /v/ because the /b/ becomes a labial-dental sound if it precedes these phonemes (as in "obvious"; Collins & Mees, 1996). Therefore, /b/ was always followed by a vowel.

Each /b/-target word was randomly paired with an/æ/-target word in a sentence for the baseline blocks. In the exposure block, the /b/-target word was again randomly paired with another /æ/-target word in another sentence, resulting in 60 sentences containing one instance of each contrast created for the first two blocks. An additional two sentences containing only one target word were created for each target word for the postexposure block. There were no particular constraints on the sentences: They were constructed by the authors, both long and short sentences were included, and the sentences were nonconstraining toward the target words. The confederate checked whether the sentences were grammatically correct before the experiments were run; she corrected one sentence.

Two presentation lists were created for each block where the sentences were presented in pseudorandom order: The pattern of the lag manipulation in the alternating block was the same for both lists, but the order of the sentences was randomized. Each list could be presented in version A or B so that the sentences read by the participant in version A were read by the confederate in version B and vice versa.

Procedure

In the confederate-present context, the experimenter went to pick up the participant and the confederate in the hall of a university building. Throughout the experiment, the confederate acted as if she was just another participant and the confederate did not speak English before the experiment started. In the confederate-absent context, the confederate was seated in another room and the participant did not see the confederate during the experiment. In this condition, participants were told that they would be listening to recordings of spontaneous speech and participants thought they were the only one being tested. Participants received oral and written instructions in Dutch to read aloud the English sentences presented on the screen. We told the participants that the experiment tested whether comprehension of sentences was better when participants read the sentences or when someone else read the sentences (i.e., the confederate). This explanation was provided to draw the participants' attention away from the true goal of the experiment.

Participants were tested in a silent room and were seated in front of a computer screen and a microphone while wearing headphones. In both the confederate-present and the confederate-absent context, the participant, the confederate, and the experimenter each worked on a laptop computer. The experimenter used his laptop to record the speech of the confederate and the participant. The confederate's and participant's laptop were used for visual stimulus presentation by means of the computer program E-prime 2.0. The confederate's and the participant's microphones were connected to a mixer, which was in turn connected to the experimenter's laptop. The recordings were made in *Audacity* with a sampling frequency of 48 kHz. The participant and confederate heard each other live over headphones both when the confederate was present in the same room and when she was seated in the other room. None of the participants in the confederate absent context noticed that the confederate's speech production was live instead of a recording. The confederate's speech was live in both conditions to keep the conditions as similar as possible on all variables except for physical confederate presence; pronunciation of the sentences was of comparable variability and the confederate could also hear the participant's speech in both versions.

Table 2 summarizes the design. In the baseline block, participants read the sentences out loud, while the sentences were read by the confederate in the exposure block. In the alternating block, the participant and confederate each read a sentence in turn. Every trial started with a fixation cross on the screen, after which a sentence was presented if it was the participant's turn to read a sentence. When the confederate read aloud a sentence, a picture of an ear and the text "Listen" was presented on the participant's screen. The sentence or the word Listen remained on the screen until the participant pressed a button, after which the next sentence was presented. A comprehension question was presented after 10% of the sentences. The participant and confederate (when present) were asked to answer the questions by pressing the F-button if the statement about the sentence was incorrect and the J-button if it was correct. To ensure that the participant and the confederate continued at the same pace with the next trial, they were asked to say "okay" before continuing after answering a question. Only the participant was asked to say okay after answering a question in the confederate absent context. After the experiment, participants were asked whether they thought they knew what the experiment was about. None of the participants suspected that the experiment was about their pronunciation, and hence neither about alignment.

[2] For one sentence with a target word that ended in word-final /b/, the word "while" followed the target word ("stub") instead of a word starting with a vowel. As the intercept "sentence" was included in the linear mixed effects model this should not lead to problems in the analyses.

Acoustic Measures and Annotation

Analyses were performed on the recordings of the participants' speech. The target sounds were annotated by hand using Praat (Broersma & Weenink, 2014) after which a script was used to extract the formant frequencies of the first and second spectral peaks (F1 and F2) and the length of annotated vowel and word segments. For /æ/, both the vowel itself and the entire word were annotated. For word-final /b/, the preceding vowel, closure duration, and the entire word were annotated.

Phoneme boundaries were determined as accurately as possible through visual and auditory inspection. Vowel boundaries were placed at F2 onset and offset in the spectrogram or, if F2 onset or offset was unclear, where two or more formants appear or drop out together (Hwang et al., 2015). The offset of the target word with /b/ was always set right after the release burst of /b/. If the release was not audible and/or visible, it was placed immediately before the onset of the next word. Closure duration was defined as the length of the segment from vowel offset until the release burst. If the release was not visible and/or audible, closure duration was not taken into account.

The Praat script determined the formants using a 0.00625 s time step and a 0.025 s window length. Formant frequencies were then aggregated so that the dataset contained one mean formant frequency for F1 and F2 for each produced phoneme (see Appendix B for a table displaying raw values of formant frequencies and durations). To be able to create a measure of /æ/ that was normalized to each participant's vowel space, we also annotated all occurrences of /ɛ/ in the experiment. Depending on the list, there were 17 or 22 occurrences of /ɛ/ in the baseline block and 43 or 52 occurrences of /ɛ/ in the postexposure block. The frequencies of F1 and F2 of /æ/ and /ɛ/ were transformed to the psychoacoustic Bark scale for analysis (Traunmüller, 1990). The participants' F1 and F2 values of /æ/ were then divided by the mean F1 and F2 formant frequency of /ɛ/ (respectively) in the same block to create the normalized measure. This measure is more informative than plain F1 and F2 of /æ/, because it shows to what extent L2 speakers of English distinguish between /æ/ and /ɛ/. The experimental set-up induced considerable noise to the recordings. However, the spectrograms showed that the speech signal was considerably stronger than the noise signal.

The duration measures used for the analyses of the production of the vowel preceding word-final /b/, closure duration, and /æ/-duration were relative (the duration of the vowel/closure divided by the duration of the word). This relative measure of vowel length was used to correct for speech rate. In the analyses, when we refer to F1, F2 or duration, we always refer to the normalized measures. All values above and below 2.5 *SD*s of a participant's mean for an item were excluded from the analysis.

Annotation took approximately 250 h; the task was divided over five researchers. Interclass correlation (ICC) was calculated for all duration measures based on the preexposure block of a randomly selected subject by means of the package ICC in R (3.4.1; R Core Team, 2013). ICC was only calculated for duration because segment duration directly reflects placement of phoneme boundaries. Two-way models were used with type "agreement" and unit as definition. There was a high degree of reliability between phoneme boundary placement for almost all measures (see Table 3 below).

The ICC of closure duration is low because of many missing values in the measurements (where only five instances of closure duration were measured by one of the annotators). The release of the /b/ was not always audible and/or visible and, therefore, this particular measure has more missing data. The percentage of annotated closure durations amounted to 69.5% (1335/1920) in the confederate data set and to 60.9% (1169/1920) in the participant data set.

Analyses

We first determined whether there were substantial differences between the Participant's and the Confederate's[3] acoustic characteristics for each target phoneme. Then, we tested whether Participants' phonetic characteristics changed after exposure to the Confederate's speech by comparing the postexposure (alternating) block and the preexposure (baseline) block, and whether the degree of change depended on social context (the presence or absence of a Confederate during the experimental session). Additionally, we tested whether phoneme production in the postexposure block was more similar to that of the Confederate immediately after the Participant had heard the Confederate's production of the phoneme (lag 0) than when four sentences intervened between perception and production (lag 4). Finally, we tested whether mere repetition of the target sounds lead to changes in Participants' production by assessing change over the course of the baseline block and whether listening to and producing target phonemes in the postexposure block lead to additional changes over the course of that block (trial number effects).

We ran the same analyses for the Confederate and additionally tested whether she also changed her target phoneme production over the course of experimental sessions (one Participant was tested each experimental session). For the Confederate, trial number effects were only assessed in the postexposure block to test whether more interaction with the participants led to (more) adaptation over the course of the postexposure (alternating block). Whether mere repetition of the target phoneme lead to changes in the confederate's target phoneme productions was not of interest here. Additionally, the exposure block was not a true baseline block like the baseline block for the participants (because the confederate already heard the participant's production during the baseline block at this point). Therefore, we did not assess the effects of trial number in the exposure block.

Our data set was analyzed with linear mixed effects models in R (Version 3.4.0); *p* values for the fixed effects and interactions in the final models were computed using the lmerTest package (Version 2.0–33; Satterthwaite degrees of freedom approximation; Kuznetsova, Brockhoff, & Christensen, 2016). First, we ran a simple model for the normalized measures of duration of the vowel preceding word-final /b/, closure duration F1, F2, and /æ/ duration separately. These simple models included the three main experimental fixed factors social context (confederate absent/present), block (preexposure and postexposure), and list (control variable: the different stimuli presentation lists) as well as the interaction between social context and block. The random intercepts were participant, word, and sentence. Participants' L2 proficiency (centered LexTALE score) and the two- and three-way interactions between proficiency, block, and context (participant data only) were only added to

[3] In the Analyses and Results section we use Participant and Confederate (with capital letter) to refer to experimental role. The terms are not capitalized when they refer to experimental factors (by participants random intercept or confederate absent/present condition).

Table 3
Interclass Correlation Information on Five Different Measures

Measure	ICC	Lower CI (2.5%)	Upper CI (97.5%)	*F* value (*df*)	*p* value
Word duration /æ/	.821	.678	.913	33.6 (22, 33.1)	<.001
Vowel duration /æ/	.672	.499	.823	12.8 (21, 68.9)	<.001
Word duration /b/	.825	.699	.910	31.5 (24, 45.8)	<.001
Vowel duration before /b/	.823	.700	.902	22.7 (23, 95.9)	<.001
Closure duration	.209	−.055	.703	2.34 (5, 23.3)	.074

Note. ICC = interclass correlation. *F*- and *p* values indicate whether the correlation significantly differs from zero.

the model if they contributed to the model fit. Similarly, experimental session and the interactions between session, block, and context were only added to the models for the confederate data set if they contributed to the model fit. Note here that every session had a new Participant but the same Confederate. Participant was never included as random intercept when session was a fixed factor in the model because the intercept captured the same information.

Subsequently, random slopes were determined by comparing models with and without each random slope with a χ^2 test (Baayen, 2008). If the models differed significantly, then the model that explained the most variance and with the lowest Akaike's Information Criterion (AIC) value was used. Random slopes were tested in a fixed order (block, condition, list, and trial number if applicable). Also, the random effects structure was simplified if running the model resulted in convergence errors.

A separate linear mixed effects model was constructed for the data from the postexposure block to test whether there was an effect of Lag. This model included the fixed factors Lag (0 or 4 sentences), social context (confederate present or absent), and their interaction. The control variable presentation list was also included as a fixed factor. Random intercepts of participant, word, and sentence were included and random slopes were once again determined by model comparison. The effects of trial number were also assessed separately in the baseline and postexposure block. The data sets and scripts used for the analyses are available online at Open Science Framework (https://osf.io/p62j4/).

Results

Target Phoneme /b/

Figure 1 below shows that the Confederate produced longer vowels preceding word-final /b/ than the Participants, whereas Figure 2 shows that the Participants had a longer mean closure duration than the Confederate. Two linear mixed effects models with speaker (Confederate vs. Participant) as fixed factor, session as random intercept, and random slope for session were constructed to test whether these differences were significant. The differences were significant for both vowel length ($\beta = -.11$, $SE = .008$, $t = -13.29$, $p < .001$) and closure duration ($\beta = .09$, $SE = .04$, $t = 20.14$, $p < .001$).

Participants. There was a main effect of block (baseline vs. postexposure) for the Participants on both measures (preceding vowel duration: ($\beta = .02$, $SE = .004$, $t = 4.56$, $p < .001$); closure duration: $\beta = -.02$, $SE = .005$, $t = -3.59$, $p < .001$). The duration of the vowel preceding word-final /b/ increased after exposure to the Confederate's speech, and the closure duration of the Participants decreased. Thus, the Participants' production of final /b/ became more like the Confederate's production on both acoustic measures. The interaction between block and social context was not significant (preceding vowel duration: $\beta = -.002$, $SE = .006$, $t = -.35$, $p = .72$; closure duration: $\beta = .004$, $SE = .007$, $t = .57$, $p = .57$; full results are presented in Table C1 in Appendix C). Finally, L2 proficiency did not improve the model fit

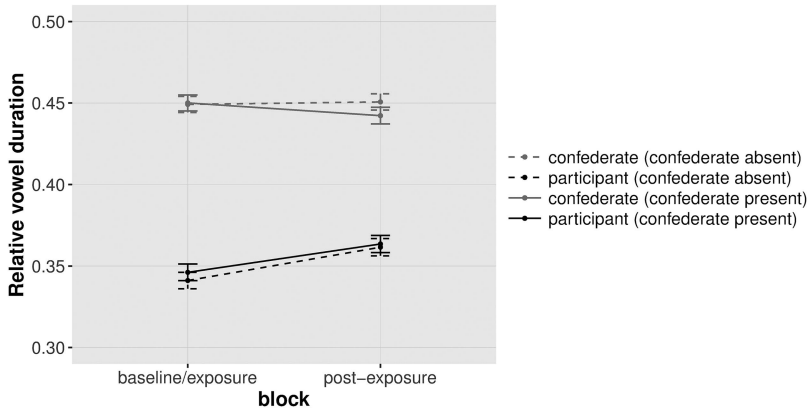

Figure 1. Relative vowel duration of the vowel preceding word-final /b/ in the baseline and postexposure block for the Participant and in the exposure and postexposure block for the Confederate. Error bars denote *SE*s.

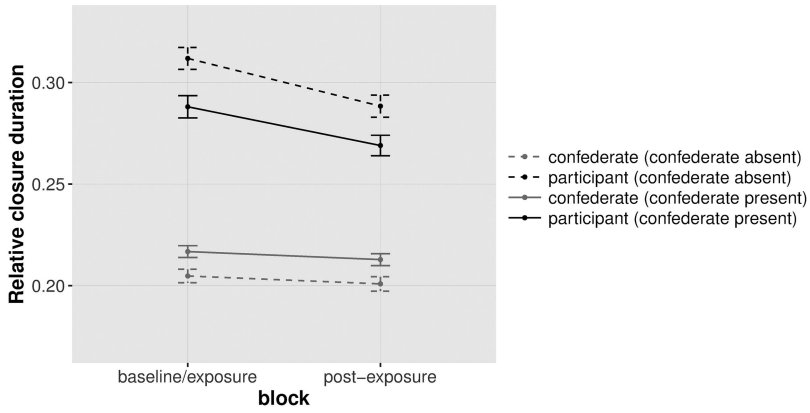

Figure 2. Relative closure duration of the vowel preceding word-final /b/ in the baseline and postexposure block for the Participant and in the exposure and postexposure block for the Confederate. Error bars denote *SE*s.

(preceding vowel duration: $\chi^2(4) = 8.1$. $p = .09$; closure duration: $\chi^2(4) = 7.05$, $p = .13$).

There was no main effect of trial number in the baseline block (p values $> .1$). The postexposure block, however, did reveal a main effect of trial number on vowel duration only ($\beta = .001$, $SE = .0004$, $t = 2.45$, $p = .01$); the vowel preceding /b/ became longer over the course of the postexposure (alternating) block. There were no interaction effect between trial number and social context in either the baseline or postexposure block (p values $> .05$).

The main effect of lag did not reach significance (preceding vowel duration: $\beta = .009$, $SE = .007$, $t = 1.26$, $p = .21$; closure duration: $\beta = .001$, $SE = .008$, $t = .16$, $p = .88$), nor did the interaction of lag and social context (preceding vowel duration: $\beta = -.006$, $SE = .009$, $t = -.7$. $p = .48$; closure duration: $\beta = -.0005$, $SE = .009$, $t = -.05$, $p = .96$).

Confederate. Experimental session improved the model fit for preceding vowel duration ($\chi^2(4) = 20.38$, $p < .001$) and for closure duration ($\chi^2(4) = 9.89$, $p = .04$). This factor was, therefore, included in the final models. There was only a significant main effect of session for vowel duration ($\beta = -.0009$, $SE = .0003$, $t = -2.95$, $p = .003$), with the Confederate's relative vowel length decreasing over the course of experimental sessions. No other main effects or two- and three-way interactions between session, block and social context were significant for closure duration or preceding vowel duration (all p values $> .05$; full results are presented in Table C2 in Appendix C). The main effect of trial number did not reach significance in the postexposure block and there was no interaction between trial number and social context on either measure (p values $> .1$).

Summary target phoneme /b/. Participants showed an adaptation effect for both the duration of the vowel preceding word-final /b/ and closure duration. The increase in the Participants' duration of the vowel preceding word-final /b/ over the course of the postexposure block suggests that Participants adapted vowel length more after hearing and producing more target sounds. No effects of social context or lag were found. The Confederate did not adapt the duration of these measures to the Participants' productions from the exposure block to the postexposure (alternating) block, but she did significantly shorten her vowels preceding /b/ (they became closer to the Participants' vowel length) after taking part in more experimental sessions.

Target Phoneme /æ/

Figure 3 to 5 show the normalized mean F1 scores, F2 scores, and the relative duration of /æ/ for the Participants and the Confederate in

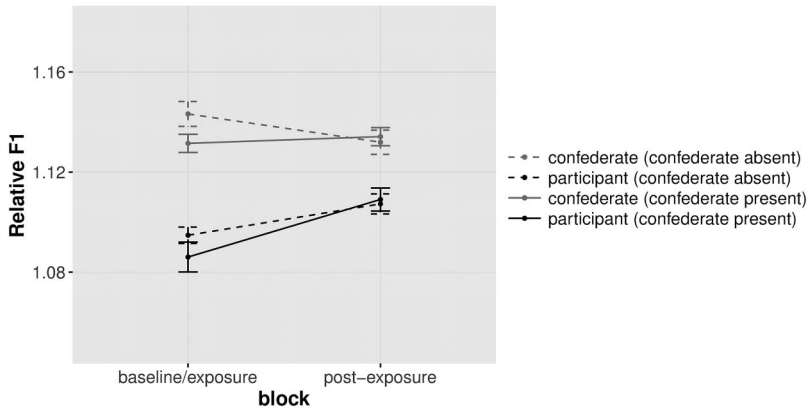

Figure 3. Relative F1 frequencies of target vowel /æ/ in Bark in the baseline and postexposure block for the Participant and in the exposure and postexposure block for the Confederate. Error bars denote *SE*s.

the confederate-present and -absent contexts before and after exposure. Figure 3 shows that, as expected, the Participants' mean F1 was lower than the Confederate's. A linear mixed effects model with speaker (Confederate vs. Participant) as a fixed factor and session as random intercept that was run for the baseline and exposure block data confirmed this ($\beta = -.05$, $SE = .005$, $t = -10.43$, $p < .001$). The Participants' F2 values in the baseline block were also significantly different from the Confederate's F2 values in the exposure block ($\beta = -.008$, $SE = .004$, $t = -2.20$, $p = .036$; see Figure 4). A final model indicated a significant difference in mean duration of /æ/ between Participants and Confederate ($\beta = -.07$, $SE = .009$, $t = -7.32$, $p < .001$; see Figure 5).

Participants. The difference between the Participants' production of /æ/ in the baseline and postexposure block was not significant for any of the acoustic measures (F1: $\beta = .01$, $SE = .01$, $t = 1.36$, $p = .18$, F2 $\beta = -.002$, $SE = .002$, $t = -.67$, $p = .16$; /æ/ duration: $\beta = .01$, $SE = .008$, $t = 1.62$, $p = .11$). The interaction between block (baseline vs. postexposure block) and social context (confederate present vs. confederate absent) was not significant either (F1: $\beta = .01$, $SE = .01$, $t = .84$, $p = .41$; F2: $\beta = -.001$, $SE = .003$, $t = -.44$, $p = .66$; /æ/ duration: $\beta = -.0002$, $SE = .01$, $t = -.03$, $p = .98$).[4] Full results are presented in Table C3 of Appendix C.

L2 proficiency and the interactions between proficiency, block, and social context did not improve the model fit for any of the acoustic measures (F1: $\chi^2(4) = 3.03$, $p = .55$; F2: $\chi^2(4) = 1.69$, $p = .79$, duration: $\chi^2(4) = 4.28$, $p = .37$). The main effect of trial number was not significant in the baseline block, nor in the postexposure (alternating) block for any measure (p values $> .1$). The interaction between trial number and social context was not significant either (p values $> .05$).

There was no effect of time lag between perception and production (zero vs. four intervening sentences) in the postexposure block (F1: $\beta = -.001$, $SE = .008$, $t = -.14$, $p = .89$; F2: $\beta = .002$, $SE = .003$, $t = .46$, $p = .65$; duration: $\beta = -.003$, $SE = .008$, $t = -.36$, $p = .72$), or an interaction between time lag and social context (F1: $\beta = .006$, $SE = .01$, $t = .6$. $p = .55$; F2: $\beta = -.007$, $SE = .004$, $t = -1.66$, $p = .10$; duration: $\beta = -.004$, $SE = .008$, $t = -.53$, $p = .60$).

Confederate. As for the Confederate, there was no significant effect of block on F1 ($\beta = .009$, $SE = .008$, $t = 1.09$, $p = .27$) or on duration ($\beta = .003$, $SE = .006$, $t = .43$, $p = .67$). The Confederate did significantly decrease her F2 from the exposure block to the postexposure (alternating) block ($\beta = -.016$, $SE = .005$, $t = -2.95$, $p = .004$). A main effect of social context was found for F2 ($\beta = -.02$, $SE = .004$, $t = -6.15$, $p < .001$), with the Confederate's F2 being lower in the present than in the absent condition. Social context was also significant for duration ($\beta = -.03$, $SE = .01$, $t = -2.48$, $p = .02$), with the Confederate producing shorter vowels in the present condition. The interaction between block and social context was not significant (F1: $\beta = -.008$, $SE = .009$, $t = -.89$, $p = .37$; F2: $\beta = -.008$, $SE = .005$, $t = 1.52$, $p = .13$; duration: $\beta = .005$, $SE = .006$, $t = .9$. $p = .37$.

The factor experimental session and the two- and three-way interactions of session, block and social context improved the model fit for F1 ($\chi^2(4) = 26.54$, $p < .001$) and for F2 ($\chi^2(4) = 77.55$, $p < .001$) and were, therefore, included in the final models for those measures. The effect of session did not contribute to the model fit for duration ($\chi^2(4) = 4.14$, $p = .39$). There was a three-way interaction between block, social context, and session for F1 ($\beta = .001$, $SE = .0004$, $t = 2.68$, $p = .008$). Post hoc tests with lsmeans showed that in the exposure block in the absent condition, F1 increased significantly over the course of experimental sessions ($\beta = .0007$, $SE = .0002$, $t = 2.87$, $p = .008$), but not in the postexposure block ($\beta = -.0004$, $SE = .0002$, $t = -1.82$, $p = .13$). In the present condition there was a significant decrease of F1 over sessions both in the exposure ($\beta = -.0006$, $SE = .0002$, $t = -2.96$, $p = .006$) and the postexposure block ($\beta = -.0005$, $SE = .0002$, $t = -2.52$, $p = .02$). There was a main effect of session on F2 ($\beta = -.0008$, $SE = .0001$, $t = -5.62$. $p < .001$) and session also interacted with condition ($\beta = .001$, $SE = .0002$, $t = 6.49$, $p < .001$). Post hoc tests with lsmeans revealed a positive trend for session in the present condition and a negative trend in the absent condition. Full results are presented in Table C4 in Appendix C.

The effect of trial number in the postexposure block was not significant for the F1 ($\beta = -.0007$, $SE = .0004$, $t = -1.74$, $p = .09$), but there was a main effect of trial number for the F2 in the postexposure block $\beta = -.0006$, $SE = .0002$, $t = -3.51$, $p < .001$). This suggests a further downward change of F2 over the course of the postexposure (alternating) block. The effect of trial number was also significant for vowel duration in the postexposure block ($\beta = -.001$, $SE = .0004$, $t = -3.48$, $p < .001$): the Confederate shortened her vowels over the course of the postexposure block. There were also a significant interaction between condition and trial number for F2 ($\beta = .004$, $SE = .0002$, $t = 2.10$, $p = .037$) and for vowel duration ($\beta = .001$, $SE = .0004$, $t = -2.36$, $p < .02$), indicating that the adjustment over trials was larger in the absent than in the present condition.

Summary target phoneme /æ/. Participants did not show a change in their pronunciation of /æ/ after exposure to /æ/ pronounced by the Confederate. Time lag between perception and production of /æ/ did not affect pronunciation either. The confederate lowered her F2 from exposure to postexposure, but there was no change in her F1, or vowel duration. The Confederate's F2 was lower and her vowel duration shorter in the present than in the absent condition. Further, the confederate's F1 increased over sessions in the exposure block in the absent condition, and decreased over sessions in the present condition. The confederate's F2 increased over sessions in the present condition and decreased over sessions in the absent condition.

Discussion

Aligning with a native speaker may be a useful mechanism for language learning. On the other hand, L2 speakers may be too dissimilar from native speakers for phonetic alignment to occur. The aim of the present study was to test whether unbalanced

[4] There were five participants whose mean relative F1 was higher than the confederate's at baseline, and there were 12 participants whose mean relative F2 was higher than the confederate's at baseline. We conducted additional analyses where participants with a higher mean F1 and F2 value at baseline were excluded. As their initial F1 and F2 values were higher than that of the confederate, one would not expect to see phonetic alignment in these participants (or maybe even reversed alignment). However, no main effects of block or interactions between block and social context were found (all p-values $> .05$).

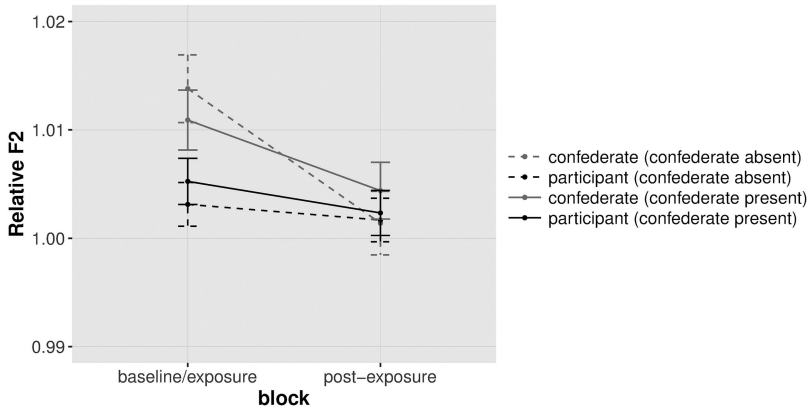

Figure 4. F2 frequencies of target vowel /æ/ in Bark in the baseline and postexposure block for the Participant and in the exposure and postexposure block for the Confederate. Error bars denote *SE*s.

Dutch-English bilinguals adapt their L2 speech after listening to a native speaker of the target language. Additionally, we tested whether a native English confederate also adapted her pronunciation to our (nonnative) participants' pronunciation. In particular, we focused on the pronunciation of the phoneme /æ/ and the vowel preceding word-final /b/ in English.

There was significant alignment of the participant to the confederate for closure duration of word-final /b/ and duration of the vowels preceding word-final /b/. Specifically, closure duration of the participants was shortened in the postexposure block compared with baseline whereas the duration of the participants' vowel preceding word-final /b/ increased. However, there was no alignment for the other target phoneme, /æ/. No main effect of block was seen on the F1, F2, or duration of /æ/ for the participant. Social context did not affect alignment of either phoneme, nor did time lag between perception and production.

The finding that L2 speakers of English adapt their pronunciation of word-final /b/ and the preceding vowel supports the findings of Hwang et al. (2015), who also found alignment of L2 speakers in L2-L1 dialogue for /b/ (on preceding vowel duration but not closure voicing duration) and /æ/ (on vowel duration and F1 but not F2). It also strengthens the claim that alignment takes place when speakers can improve their L2 pronunciation by adapting to L1 speech. As demonstrated by the lack of a trial number effect in the baseline block, the adaptation of word-final /b/ was not merely an effect of repeated production of the phoneme.

However, the lack of alignment on the target vowel /æ/ suggests that alignment by L2 speakers does not occur under all circumstances. Perhaps our participants could not sufficiently perceive the difference between their own speech and that of the native speaker. Dutch native speakers often have difficulty distinguishing /æ/ and /ɛ/ in speech perception (Broersma, 2005; Weber & Cutler, 2004). If the difference in pronunciation cannot always be perceived by Dutch speakers, then it might be very hard if not impossible for them to adjust their phoneme boundaries of this particular vowel. In contrast, Dutch listeners have no difficulty distinguishing /b/ and /p/, as /b/ does occur in Dutch (only not at the end of the word).

The acoustic characteristics of the participants' word-final /b/ in the postexposure block were not affected by the number of sentences (zero or four) intervening between the participants' and confederate's production of the target phoneme. This find-

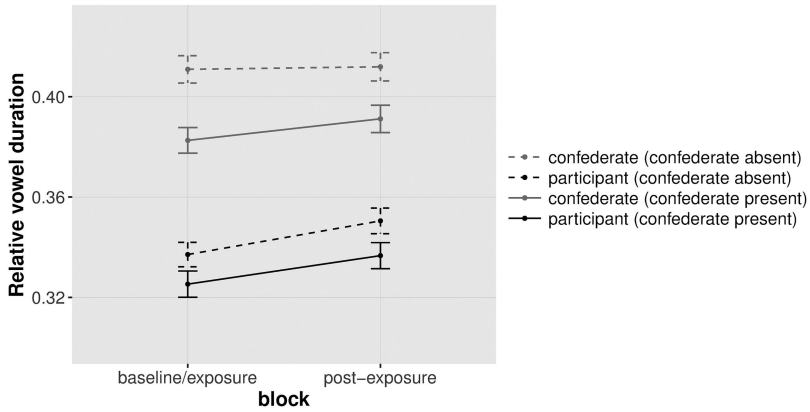

Figure 5. Relative vowel duration of the target vowel /æ/ (duration of the vowel divided by duration of the word) in the baseline and postexposure block for the Participant and in the exposure and postexposure block for the Confederate. Error bars denote *SE*s.

ing extends the observations of Hwang et al. (2015), who found alignment in L2 speech after immediate priming by the L1 confederate without including a lag between target words. We found alignment of word-final /b/ both in the immediate condition (lag 0) and the delayed condition (lag 4). An account in terms of automatic priming would predict time lag effects. Possibly, the influence of an exposure to a native phoneme is relatively long-lasting, so that the confederate's production four trials back still affects the participant's current production. However, it is also possible that the cumulative influence of the confederate's productions during the exposure phase was strong enough to last during the postexposure phase, so that any new exposure had little further effect. Also, simulation theory (Gambi & Pickering, 2013) would predict that when episodes of comprehension are tightly interwoven with episodes of production (like in our postexposure/alternating block), simulation should be enhanced. This would perhaps not predict an effect of time delay between perception and production of a specific phoneme, but an effect of time delay between speech perception and production in general. In the postexposure block in our study, the time delay between speech perception and production was always short. The effect of trial number on the length of the vowel preceding word-final /b/ (further lengthening of the vowel over the course of the postexposure block), supports this claim.

The present study also tested whether there was a difference in the amount of phonetic alignment between an L2 speaker and an L1 speaker when the L1 speaker was physically present or absent. Based on the Interactive Alignment Model, priming should result in alignment, irrespective of the social context. However, if alignment is not solely based on priming but is also modulated by contextual factors (e.g., social context, motivation, or beliefs about an interlocutor), the presence of a confederate may boost alignment. Hence, we hypothesized that the actual *presence* of the confederate would have an influence on the amount of phonetic alignment. Yet, no social context effects (effects of confederate presence) were found for the participants. Gambi and Pickering (2013) suggest that phonetic adaptation through simulation depends on the allocation of limited attentional resources. Perhaps in our study, because of disadvantages associated with L2 processing, the nonnative speakers had less resources available to further adjust their pronunciation of /b/ to that of the confederate's when the confederate was present.

The confederate did not show consistent alignment with the participants. The confederate adjusted her F2 of /æ/ from the exposure block to the postexposure (alternating) block (in the direction of the participant mean). Within the postexposure block (the alternating block) the confederate further lowered her F2 value and she also shortened the vowel /æ/ over the course of the postexposure block (in the direction of the participants). The confederate's pronunciation also changed over experimental sessions, but there was no systematic convergence with the participants. The confederate's vowel before /b/ became shorter across sessions (closer to the participants). Her F1 of /æ/ increased across sessions in the absent condition in the exposure block (diverging from the participants), but not in the postexposure block. She also lowered her F1 across sessions in the confederate present condition in both blocks (converging with the participants). F2 became significantly lower across sessions in the absent condition (diverging from the participants) and higher in the present condition.

Our findings partially support the Interactive Alignment account (Garrod & Pickering, 2004), which assumes that alignment is a rather automatic process. On the one hand, it is supported specifically by the findings that there was no support for a modulation of alignment on /b/ by social context, suggesting that alignment occurs automatically without considering the situation. Moreover, the participants were unaware of the goal of the experiment. On the other hand, the finding that the confederate did not align her speech toward that of the participants does not support the automaticity of alignment.

Simulation theory (Gambi & Pickering, 2013) can account for this apparent inconsistency if we assume that L2 speakers aspire to be more similar to L1 speakers (and, therefore, perceive themselves as being more similar), whereas L1 speakers perceive themselves to be very dissimilar from L2 speakers. Gambi and Pickering (2013) suggest that when the perceived difference between two interlocutors is too large, interlocutors may rely less on simulations of the other person's speech. If less simulation occurs during speech comprehension, then there should also be less influence of simulations on one's own speech production. Also, simulation of an L2 speaker's utterances by an L1 speaker may simply fail because the L1 speaker lacks experience with the L2 speaker's utterances. Even though word final /p/ exists in English like in Dutch (e.g., *hip hop*), the devoiced pronunciation of word final /b/ in English words (e.g., *blop* instead of *blob*) by L2 speakers may be unfamiliar/unexpected to a native speaker. Therefore, the L1 speaker may be slow to adjust her predictions of the L2 speakers' utterances and, therefore, alignment may fail.

The current study focused on alignment by L2 participants, rather than native speakers and, therefore, only included one confederate. Therefore, the lack of consistent alignment in the confederate's speech data set might also be because of individual characteristics of the confederate. In this study we wanted the participants to be exposed to the same speaker to reduce variability, but future research with multiple confederates could point out whether adaptation of an L1 speaker toward an L2 speaker occurs under some circumstances.

Gambi and Pickering (2013) suggest that when there is more information available at linguistic levels other than the phoneme level, limited availability of attentional resources may cause predictions to be based on those levels (such as the word or sentence level). Phonetic imitation may, therefore, be less pronounced. Perhaps this can explain why Hwang et al. (2015) did find adaptation of /æ/ in nonnative speakers when primed by a native speaker, whereas we did not. In their experiment, confederate utterances were very simple (e.g., "what is below Hob"). Our stimuli contained longer and more complex sentences and participants may, therefore, have made use of predictions at other linguistic levels, making them less sensitive to variations at the phonetic level.

A potential limitation of the current study is that the baseline block was not entirely identical across conditions. During the baseline blocks, the confederate was present in the same room as the participant in the confederate present condition whereas she was absent in the absent condition. The sole presence of the confederate might have influenced pronunciation of the partic-

ipant in the baseline block, for example by motivating the participant to produce the sentences with a more native-like accent. That being said, the confederate did not speak English (nor Dutch) up until the exposure block, meaning that the confederate's speech could not have affected the participants' utterances at baseline. Moreover, there was no main effect of social context nor an interaction between block and social context for the measures that showed alignment (vowel duration preceding word-final /b/ and closure duration). Therefore, we argue that this inconsistency would not have greatly affected the results.

In conclusion, results from the current study show that speech production in L2 is influenced by exposure to speech produced by a native speaker of that language. However, the effect depended on the particular phoneme, possibly related to the degree to which participants can perceive the relevant phonemic distinction. Adaptations seem to last over at least four intervening trials. There was no compelling evidence that such influences are affected by social factors.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511801686

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40,* 177–189. http://dx.doi.org/10.1016/j.wocn.2011.09.001

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2012). Effects of phonological feedback on the selection of syntax: Evidence from between-language syntactic priming. *Bilingualism: Language and Cognition, 15,* 503–516. http://dx.doi.org/10.1017/S1366728911000162

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition, 127,* 287–306. http://dx.doi.org/10.1016/j.cognition.2013.02.005

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11,* 303–328. http://dx.doi.org/10.1017/S0142716400008912

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106,* 707–729. http://dx.doi.org/10.1016/j.cognition.2007.04.005

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition, 121,* 41–57. http://dx.doi.org/10.1016/j.cognition.2011.05.011

Broersma, M. E. (2005). *Phonetic and lexical processing in a second language*. Nijmegen, the Netherlands: Radboud Universiteit Nijmegen.

Broersma, P., & Weenink, D. (2014). *Praat: Doing phonetics by computer* (Version 5.3.78). Retrieved from www.praat.org

Collins, B., & Mees, I. (1996). *The phonetics of English and Dutch* (rev. ed.). Leiden: Brill.

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes, 23,* 528–556. http://dx.doi.org/10.1080/01690960801920545

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America, 119,* 1950–1953. http://dx.doi.org/10.1121/1.2178721

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology, 4,* 340. http://dx.doi.org/10.3389/fpsyg.2013.00340

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8,* 8–11. http://dx.doi.org/10.1016/j.tics.2003.10.016

Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139166126

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language, 81,* 72–90. http://dx.doi.org/10.1016/j.jml.2015.01.001

Kim, M. (2012). *Phonetic accommodation after auditory exposure to native and nonnative speech*. Evanston, IL: Northwestern University. Retrieved from http://www.linguistics.northwestern.edu/documents/dissertations/linguistics-research-graduate-dissertations-kimdissertation2012.pdf

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology, 2,* 125–156. http://dx.doi.org/10.1515/labphon.2011.004

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition, 107,* 54–81. http://dx.doi.org/10.1016/j.cognition.2007.07.013

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models*. Retrieved from https://CRAN.R-project.org/package=lmerTest

Lametti, D. R., Krol, S. A., Shiller, D. M., & Ostry, D. J. (2014). Brief periods of auditory perceptual training can determine the sensory targets of speech motor learning. *Psychological Science, 25,* 1325–1336. http://dx.doi.org/10.1177/0956797614529978

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44,* 325–343. http://dx.doi.org/10.3758/s13428-011-0146-0

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language, 33,* 42. http://dx.doi.org/10.2307/410949

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America, 96,* 2076–2087. http://dx.doi.org/10.1121/1.410149

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *The Journal of the Acoustical Society of America, 78,* 1949–1957. http://dx.doi.org/10.1121/1.392651

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238. http://dx.doi.org/10.1016/S0010-0285(03)00006-9

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119,* 2382–2393. http://dx.doi.org/10.1121/1.2178720

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences, 3,* 136–141. http://dx.doi.org/10.1016/S1364-6613(99)01293-0

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36,* 329–347. http://dx.doi.org/10.1017/S0140525X12001495

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51,* 1296–1303. http://dx.doi.org/10.1121/1.1912974

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, Retrieved from http://www.R-project.org/

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America, 88,* 97–100. http://dx.doi.org/10.1121/1.399849

Trofimovich, P., & Kennedy, S. (2014). Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. *Bilingualism: Language and Cognition, 17,* 822–836. http://dx.doi.org/10.1017/S1366728913000801

Weber, A., Betta, A. M. D., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics, 46,* 34–51. http://dx.doi.org/10.1016/j.wocn.2014.05.002

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50,* 1–25. http://dx.doi.org/10.1016/S0749-596X(03)00105-0

Appendix A

Stimuli Sentences and Target Words

Sentence	Target word /æ/	Target word /b/
The Russian mob of New York was glad the police did not arrest them.	Glad	Mob
They prescribe a type of medicine that decreases gas in your bowels.	Gas	Prescribe
We rob all people with a hammer, said the criminal.	Hammer	Rob
The rich snob often paints a portrait of a landscape outside.	Landscape	Snob
The man was sitting on a stub while thinking about his future.	Man	Stub
While being in the pub on Main Street, he tends to slap people.	Slap	Pub
The woman decided to show a boob on the tram in the city center.	Tram	Boob
This band tours around the globe every 2 years.	Band	Globe
A friend of mine broke his rib on his left side because of a bat on the baseball field.	Bat	Rib
A tube of sand was used during the experiment.	Sand	Tube
Suzanne's job in the music industry was to rap on stage.	Rap	Job
Much of the fat was reduced with a probe inserted into the tissue by a doctor.	Fat	Probe
The panther lay on a stone in the form of a cube in the jungle.	Panther	Cube
He told me to rub a lamp to see a genie.	Lamp	Rub
I always enjoyed it when I had to dub a movie.	Had	Dub
Either choose a robe or a mantle, but not both.	Mantle	Robe
He felt a throb in his head due to the scam of the criminal.	Scam	Throb
They plan to bribe all the supervisors of the company.	Plan	Bribe
Melissa keeps one hand in the hot tub only because she likes the warmth.	Hand	Tub
All she did was sob in the shadow of the tree.	Shadow	Sob
The club in Denver purchased a car ramp for the parking lot.	Ramp	Club
The knob on the door in the old building was flat like a leaf.	Flat	Knob
The cub of the cat was too tired to play.	Cat	Cub
A web of a spider is its best trap to hunt its prey.	Trap	Web
Sergio always forgets to scrub around the gap in the floor.	Gap	Scrub
They organized a sports match with the tribe of Indians in the morning.	Match	Tribe
The hat of the old woman was covered with a blob of bird poop.	Hat	Blob
Bob often showed her a map of the subway.	Map	Bob
The babe in the cradle loves to play with the small pan in the kitchen.	Pan	Babe
I scan the crib in order to find little Lisa's favorite toy.	Scan	Crib
Her plan was to expose a boob on stage.	Plan	Boob
They needed a hammer to open the knob on the door.	Hammer	Knob
There was a shadow of the king's robe on the road.	Shadow	Robe
It is very difficult to dub a rap in a movie.	Rap	Dub
The teenager saw a total babe entering the tram to the center.	Tram	Babe
He knew the man loved to go to the club in London to perform.	Man	Club
There was a spider web on the old fur mantle in my mother's closet.	Mantle	Web
The guy lost his job of course, since he refused to remove his hat when serving customers.	Hat	Job
He knew it was a trap when Bob ordered him to lock the door.	Trap	Bob
He bruised his rib in June because he did not notice a gap in the street.	Gap	Rib
He put a cube of butter into the pan to melt.	Pan	Cube
The big bat from the cave bit the poor lion cub only out of fear.	Bat	Cub
I think I had a stub of a pencil in my drawer somewhere.	Had	Stub
The skateboarder preferred the tube over the ramp since it was much more exciting.	Ramp	Tube
The cat enjoys it when you rub its stomach.	Cat	Rub
The mob in Sicily is involved in the theft of gas from cars.	Gas	Mob

(*Appendices continue*)

Appendix A (*continued*)

Sentence	Target word /æ/	Target word /b/
He stepped out of the tub in order to observe the landscape through a window.	Landscape	Tub
I'm glad because I will never need to talk to this snob again.	Glad	Snob
The tribe occupying the strip of desert used sand to clean their pots.	Sand	Tribe
The officer showed the suspect the map in order to probe into what really occurred.	Map	Probe
The artist used his hand to remove a blob of paint.	Hand	Blob
The doctor needed to prescribe a number of drugs to the fat patient because he was diabetic.	Fat	Prescribe
It was an awful scam to try and sell the pub on the block which would be demolished.	Scam	Pub
Please turn on the lamp so I will be able to find my country on the globe in the corner.	Lamp	Globe
Since he expected his brother to rob a neighbor's flat he called the police.	Flat	Rob
The drummer of the band was told to scrub all of the dirt off of the stage.	Band	Scrub
The zoo keeper couldn't hear the panther's heart throb in his chest.	Panther	Throb
Sometimes he told her he would slap her if she would sob in public.	Slap	Sob
He made a fire using a match next to the crib in the nursery.	Match	Crib
She urged me to send a scan of the article on how to bribe a teacher.	Scan	Bribe
A large gas explosion occurred in the shop.	Gas	
Much whisky was drunk on the party instead of beer.		
Her cleavage revealed a perfect boob in a pretty red bra.		Boob
She took her boob out of her shirt in order to feed her baby.		Boob
I don't want you to probe into my business.		Probe
Shell wants to start searching for gas in the North Pole.	Gas	
There was a huge gap between his teeth.	Gap	
To bridge a gap, the directors paid the employees more.	Gap	
He spilled some wine on her dress.		
Gently insert the probe into the mouth when the patient is asleep.		Probe
The police arrested important members of the Chinese mob in their homes.		Mob
He waited desperately for the lord's sign because he did not know what to do.		
My mother uses the large pan to cook the meat.	Pan	
The pan caused a fire in the kitchen.	Pan	
I was glad the problem could be solved.	Glad	
The mob in New York is increasing its power in some neighborhoods.		Mob
The web of the tiny spider reached all the way to the other side of the porch.		Web
Don't get caught up in his web of lies again.		Web
All tennis balls were collected by the assistant.		
You should be glad he did not sue you.	Glad	
Desert sand is able to get inside your watch.	Sand	
In winter, the children go out to play in the snow.		
The red wine stain may disappear if you rub a bit of salt on it.		Rub
It would be great if you could rub a bit of sun block on my shoulders.		Rub
I'd love to own a house with a tub in the bedroom.		Tub
A lot of sand is used for the new garden.	Sand	
Thor is armed with a large hammer according to myth.	Hammer	
A yellow hammer is a kind of bird.	Hammer	
Ben is too young to be a lawyer.		
My uncle built a tub in his own yard.		Tub
The girl tried to bribe an officer in the parking lot.		Bribe
There was a big explosion in Syria because of terrorists.		
Bobby's right hand was scarred by the fire.	Hand	
Would you give me a hand with this ceiling?	Hand	
Nobody wears a hat these days.	Hat	
It seemed like she wanted to bribe a lawyer but I'm not sure.		Bribe
The patient's tongue was so swollen he needed to breathe through a tube in his throat.		Tube
There was a tube in there connecting the vessel to another one.		Tube
He kicked in the door with his heel.		
I take my hat off for this accomplishment.	Hat	
The bear walked right into the trap of the hunter.	Trap	
The book which stood on the shelf fell on the floor.		
The tribe of Indians dispersed in the woods to confuse the explorers.		Tribe
The spiral shaped scar on his shoulder meant he was part of the tribe of hunters.		Tribe
When I was a child there was a globe in my room with a light in it.		Globe
This useless trap did not kill the prey.	Trap	
Julia found a man on the street who was shot.	Man	

(*Appendices continue*)

Appendix A (*continued*)

Sentence	Target word /æ/	Target word /b/
The common man does not know much about neurobiology.	Man	
Emma is talking about the tigers she saw today on her trip to the zoo.		
Let's spin the globe in order to find a nice location for our spring trip.		Globe
Tomorrow in the spa we could use sea salt to scrub our skin.		Scrub
People whisper when they do not want to be heard.		
The criminal continued his scam on the street.	Scam	
A good scam deprives you of all your accessories.	Scam	
The mantle of the king was far too short.	Mantle	
The maid really needs to scrub all the restrooms before the guests arrive.		Scrub
She wore a gorgeous robe accompanied by the perfect pumps.		Robe
The bishop couldn't find his robe anywhere this morning.		Robe
It was too hot to sit outside to drink coffee.		
We covered the wounded soldier in a mantle of silk.	Mantle	
Our kitten resembled a panther when she hunted.	Panther	
Her colleague told her about their new boss.		
We should encourage them not to dub all French movies in order to boost learning.		Dub
You may know her voice because she is often paid to dub a movie.		Dub
You've been behaving like a snob all week.		Snob
A panther is hard to see in the dark.	Panther	
The biker used the ramp during the race.	Ramp	
Your ramp caught fire since it is made of wood.	Ramp	
The computer broke down because of a virus.		
The waiter serving us yesterday was a snob anyway.		Snob
The model would like the surgeon to remove a rib in order to look slimmer.		Rib
Everyone thought the white elegant outfit of the bride was beautiful.		
She used to slap her in the face.	Slap	
A hard slap is said to help you focus.	Slap	
Blake needed a CT-scan to find the tumor.	Scan	
A rib eye steak is what I love most in the world.		Rib
She urged the doctor to prescribe a pill from a different company.		Prescribe
This doctor does not prescribe any medicine for a cold with good reason.		Prescribe
Everyone listens attentively to the guide talking about the old church.		
You will need a scan of this document.	Scan	
All of the pirates sought the treasure map of the island.	Map	
My father always wants to be the best in chess.		
I will be fired next week but I didn't really like my job anyway.		Job
These days it is very difficult to find a job in my field.		Job
They heard a sudden throb a second before the motor died.		Throb
Only a map will show us the way out of this maze.	Map	
When I was young I had a teddy bear called Charly.	Had	
Did you say you had a house with a swimming pool?	Had	
Her father loves to take his luxurious car for a spin.		
This morning the wound started to throb a little.		Throb
The leopard left his cub alone to go on a hunt.		Cub
He went to the shop to buy a new book.		
My brother joined a band in order to become popular.	Band	
The lead singer in a band mostly determines its success.	Band	
Those two always try to match their outfits.	Match	
When you see a bear cub alone you need to be cautious because the mother will not be far.		Cub
To get in through the door you need to turn the knob on the other side.		Knob
I'm not sure how to open it, I don't find the knob on this window.		Knob
Marc goes to the therapist living in a nearby village.		
I could never match her chess skills.	Match	
The tram in the Hague makes me nauseous.	Tram	
Uncle Jerry needs a dentist because his tooth hurts.		
Today either Bob or Marc will win a bike in the tournament.		Bob
You did not mention Bob all of a sudden leaving his wife for another.		Bob
I'll be out partying in the club on Times Square tonight.		Club
A Belgian tram does not show its current location.	Tram	
Suzy got fat because she ate too much junk food.	Fat	
My neighbors' fat dog was regularly overfed.	Fat	

(*Appendices continue*)

Appendix A (*continued*)

Sentence	Target word /æ/	Target word /b/
He saw some money lying on the floor in front of him.		
Would you like to join the club of supporters?		Club
I never once witnessed someone rob a store.		Rob
The musical on Broadway was amazing.		
The lamp in the changing room was broken.	Lamp	
Jacob's chamber was lit only by the lamp on his desk.	Lamp	
Turkish people make flat bread in a great oven.	Flat	
Please don't tell me you gave him permission to rob all of them.		Rob
Where the little crib of the girl used to be, there was now a desk.		Crib
The rock star owns a crib in Florida the size of Disney World.		Crib
The concerned uncle comforted the toddler on his first day of school.		
People thought the world was flat in the middle ages.	Flat	
A scary bat rested on the ceiling of the cave.	Bat	
Her mother likes her new scarf very much.		
When his work is finished he goes to the pub in a village nearby.		Pub
I bought an old pub in need of remodeling.		Pub
Put a cube of ice on the wound to reduce the pain.		Cube
Billy could never hold the bat the right way.	Bat	
A lot of people rap, but only few possess skill.	Rap	
His rap music united two competing neighborhoods.	Rap	
Her sister is a successful model working in New York.		
Today in school we learned how to draw a cube in 3D.		Cube
Where the soldier's arm used to be there was only a stub of about 3 inches long now.		Stub
He interviewed the victim of the assault.		
I plan to finish my thesis next year.	Plan	
If everything goes according to plan, we should win the cup.	Plan	
Germany possesses a beautiful landscape, especially in the west.	Landscape	
It is impolite to toss your cigarette stub on the ground.		Stub
I saw the boy sob a long time when his mother said goodbye to him on his first day of school.		Sob
She did not want to show her tears but she could not help but sob all evening.		Sob
He tossed his broom on the floor because he was on strike.		
The artist got inspired by the landscape of Spain.	Landscape	
Their annoying cat always walks in our garden.	Cat	
I need some tissues to clean this mess.		
There was always a hot babe in the company of the movie star.		Babe
He hoped to hold the babe in his arms for the first time before he went to bed.		Babe
He found a mysterious blob of jelly in the dirty old fridge.		Blob
Their cat loved to hunt mice.	Cat	
He loves to lurk in the shadow of the school.	Shadow	
My cousin always tries to catch her shadow on the street.	Shadow	
The binder contained information on the budget.		
The blue blob over there is the lake she was talking about.		Blob

Note. Stimuli sentences and target words. When no target word is specified in the second or third column, the sentence in the first column is a filler sentence.

(*Appendices continue*)

Appendix B

Raw Values

Condition	F1 /æ/	F2 /æ/	Vowel duration /æ/	Vowel duration /b/	Closure duration /b/
Confederate absent, baseline	868 (113)	1873 (241)	124.58 (35.3)	131 (44)	60 (12)
Confederate absent, postexposure	844 (108)	1821 (221)	123.67 (38.36)	131 (48)	60 (16)
Confederate present, baseline	886 (86)	1820 (209)	125.66 (34.17)	134 (41)	63 (12)
Confederate present, postexposure	858 (83)	1802 (197)	121.00 (37.14)	131 (46)	62 (14)
Participant absent, baseline	751 (79)	1833 (162)	104.16 (28.93)	101 (40)	90 (27)
Participant absent, postexposure	760 (79)	1837 (160)	101.84 (27.48)	99 (37)	80 (23)
Participant present, baseline	749 (169)	1848 (221)	107.67 (32.54)	104 (38)	87 (30)
Participant present, postexposure	759 (112)	1826 (180)	106.34 (34.53)	104 (38)	78 (24)

Note. Raw values of F1 and F2 in Hz, vowel duration, and closure duration in milliseconds divided by speaker, social context, and block. The *SD*s are presented in parentheses.

Appendix C

Table C1
Estimates, SEs, t Values, and p Values for the Fixed and Random Effects of the Final General Linear Mixed Effect Model for the Dependent Measures of /b/, Participant Data Set

	Closure duration				Preceding vowel duration			
Fixed effects	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Intercept	.297	.018	16.140	<.001	.363	.021	17.396	<.001
Block	−.019	.005	−3.585	<.001	.020	.004	4.556	.000
Social context	−.018	.014	−1.331	.192	.005	.014	.372	.712
List 2	.028	.018	1.525	.138	−.032	.019	−1.656	.106
List 3	.020	.019	1.106	.277	−.039	.019	−2.077	.046
List 4	−.004	.019	−.200	.843	−.017	.019	−.873	.388
Block × Social Context	.004	.007	.566	.571	−.002	.006	−.354	.723
Random effects	Variance	*SD*			Variance	*SD*		
Sentence (intercept)	<.001	<.001			<.001	.005		
Participant (intercept)	.001	.035			.001	.036		
Word (intercept)	.004	.060			.006	.080		
Block	/	/						
List 2	/	/			.001	.029		
List 3	/	/			<.001	.019		
List 4	/	/			.001	.026		

(*Appendices continue*)

Table C2
Estimates, SEs, t Values, and p Values for the Fixed and Random Effects of the Final General Linear Mixed Effect Model for the Dependent Measures of /b/, Confederate Data Set

	Closure duration				Preceding vowel duration			
Fixed effects	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Intercept	.218	.009	23.025	<.001	.468	.018	26.416	<.001
Block	−.009	.008	−1.176	.240	−.008	.011	−.759	.449
Social context	.001	.006	.138	.890	−.007	.008	−.881	.379
Session	<.001	<.001	−.435	.664	−.001	<.001	−2.947	.003
List 2	−.003	.004	−.699	.485	.004	.007	.553	.581
List 3	−.007	.003	−2.565	.010	−.008	.004	−2.175	.030
List 4	−.011	.004	−2.771	.006	.004	.007	.502	.617
Block × Social Context	.006	.008	.720	.472	−.001	.011	−.077	.938
Block × Session	<.001	<.001	.903	.367	<.001	<.001	1.255	.210
Social Context × Session	<.001	<.001	1.719	.086	<.001	<.001	.618	.537
Block × Social Context × Session	<.001	<.001	−.764	.445	<.001	.001	−.843	.399
Random effects	Variance	*SD*			Variance	*SD*		
Sentence (intercept)	<.001	.022			.001	.039		
Social context	<.001	.007			<.001	.012		
Word (intercept)	.002	.041			.007	.085		
Social context	<.001	.009			<.001	.012		

Table C3
Estimates, SEs, t Values, and p Values for the Fixed and Random Effects of the Final General Linear Mixed Effect Model for the Dependent Measures of /æ/, Participant Data Set

	F1				F2				Duration			
Fixed effects	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Intercept	1.103	.018	62.632	<.001	.995	.009	111.811	<.001	.361	.022	16.603	<.001
Block	.013	.009	1.361	.183	−.002	.002	−.674	.501	.012	.008	1.615	.114
Social context	−.009	.016	−.547	.589	.002	.007	.279	.782	−.013	.016	−.780	.442
List 2	.006	.021	.280	.781	.015	.010	1.467	.152	−.035	.019	−1.811	.079
List 3	−.020	.021	−.962	.343	.002	.010	.206	.838	−.029	.019	−1.534	.135
List 4	−.018	.021	−.850	.402	.016	.010	1.592	.121	−.028	.019	−1.470	.151
Block × Social Context	.011	.013	.839	.407	−.001	.003	−.438	.661	<.001	.010	−.026	.979
Random effects	Variance	*SD*			Variance	*SD*			Variance	*SD*		
Sentence (intercept)	<.001	<.001			<.001	.005			<.001	.015		
Block	/	/							/	/		
Participant (intercept)	.002	.043			<.001	.020			.002	.045		
Block	.001	.029							<.001	.022		
Word (intercept)	.001	.023			<.001	.021			.006	.077		
Block	<.001	.013							<.001	.004		
Social context	<.001	.013										
List 2	/	/			/	/			/	/		
List 3	/	/			/	/			/	/		
List 4	/	/			/	/			/	/		

(*Appendices continue*)

Table C4

Estimates, SEs, t Values, and p Values for the Fixed and Random Effects of the Final General Linear Mixed Effect Model for the Dependent Measures of /æ/, Confederate Data Set

	F1				F2				Duration			
Fixed effects	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Intercept	1.124	.019	6.080	<.001	1.029	.011	89.764	<.001	.424	.021	19.789	<.001
Block	.009	.008	1.095	.275	−.016	.005	−2.949	.004	.003	.006	.434	.665
Social context	.010	.008	1.296	.197	−.023	.004	−6.145	<.001	−.027	.011	−2.481	.017
Session	.001	<.001	2.868	.004	−.001	<.001	−5.616	<.001	/	/	/	/
List 2	.011	.005	2.222	.032	−.002	.003	−.488	.627	−.018	.015	−1.169	.249
List 3	−.002	.003	−.461	.647	.001	.002	.365	.715	−.005	.014	−.328	.745
List 4	.018	.005	3.530	.001	−.002	.003	−.678	.499	−.028	.015	−1.853	.071
Block × Social Context	−.008	.009	−.890	.374	.008	.005	1.517	.130	.005	.006	.900	.369
Block × Session	−.001	<.001	−3.318	.001	<.001	<.001	.932	.352	/	/	/	/
Social Context × Session	−.001	<.001	−4.122	.000	.001	<.001	6.491	<.001	/	/	/	/
Block × Social Context × Session	.001	<.001	2.676	.008	<.001	<.001	−.357	.721	/	/	/	/
Random effects	Variance	*SD*			Variance	*SD*			Variance	*SD*		
Sentence (intercept)	.001	.024			<.001	.019	<.001	.019	.001	.026		
Social context	<.001	.010			<.001	.008	<.001	.008	.000	.006		
Participant (intercept)	/	/			/	/			.001	.027		
Word (intercept)	.009	.097			.003	.059			.009	.097		
Block	<.001	.015			<.001	.006			/	/		
Social context	.001	.027			<.001	.006			<.001	.014		
List 2	<.001	.014										
List 3	<.001	.009										
List 4	<.001	.015										

Received November 1, 2016
Revision received May 9, 2018
Accepted May 22, 2018 ■