

The perception of dental and alveolar stops among speakers of Irish English and American English¹

NIAMH E. KELLY

American University of Beirut

(Received 18 October 2015; revised 17 January 2017)

Most speakers of Irish English use a dental stop for words containing <th>, a sound that is generally pronounced as [θ] and [ð], in other varieties of English (Wells 1982; Ó hÚrdail 1997). Alveolar stops [t,d] and dental stops [t̪,d̪] are articulatorily and acoustically similar, and thus it is unusual for a language to use them contrastively (e.g. Ladefoged 2001). Despite this, Irish English contrasts them and speakers of this dialect have no trouble distinguishing them. This raises the question as to whether speakers of a dialect which does not use this contrast can distinguish them. To investigate this, speakers of Irish English and American English participated in an identification task involving words produced by an Irish English speaker. American English speakers had a high accuracy but did significantly worse than Irish English speakers, and both groups did significantly worse when the contrast was in final position than when it was in initial position. A small-scale production experiment examined words with this contrast and the vowel /a/, with the finding that for speakers of both dialects, the vowel is longer in words ending in <th> than <t>. The findings are discussed in the context of linguistic experience, and the effect of surrounding consonants on vowel duration.

Keywords: perception, stops, plosives, dental, alveolar, dialect variation

1 Introduction

Words represented in English orthography with <t> are pronounced with an alveolar stop in English, while words represented by <th> are pronounced with dental fricatives, [θ] and [ð], in most varieties of English (Wells 1982; Ó hÚrdail 1997). (Some research has shown that these fricatives are more consistently produced with an interdental gesture in American English and more often with the tongue behind the teeth in British English (Ladefoged & Maddieson 1996; McGuire & Babel 2012). To avoid overly specific descriptions here, the term ‘dental fricative’ is employed to refer to both kinds of production.) However, in some varieties, including Irish English (henceforth, IE) these dental fricatives correspond to surface dental stops. While most dialects of IE consistently differentiate alveolar and dental stops in production and perception, impressionistic investigations indicate that native speakers of American English have difficulty perceiving the difference. Indeed, Wells (1982: 429) even mentions that the contrast between the alveolar and dental stop ‘is not always obvious

¹ The idea for this project came about during a conversation with Mark Aronoff. I would also like to thank Miquel Simonet for his workshop on how to use PsychoPy. Special thanks to Deirdre Kelly, who recorded the stimuli, and to all participants in Ireland and at the University of Texas at Austin. Finally, special thanks are due to the editor and reviewers, whose comments greatly improved the manuscript.

to a non-Irish ear'. This apparent difference between the groups points to the effect of language experience on the perception of contrasts, which has been investigated from the perspective of segments (e.g. Flege 1984; Mann 1986; Kuhl 1991; McAllister *et al.* 2002; Best & Tyler 2007) and suprasegmentals (e.g. Xu *et al.* 2006; Francis *et al.* 2008; van Dommelen & Husby 2009; DiCanio 2012). Work on this issue has demonstrated that speakers are generally better at perceiving contrasts that are present in their native language than speakers who do not have this particular contrast in their language. The current investigation focuses not on a difference between two languages but a difference in the phonetic instantiation of a contrast between two dialects of the same language, namely Irish English and American English. This investigation adds to the literature on the relationship between production and perception. It also informs the literature on IE phonology, particularly from the less-investigated perspective of perception of language.

The following subsection (1.1) will describe the alveolar/dental contrast in IE and the corresponding contrast in American English (henceforth, AmE). Section 1.2 describes other languages that have an alveolar/dental contrast, and section 1.3 discusses work on the relationship between language experience and perception. Sections 2 to 4 then describe the experiment and results, and section 5 provides a discussion of the findings in the context of the relevant literature.

1.1 Coronals in Irish English and American English

Descriptions of Irish English note that for most speakers across the country, the segments generally written with <th> (which are dental fricatives [θ] and [ð] in most varieties of English) are pronounced as dental stops (Wells 1982). For instance, the word *tin* has an alveolar stop while the word *thin* has a dental stop for most speakers of IE (Ó hÚrdail 1997). Some IE speakers have a merger to an alveolar stop (Ó hÚrdail 1997), and this merger is usually associated with speakers from rural areas, as well as the south of the country, and inner-city Dublin.² Speakers from the west (where the participants in the current investigation are from) are known to have the alveolar/dental contrast (Hickey 2008). Hickey (2008) posits that the fricatives present in British English were replaced with dental stops by native speakers of Irish when they began to use English, since this was the closest equivalent in their first language.

It should also be mentioned here that a widespread phenomenon in IE is what is known as t-lenition (Hickey 1984; Pandeli *et al.* 1997; Kallen 2005; Hickey 2008). Words in which the voiceless alveolar stop /t/ follows a stressed vowel (*bat*, *about*, *butter*) and in words where /t/ is the coda of a weak syllable that follows a stressed

² This merger is often present in imitations of IE by speakers of other varieties of English. In reality, more IE speakers have an alveolar/dental contrast than a merger, so the perception that all IE speakers have a merger is either based on stereotyped versions of the IE accent, or due to the fact that speakers of other dialects actually do not perceive the alveolar/dental contrast.

syllable (such as *bucket*), the alveolar stop is often lenited, and produced as what has been described as a ‘slit-fricative’ (Kallen 2013). This is not otherwise relevant to the current study, so in order to focus on the issue in question, it was ensured that this sound did not occur among the stimuli used in the current perception experiment. This was done by choosing stimuli that had the alveolar stop rather than the slit-fricative. (This was done on an auditory basis by the author, since the slit-fricative is highly salient.)

General American English does not generally use the dental stop for words written with <th>; instead, it has a dental fricative (usually interdental) (Aschmann 2015), except for some northern dialects and AAVE (Edwards 2008; Gordon 2008; Nagy & Roberts 2008; Thomas 2008). This fricative is sometimes produced as a dental stop in running speech (Lavoie 2002). Jongman *et al.* (1985) also note that (American) English speakers sometimes produce alveolar stops as dental stops, presumably because since AmE does not contrast these as phonemes, speakers’ productions of alveolars are freer to move into the dental space. (This is different from Malayalam, where the acoustic properties (in this case, burst amplitude) are kept consistent for either alveolar or dental stops, because they are separate phonemes in this language.) A palatographic study with speakers of AmE from the Pacific coast found that across 20 speakers’ productions of words with the alveolar stop such as *tap* and *pat*, 7.7 per cent of articulations of <t> were dental, while 2.5 per cent of articulations of <d> were dental (Dart 1991, 1998). As such, AmE speakers have some experience with the dental stop, but only as allophones.³ Furthermore, their experience with the dental stop allows it to conceivably be perceived as either an alveolar stop or a dental fricative, although the perception of it as an alveolar stop could be argued to be more likely based on the small acoustic difference between an alveolar stop and a dental stop, and the apparent higher frequency with which the dental stop is used as an allophone of the alveolar stop. It therefore remains to be determined whether they will be able to consistently categorise it as one of these (the dental fricative, in this case). The current investigation examines how well they can distinguish dental vs alveolar stops when the two are being specifically contrasted.

Likely due to limited experience with the dental stop, AmE speakers appear to have difficulty distinguishing the alveolar stop from the dental stop, something that IE speakers are impressionistically found to do with a high level of accuracy. Since IE has a dental stop where AmE has a dental fricative, the difference between the dialects is not in the distribution of these two sounds but in the phonetic realisation of the same phoneme. Under this analysis, these dialects are seen as having the same underlying system, but differences in the realisation of certain phonemes. This panlectal/polylectal approach is the most straightforward analysis, given the historical development of American English, which came from varieties of English from Britain and Ireland, as well as the consequential similarities and intelligibility between the two dialects, and the fact that the distribution of dental fricatives in AmE lines up with the distribution

³ The author would like to thank an anonymous reviewer for this insight.

of dental stops in IE. Essentially, the dental stop is a phoneme for IE speakers, while it is a possible allophone for AmE speakers.

1.2 *Languages with an alveolar/dental contrast*

Although dental and alveolar stops are articulatorily and acoustically similar, there are other languages that contrast them. Cajun English is another variety of English known to exhibit dental stops rather than fricatives (Dubois & Horvath 1998). Languages other than English that have an alveolar/dental contrast include Kimvita Swahili, spoken in Mombasa (Hayward *et al.* 1989), Dinka, spoken in South Sudan (Remijsen & Manyang 2009), some Dravidian languages, such as Malayalam (Haowen 2010), as well as some languages of Australia, such as Yolngu (Morphy 1983), Nunggubuyu (Hughes & Leeding 1971), Yanyuwa (Huttar & Kirton 1981) and Ngiyambaa (Donaldson 1980). Therefore, the contrast is clearly sufficiently perceptible to be used, although it is typologically unusual.

1.3 *Linguistic experience and perception*

Research into the effect of linguistic experience has demonstrated that speakers are generally better at distinguishing contrasts that occur in their own language than those that do not (e.g. Flege 1984, 1995; Polka 1989, 1991; Strange 1992, 1995; Best & Strange 1992; Best 1994, 1995; Best & Tyler 2007). Work by Werker & Tees (1984) and Kuhl *et al.* (1992) has shown that linguistic experience affects how infants process vowels and consonants within the first year of life. This effect of linguistic experience is also found in adults. Werker & Tees (1984) found that native adult speakers of English had difficulty distinguishing velar stops from uvular stops, since English does not have uvular stops, while speakers of Thompson Salish, which has this contrast, had no difficulty distinguishing the sounds. Likewise, speakers of Japanese often have difficulty distinguishing English /l/ and /r/ because these are not separate phonemes in Japanese (Mann 1986).

On the other hand, contrasts that are in one's native language can also facilitate learning of a similar contrast in another language. McAllister *et al.* (2002) found that native speakers of Estonian were more proficient at acquiring the segment length distinction in Swedish than either English or Spanish speakers, due to the fact that Estonian also has a length distinction. Similarly, speakers of Spanish and Russian tend to use duration rather than spectral cues to distinguish English tense and lax vowels, since their native languages use duration more than spectral differences (Kondaurova & Francis 2008).

While no work appears to have examined perception of the contrast between dental and alveolar stops by listeners who do not have it in their native dialect, one study found that speakers of AmE had difficulty perceiving the contrast between Hindi dental and retroflex stops (Pruitt *et al.* 2006). In their study, AmE participants identified dental and retroflex stops with only 50–60 per cent accuracy.

The authors interpret this as the AmE speakers categorising both Hindi stops as alveolar stops.

Similar patterns have been found for suprasegmentals. For example, Mandarin speakers were found to categorically perceive F0 contours that were along a continuum from level to rising, while English speakers did not show categorical perception, due to the fact that Mandarin has lexical tone while English does not (Xu *et al.* 2006). The language experience of Mandarin speakers caused them to perceive the tonal contours linguistically and thus divide them into two categories. Research on the perception of pitch in Yucatec Maya has shown a correlation between what speakers use in tone production and what cues they use in perception (Frazier 2009).

Overall, these results show an effect of linguistic experience on the perception of sound contrasts. Based on this research, an identification experiment was set up to compare IE and AmE speakers in terms of their perception of the alveolar/dental contrast. It was hypothesised that since IE speakers have this phonemic contrast while AmE speakers do not, IE speakers would be more accurate and have a faster reaction time than AmE speakers at identifying words that differ only in whether they have an alveolar or dental stop.

2 Method

2.1 Stimuli

Stimuli were recorded by a female native speaker of Irish English, aged 30, who produced the words in isolation. All words appear in the appendix. Stimuli consisted of 24 words in total, that is, 12 minimal pairs containing the dental/alveolar contrast in either word-initial or word-final position. The breakdown of position and voicing was as follows:

- Initial voiceless stops: 3 pairs
- Final voiceless stops: 3 pairs
- Initial voiced stops: 3 pairs
- Final voiced stops: 3 pairs

Stimuli were limited to 3 pairs for each condition to ensure that there were no extra repetitions of some target words. The pairs, for example *boat/both*, were then manipulated when necessary so they had the same pitch, intensity and vowel duration. (A noted vowel duration difference in some pairs will be discussed in [section 4](#).)

2.2 Participants

Participants were 10 (7 females, 3 males) adult native speakers of Irish English and 10 (1 male, 9 females) of American English, aged 18–30. The IE participants were from Galway City, on the west coast of Ireland, and the AmE participants were all speakers

of General American English (that is, none had impressionistically strong Southern accents, for example), living in Austin, Texas. None was proficient in any language other than English. Participants were recruited by word of mouth. All participants completed a consent form and a language background questionnaire.

2.3 Procedure

Participants took part in a two-alternative forced choice task. The experiments were presented in PsychoPy (Peirce 2007). Participants were seated in a quiet room in front of the experimenter's laptop.

In the identification task, the participant heard one word at a time and was asked to choose which of two words on-screen they had heard. The word with the alveolar stop (written with <t> or <d>, such as *den*) always appeared on the left of the screen and that with the dental stop (<th>, such as *then*) always appeared on the right. Each word was repeated eight times, meaning that each condition contained 24 tokens. That is, alveolar initial voiceless had 24, dental initial voiceless had 24, and so on. These were presented in eight blocks of 24 (192 in total), so one example of each token appeared in every block. There were two initial practice trials using words that were not part of the experiment. The participant's response for one trial initiated the following trial. There was a blink of the screen to indicate that a new trial had begun. If they did not respond within 1.5 seconds after the target word, it moved to the next trial.

After completing the perception experiment, participants were asked to read out the target words while being recorded. This was to examine their productions to determine whether the IE speakers had the alveolar/dental contrast or a merger, and to ensure that the AmE participants had a contrast between an alveolar stop and a dental fricative. The target words from the perception experiment were presented in list form on three powerpoint slides. All the words appeared once per slide, in different orders. Participants read out the words while being recorded through the program Audacity. In total, the perception experiment took about 15 minutes and the production experiment approximately 2 minutes.

2.4 Measurements

2.4.1 Perception

Identification results were measured by accuracy. These were subjected to a logistic regression using the *glmer* function in R (R Development Core Team 2008), where the dependent variable was accuracy (1 or 0). Likelihood ratio tests using the *anova* function determined which model was a best fit for the data. Reaction times for correct responses were also analysed, using a linear regression in the *lmer* function in R.

2.4.2 Production

Participants' production of the target words was examined impressionistically by two trained phoneticians to determine whether they all produced a contrast in these words

and if so, what sounds they used. The AmE speakers all contrasted alveolar stops with dental fricatives, as expected. All IE speakers had the alveolar/dental contrast. (Some IE speakers produced the ‘slit-fricative’ instead of the alveolar stop in final position in some words.)

A further analysis focused on vowel duration in two of the minimal pairs. It was noted that the original stimuli (produced by the IE speaker) that contained a final voiceless dental stop and an open vowel (such as *bath*, *path*) were produced with a longer vowel than their alveolar stop counterparts (*bat*, *pat*). (The other word pairs did not reveal any such difference.) For that speaker, those stimuli with a final alveolar stop had an average vowel duration of about 100 milliseconds while those with a final dental stop were about 200 milliseconds. For this reason, these stimuli were not used in the perception experiment, but participants were asked to produce them, in order to allow for an analysis of vowel duration.

Finally, to provide further acoustic information on the difference between the dental and alveolar stops, the IE participants’ production of voiceless initial stops was examined in terms of voice onset time (VOT). Here, VOT was measured as the time (in milliseconds) from the initial burst release, marked by a sudden onset of noise in the waveform, to the onset of voicing, as described in Lisker & Abramson (1970: 9–10): ‘the interval between the release burst and the onset of laryngeal pulsing’. Previous work (Lisker & Abramson 1964) has shown different amounts of VOT of initial voiceless stops based on place of articulation, thus explaining the decision to examine this measure in the current study. This work found that in English, /p t k/ had VOTs of 60, 70 and 80 milliseconds, respectively. For languages as diverse as, for example, English, Hungarian and Korean, they found that VOT increased as the place of articulation became further back in the vocal tract. Cho & Ladefoged (1999), referring to Maddieson (1997), explain that this pattern may be caused by the fact that for a velar stop, there is a larger body of air in the mouth than there is for an alveolar or bilabial stop, so this takes longer to be released. Similarly, since the space behind a velar closure is smaller than for a bilabial or alveolar stop, the air pressure behind the closure is higher, so it takes longer for this high pressure to fall. If this pattern holds, in the current study it would be expected that dental stops have a shorter VOT or less aspiration than alveolar stops. The production results will be discussed in section 4.

3 Perception results

3.1 Results: accuracy

Overall correct responses by participant are shown in figures 1 and 2. A high level of accuracy can be seen for both groups.

Logistic regression models were built up and compared using likelihood ratio tests. These were built by adding terms one by one and comparing models to determine

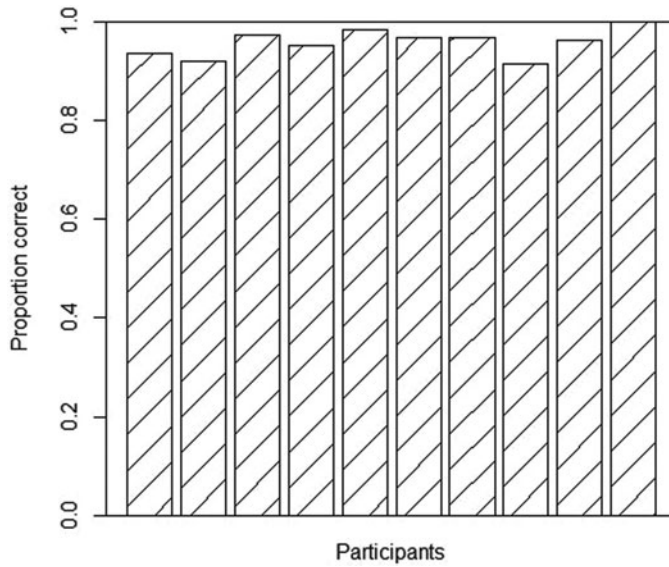


Figure 1. Identification results for IE speakers

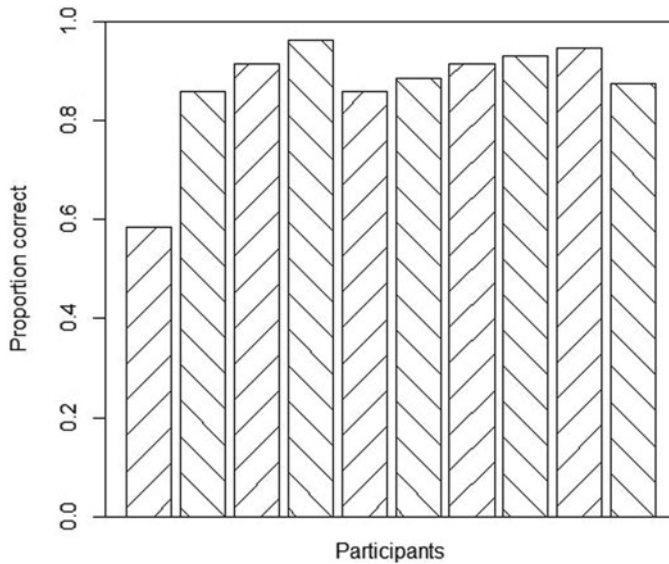


Figure 2. Identification results for AmE speakers

which model best explained the data. Terms tested were: COUNTRY (IE or AmE), BLOCK (to test for a training effect), POSITION (whether the contrast was initial or final in the word) and VOICING (/t/ vs /d/). Participant was included as a random effect. While a model including only COUNTRY as a predictor variable was better than

Table 1. *Statistical results for the best model*

| Variable | Coef. | z-value | p-value |
|-----------|-------|---------|---------------|
| Intercept | 3 | 11.8 | $p < 0.001^*$ |
| Position | 0.8 | 6 | $p < 0.001^*$ |
| Country | -1.2 | -3.5 | $p < 0.001^*$ |

Note: The reference level for position is 'initial' and country is 'IE'.

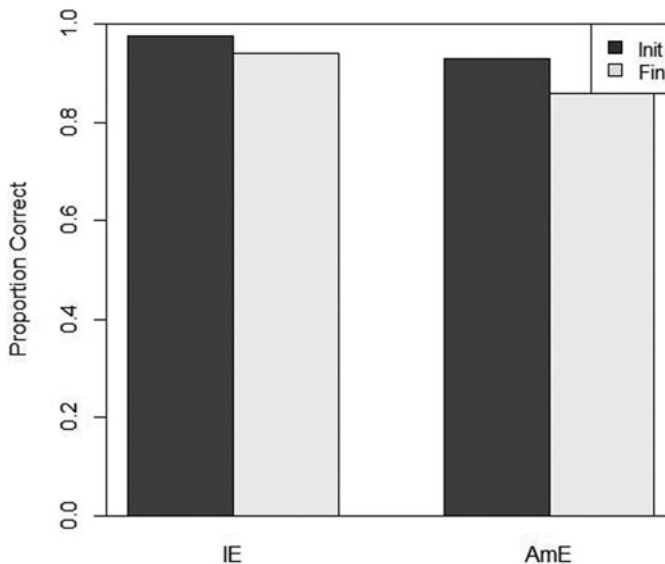


Figure 3. Average results pooled across participants, by country and position

the null model ($p = 0.001$), and a model including only POSITION was significantly better than the null model ($p < 0.001$), the best model to explain the data was one that included both COUNTRY and POSITION ($p = 0.001$) (table 1): `glmer(accuracy ~ Country + Position + (1|Participant), family=binomial)`. These results showed that, as in figure 3, while both groups of participants had high accuracy when the contrast was in initial position (for example, *tin* vs *thin*), the AmE group did worse when the contrast was in final position (for example, *tent* vs *tenth*) (average accuracy 86 per cent).

Pairwise post-hoc tests using the *lsmeans* package gave the results shown in table 2. It was found that both groups of participants did significantly better in the initial condition than in the final condition. The groups were significantly different from one another in both conditions also.

Table 2. *Pairwise statistical results by country and position*

| Pair | p-value |
|-------------------------|----------------|
| IE initial, IE final | $p < 0.001^*$ |
| AmE initial, AmE final | $p < 0.001^*$ |
| IE initial, AmE initial | $p = 0.0028^*$ |
| IE final, AmE final | $p = 0.0028^*$ |

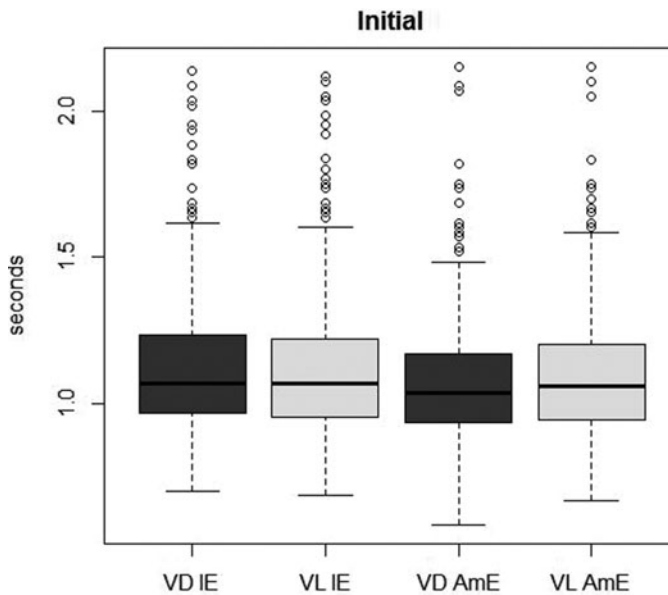


Figure 4. Reaction time in initial position, for voiced (VD) and voiceless (VL) segments by country

3.2 Results: reaction time

For all the correct answers, reaction time was examined. Since the listeners had to wait for the end of the word to respond when the contrast was in final position, the initial and final position conditions were examined separately. Figures 4 and 5 show the results for initial and final position, respectively.

A linear regression was run on the reaction time for correct answers, using the *lmer* function in R. As mentioned, initial and final position were examined separately. Likelihood ratio tests were used to determine which model best explained the data. As before, the models were built up by term (COUNTRY and VOICING) and then compared to see which model was the best fit. For initial position, no model fitted the data better than the null model, indicating that there was no significant effect of COUNTRY or VOICING on reaction time. For final position, these tests showed that a model with an

Table 3. *Statistical results for the best reaction time model for final position*

| Variable | Coef. | t-value | p-value |
|-------------|-------|---------|---------------|
| Intercept | 1.3 | 40.3 | $p < 0.001^*$ |
| Country | -0.09 | -1.9 | $p = 0.073$ |
| Voicing | -0.1 | -7.2 | $p < 0.001^*$ |
| Interaction | 0.08 | 3.8 | $p < 0.001^*$ |

Note: The reference level for voicing is 'voiced' and country is 'IE'.

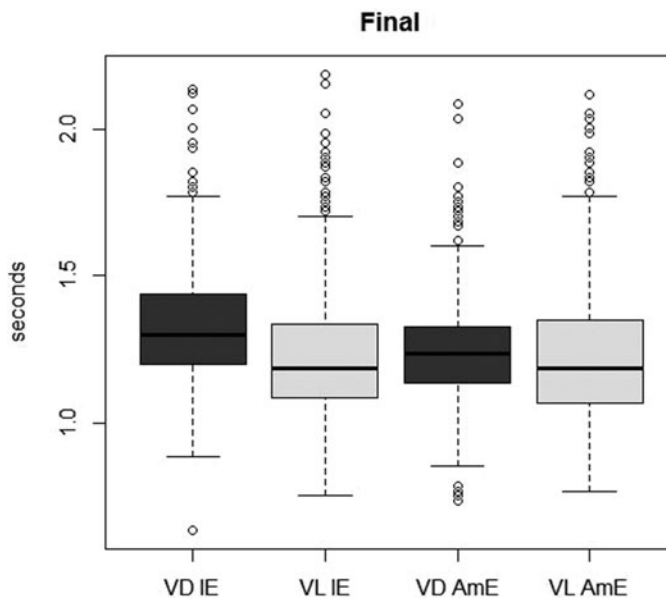


Figure 5. Reaction time in final position, for voiced (VD) and voiceless (VL) segments by country

interaction of COUNTRY by VOICING was the best fit (R code: `lmer(accuracy ~ Country * Voicing + (1|Participant), data=final)`). These results are shown in table 3. Pairwise comparisons using *lsmeans* showed that when the contrast was in final position, the AmE participants did not differ based on voicing, and did not differ from the IE participants, but the IE participants responded significantly faster in the voiceless condition than the voiced condition.

4 Production results

The first part of the analysis of the audio recordings focused on the open vowel pairs, since a vowel duration difference was noted for the original IE speaker. Recordings

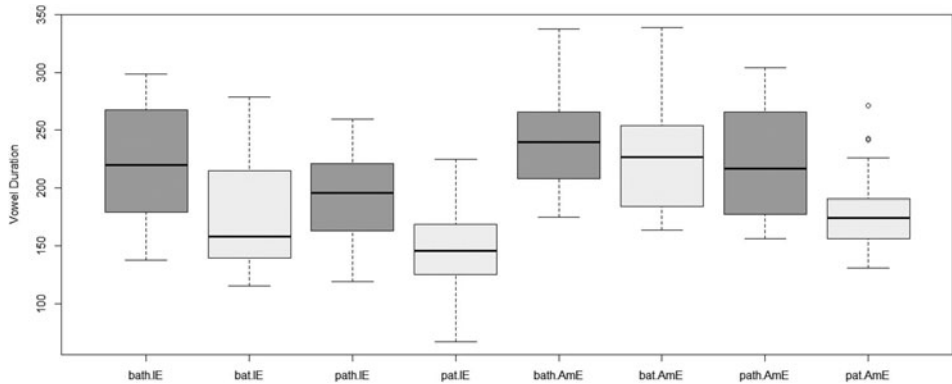


Figure 6. Average vowel duration pooled across participants, by country, final sound (<t> or <th>) and initial stop (/b/ or /p/). IE speakers are the four boxes on the left, AmE speakers the four boxes on the right. The darker boxes are those ending in <th> (*bath/path*) and the lighter boxes ending in <t> (*bat/pat*).

of 9 IE participants (7 females, 2 males) and 8 of the AmE participants (7 females, 1 male) were analysed.⁴ The vowels were manually labelled and then a Praat script (Boersma & Weenink 2011) was used to measure vowel length in these pairs. As mentioned above, the words were read out, in isolation, from Powerpoint slides. Vowel duration was measured as beginning with voicing pulses and ending where the spectrogram got lighter and the waveform less uniform, signalling the end of voicing. This meant that for words beginning with <p>, the aspiration was not included as part of the vowel. For the IE speakers' words with a final voiceless stop, pre-aspiration occurred in most instances. This was included as part of the vowel. There were three repetitions of each token (*bat*, *pat*, *bath*, *path*) per speaker, giving $3 \times 4 \times 17 = 204$ tokens in total. Two tokens were discarded due to noise, leaving 202 tokens for analysis.

Figure 6 shows average vowel duration for *bat/pat* versus *bath/path*, by speaker country. It can be seen that for both groups, *bath* had the longest vowel, and *pat* the shortest. For the IE speakers, the average duration of *bath* is longer than that of *bat* (first two boxes), while for AmE speakers the difference, if any, is much smaller. For both groups, *path* appears longer than *pat*.

Figures 7–14 show spectrograms for all four words for one female IE speaker and one female AmE speaker.

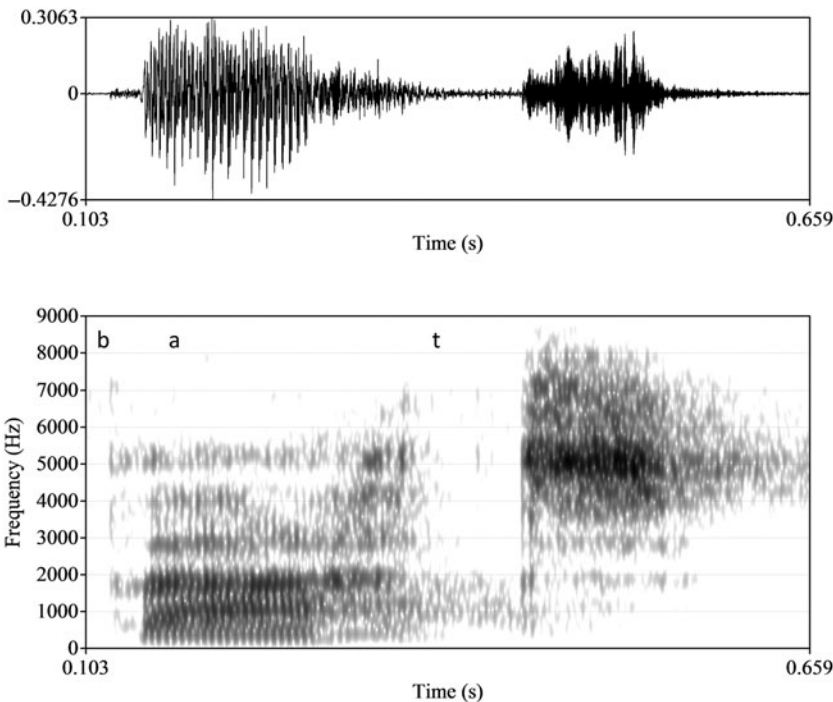
These measurements were subjected to a linear regression, with COUNTRY, FINAL (<t> or <th>) and INITIAL (whether the initial consonant was /b/ or /p/) as possible predictor variables. Models were built up term by term, as before, and compared using likelihood ratio tests (*anova*) to determine which model best explained the data. Speaker was included as a random factor. The model that best explained the data was

⁴ Only 8 of the 10 AmE participants were recorded, due to timing issues.

Table 4. *Statistical results for the best duration model*

| Variable | Coef. | t-value | p-value |
|-----------|-------|---------|---------------|
| Intercept | 218.9 | 18.8 | $p < 0.001^*$ |
| Initial | -31.9 | -8 | $p < 0.001^*$ |
| Final | -38 | -9.5 | $p < 0.001^*$ |
| Country | 33.5 | 2 | $p = 0.065$ |

Note: The reference level for Initial is /b/, for Final is <th>, and country is IE.

Figure 7. Spectrogram of *bat* for one IE speaker

one that included all three terms⁵ (R code: `lmer(duration ~ Initial + Final + Country + (1|Speaker))`), and the summary of this model is shown in table 4.

Pairwise tests using the *lsmeans* function showed that for IE speakers, there were significant differences between *bath* and *bat*, between *path* and *pat*, and also between *bath* and *path*, and between *bat* and *pat*. Parallel results were found for AmE speakers. This shows that both final sound and initial sound had significant effects on vowel duration, where the vowel in words ending in <th> was longer than that of words

⁵ A model including COUNTRY, FINAL and INITIAL was slightly but significantly better than one just including FINAL and INITIAL, $p = 0.047$.

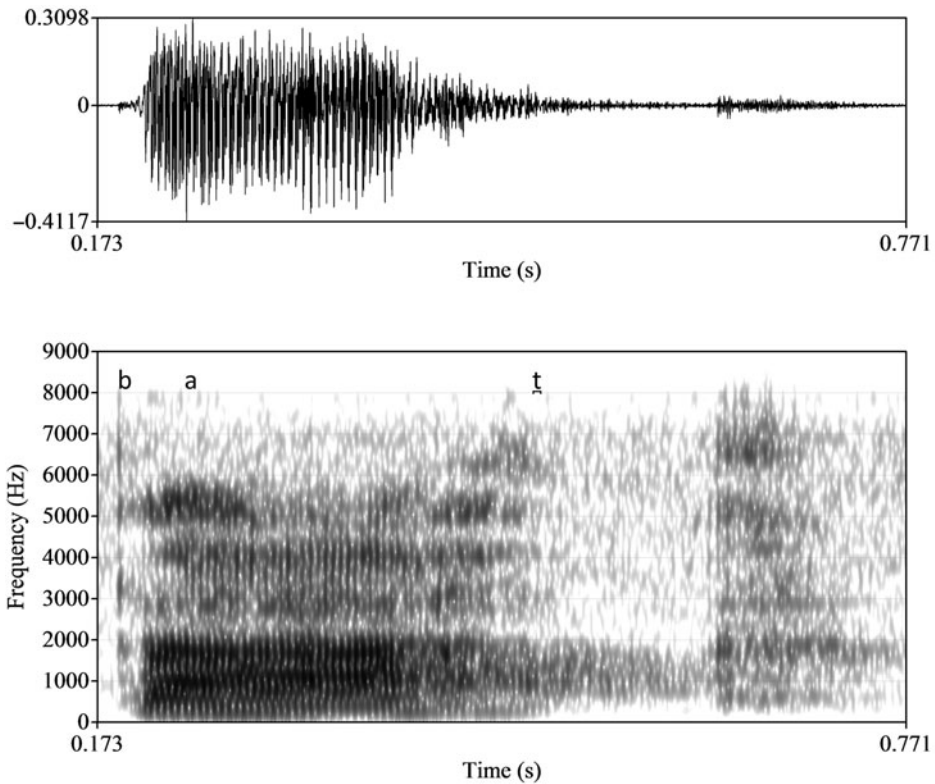
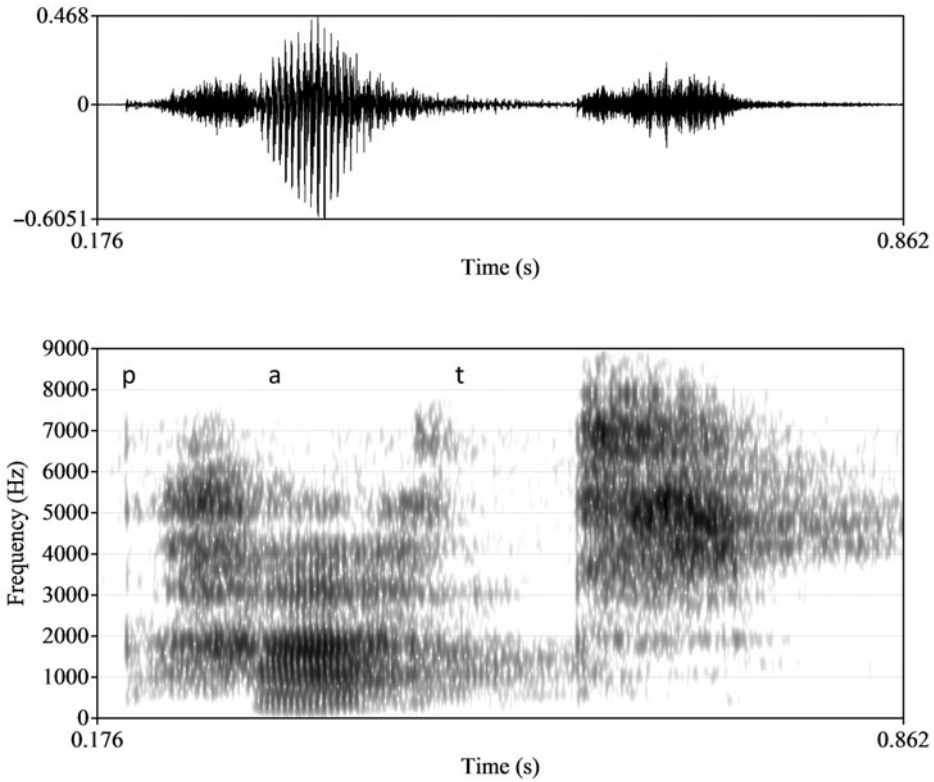


Figure 8. Spectrogram of *bath* for one IE speaker

ending in <t> for both groups, and the vowel in words beginning with /b/ was longer than that of words beginning with /p/. There were no significant differences between the IE and AmE speakers for any word. As mentioned above, the IE speakers generally had pre-aspiration in words with final voiceless stops (average 72 milliseconds, both alveolar and dental), and this was included as part of the vowel. This did not occur in words with final voiced stops, or for any of the AmE speakers. Since this occurred for both alveolar and dental stops, it does not account for the difference in vowel duration between these two. Table 5 shows the pairwise results.

These results further show that the vowel duration difference between *bath* and *bat* was not just present for IE speakers, but also for AmE speakers. For IE speakers, the final segment difference between the two words is in place of articulation, between an alveolar and dental stop, while for the AmE speakers it is a difference in both place and manner of articulation, between an alveolar stop and a dental fricative.

The amount of VOT when the voiceless alveolar and dental stops were word-initial was measured, with the finding that for all speakers, the alveolar stop <t> had longer VOT (by 12–22ms) than the dental stop <th> before the same vowel. The pooled

Figure 9. Spectrogram of *pat* for one IE speakerTable 5. *Pairwise statistical results for duration*

| Pair | p-value |
|-----------------------------------|---------------|
| IE <i>bath</i> , IE <i>path</i> | $p < 0.001^*$ |
| IE <i>bath</i> , IE <i>bat</i> | $p < 0.001^*$ |
| IE <i>path</i> , IE <i>pat</i> | $p < 0.001^*$ |
| IE <i>bat</i> , IE <i>pat</i> | $p < 0.001^*$ |
| AmE <i>bath</i> , AmE <i>path</i> | $p < 0.001^*$ |
| AmE <i>bath</i> , AmE <i>bat</i> | $p < 0.001^*$ |
| AmE <i>path</i> , AmE <i>pat</i> | $p < 0.001^*$ |
| AmE <i>bat</i> , AmE <i>pat</i> | $p < 0.001^*$ |
| IE <i>bath</i> , AmE <i>bath</i> | $p = 0.5$ |
| IE <i>path</i> , AmE <i>path</i> | $p = 0.5$ |
| IE <i>bat</i> , AmE <i>bat</i> | $p = 0.5$ |
| IE <i>pat</i> , AmE <i>pat</i> | $p = 0.5$ |

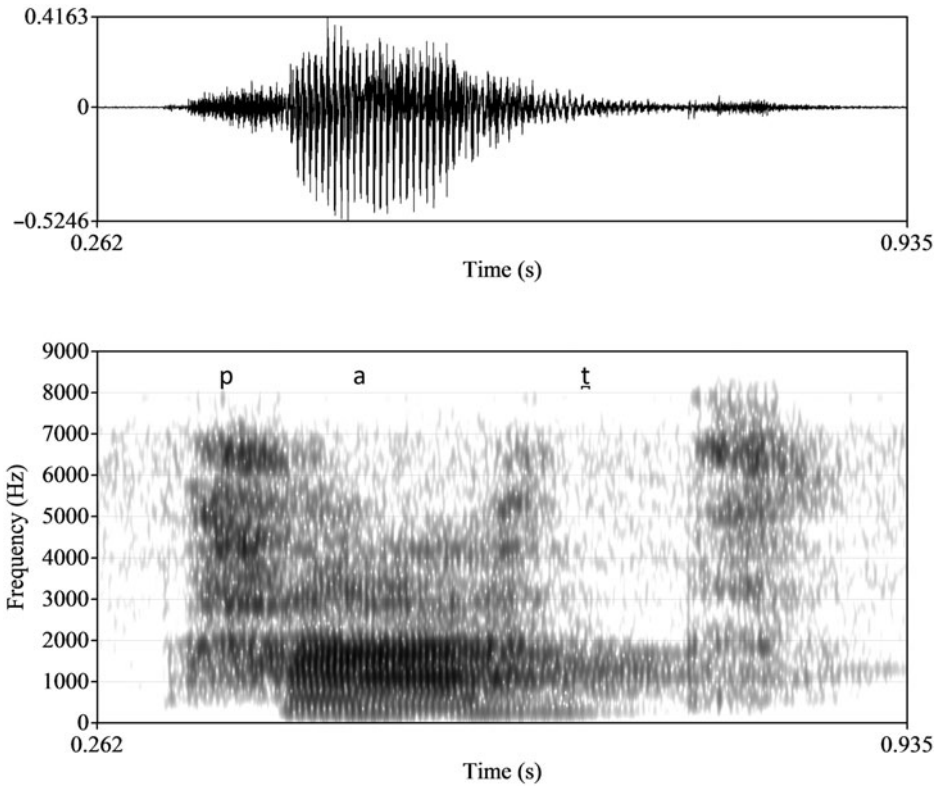


Figure 10. Spectrogram of *path* for one IE speaker

Table 6. *VOT (milliseconds) pooled across IE speakers*

| | tanks | thanks | tin | thin | taught | thought |
|-----|-------|--------|-----|------|--------|---------|
| VOT | 97 | 76 | 106 | 94 | 111 | 89 |

results voiceless for initial dental and alveolar stops produced by IE speakers are shown in [table 6](#).

These results will be discussed in [section 5.2](#).

5 Discussion

5.1 Perception

The perception results consisted of two parts: accuracy and reaction time.

Overall, the AmE participants did better than was expected at the identification task. The hypothesis was that due to phonetic similarity, AmE speakers would map alveolar

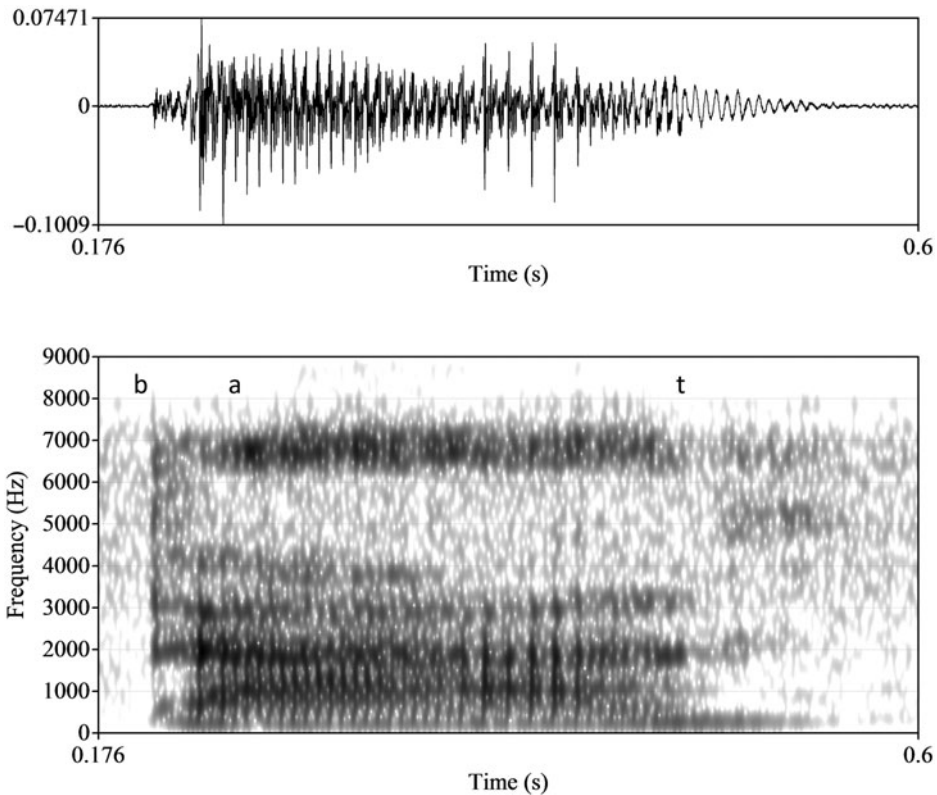


Figure 11. Spectrogram of *bat* for one AmE speaker

and dental stops onto the same (alveolar) category (Flege 1995), but the results indicate that this was not the case.

Furthermore, no effect of BLOCK was found, indicating that there was no training during the experiment, and participants did well from the beginning. It may be that since the words were produced in isolation, listeners were able to attend more to the contrast than they would in running speech. This suggests that the AmE participants were perhaps focusing on the acoustic properties of the relevant stops, rather than processing them as words. Similar to what Dupoux *et al.* (1997) found with French speakers processing word accent, it may be that the AmE participants here were able to perform well at the task because their responses were required immediately. If the task were made more difficult by increasing the memory load of participants, it is possible that AmE participants would not be able to encode the information at a higher level. While the current results show that language experience has an influence, by the result that IE speakers were more accurate than AmE speakers, the fact that these varieties are two versions of the same language may have facilitated the AmE speakers in the task. Much research has focused on comparing speakers of different languages rather than different dialects of the same language. It may be that the AmE speakers

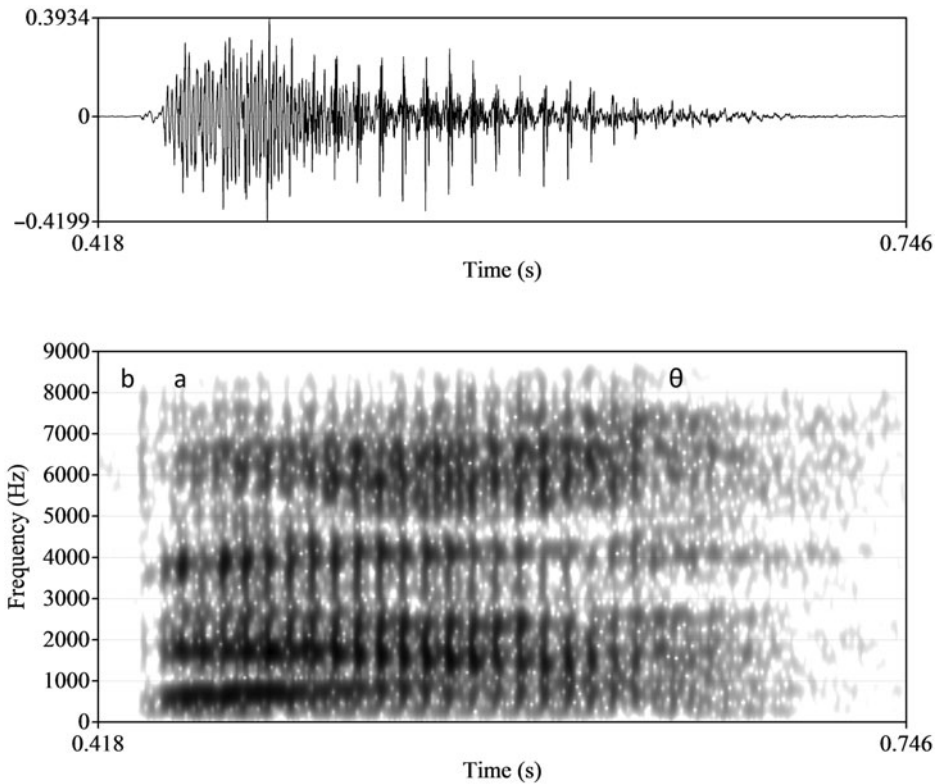


Figure 12. Spectrogram of *bath* for one AmE speaker

were aided by understanding the lexical items. Perhaps an experiment using nonce words or just auditory materials would show different results. However, even using nonce words presented in English orthography would drive the listeners towards some form of linguistic processing. For this reason, a setup with only auditory presentation of materials may be preferable.

As noted in the introduction, AmE speakers may have the dental stop allophonically. Therefore, the experiment required these listeners to identify an orthographic word such as <both> as having a final dental stop, meaning that the listeners may in fact be using their awareness of how an Irish person would say the word.⁶ An experiment involving just sounds rather than whole words could perhaps determine whether the AmE participants are in fact doing this. While the IE participants were asked to distinguish contrasting sounds or separate phonemes, the AmE participants had the more challenging task of distinguishing possible allophonic variants of the same phoneme (for these participants, both [t] and [t̪] could be realisations of /t/). Their success at doing this indicates that perhaps their possible exposure to the dental

⁶ Thanks to an anonymous reviewer for this point.

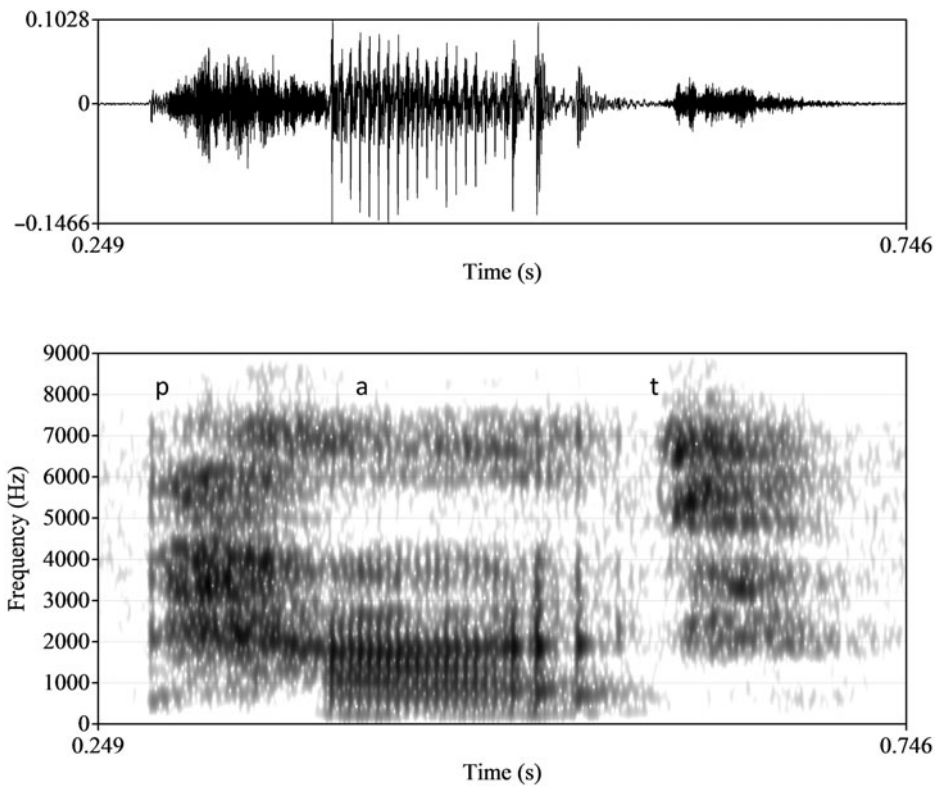


Figure 13. Spectrogram of *pat* for one AmE speaker

stop (even just as an allophone) assisted their processing of this sound. Especially comparing the current results to those of Werker & Tees (1984), where native speakers of English were poor at distinguishing velar stops from uvular stops, it appears that in the current study, AmE participants may have done relatively well because they may have experience with the dental stop, even though only allophonically.

It was also found that both AmE and IE participants did significantly worse when the contrast was in final position than when it was in initial position. However, accuracy in the final condition was still high. It is unclear why the IE speakers had a significantly faster reaction time when the segments were voiceless, but only in final position, while there was no such difference found for the AmE participants.

5.2 Production

The audio recordings showed that both IE and AmE speakers produced words ending in <th> (*bath*, *path*) with significantly longer vowels than words ending in <t> (*bat*, *pat*). Words beginning with /b/ were also longer than their counterparts beginning with /p/.

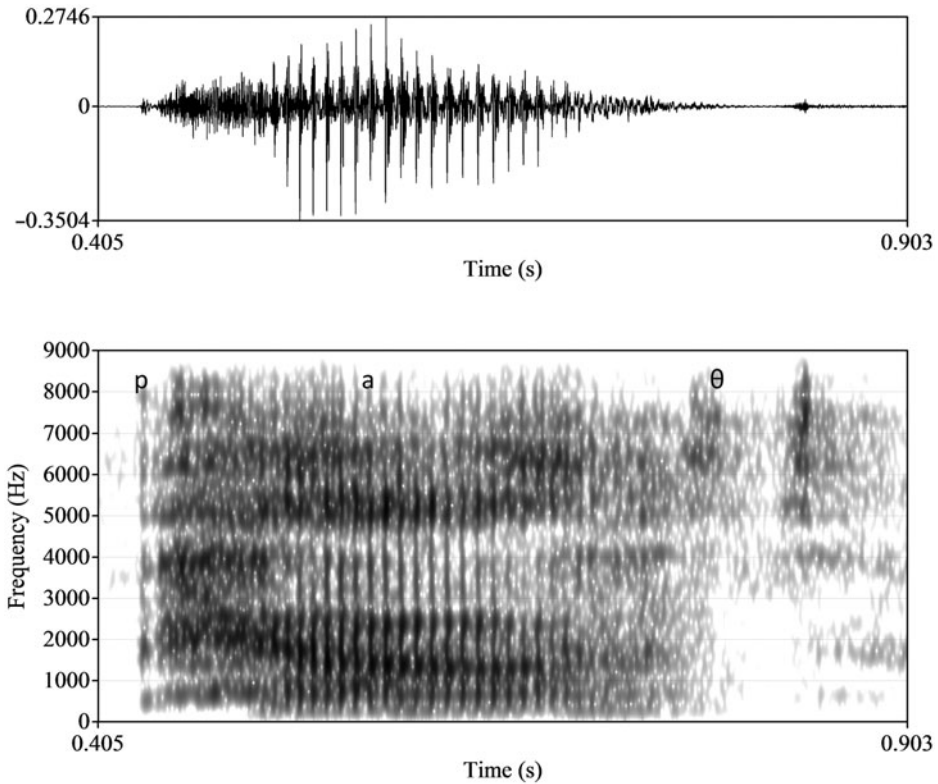


Figure 14. Spectrogram of *path* for one AmE speaker

While conclusions based on these results are preliminary due to the words being produced in isolation, the results indicate that there is indeed a vowel duration difference for both groups of speakers based on the final segment. The phonetic explanation of these results differs based on the group's production of the final sound. While it is well known that vowels are longer before voiced than voiceless consonants (House & Fairbanks 1953; Peterson & Lehiste 1960; House 1961), the difference in the current study was in place and/or manner, but not voicing. With regard to the AmE speakers, a few studies have also found that vowels are longer before fricatives than before stops (Raphael 1972; Sokolović-Perović 2009), so the results for this group are not surprising, and indeed support the limited research found on this topic. However, it appears that no research has shown a vowel duration difference before stops based on place of articulation.

The question arises as to whether the vowel duration difference or the dental/alveolar difference is the more important cue for listeners. The results of the perception study indicate that IE listeners have no difficulty in distinguishing the two types of stop even in word-final position; however, the words that contained a vowel duration difference were not included in the perception experiment. It may be that in IE, a final contrast

in place of articulation is known to be perceptually small, so for some vowels at least, there is a compensation by using vowel duration as a further acoustic correlate of the contrast. A further study involving manipulation of vowel duration could be used to determine whether listeners use this cue or the stop cue to distinguish words such as *bat* and *bath*.

While not the main focus of the study, a further finding was that for both groups, the vowel in words beginning with the voiced bilabial stop was significantly longer than the vowel in their counterparts beginning with the voiceless bilabial stop. Recall that aspiration was not included as part of vowel duration. (For the voiceless stop, aspiration was 70–90 milliseconds, while for the voiced stop, aspiration was 0–20 milliseconds.) Peterson & Lehiste (1960) examined the effect of multiple possible acoustic correlates on vowel duration. In terms of preceding stops, when aspiration was included in the measurement, vowels were longer after voiceless stops than voiced stops, as would be expected. When aspiration was not included, ‘the syllable nucleus following a voiceless plosive is usually shorter than that following a voiced plosive’ (1960: 701). The authors conclude that there is no reliable effect of voicing of the initial consonant on the duration of the following vowel. Heffner (1940) compared vowel duration in minimal pairs beginning with voiced versus voiceless stops. Again, when counted from the initial release rather than the beginning of voicing, vowels were longer after voiceless stops. However, when only the voicing part of the vowel was measured – as in the current study – he found different results for different pairs, with sometimes the vowel being longer after an initial voiced consonant and sometimes longer after a voiceless consonant. Finally, different speakers also showed different patterns in that study. In the current investigation, it was consistently found that vowels were longer after voiced stops, but a caveat is that only two words pairs were examined.

For VOT, it was found that for IE speakers, the dental stop had a shorter VOT than the alveolar stop. These results are in line with previous work, whereby the further back in the vocal tract the closure occurs, the longer the VOT (e.g. Lisker & Abramson 1964). An alveolar closure is further back than a dental closure; however, the difference between a dental and alveolar place of articulation is smaller than the difference between, for example, a bilabial and alveolar place of articulation, yet, the universal pattern holds. It is likely that this difference of 12–22 milliseconds is used as a cue by listeners to distinguish the stops. In a study examining listeners’ perception of place of articulation of voiced bilabial, alveolar and velar stops, Hazan & Rosen (1991) note individual variability in the cues listeners use, with some relying more on burst duration and some on formant transitions. Future perception experiments manipulating VOT would be required to determine how much listeners weight the VOT duration cue for the dental versus alveolar contrast here.

5.3 *Conclusions and future research*

The perception experiment in the current study had three main findings. Contrary to the hypothesis, AmE participants had a high level of accuracy in identifying words that

contained either an alveolar or a dental stop, even though this contrast is not in their native dialect. It is possible that their exposure to the dental stop as an allophone of either the alveolar stop or the dental fricative may have aided their processing of this sound. In line with the hypothesis, they did significantly worse than the IE participants. Both groups also did significantly worse when the contrast was in final position than when it was in initial position.

The production experiment found that for both groups of speakers, vowels were longer in *bat* and *bath* than in *pat* and *path*. Vowels were also longer in words ending in <th> than in words ending in <t> (*bat*, *pat*). For AmE speakers, this was explained by vowels being longer before fricatives than before stops, but for IE speakers, the finding that the vowel /a/ was longer before a dental stop than before an alveolar stop is new. This could be simply due to a place of articulation difference, although this has not been discussed in previous work, or it may even perhaps be an emerging vowel contrast in the language variety (although an argument for this would require further research in both production and perception). A perception experiment manipulating vowel duration could be used to determine whether this or place of articulation of the final stop is a more salient cue for listeners to distinguish words such as *bat* and *bath*. An examination of such words produced in natural speech rather than in wordlists would also clarify whether this vowel duration difference occurs in running speech. The finding that for word-initial voiceless stops, the VOT was longer for alveolar stops than for dental stops is in line with research on VOT and place of articulation. Once again, a perception experiment manipulating VOT would be necessary to determine if this is the main cue that listeners use to distinguish these two stops.

The current investigation adds insight into the relationship between production and perception. It appears that no other perception studies have yet been conducted on Irish English, particularly comparing speakers of two different dialects of English. The results found here indicate that while linguistic experience has an effect – shown by the higher accuracy among IE participants than among AmE participants – the AmE participants did well even without this experience, so were likely to be using a more acoustic than linguistic perception to hear the difference, and/or in fact were tuning into differences in the phonetic detail of the segments as opposed to making their judgements solely using their phonological system. Finally, the current study adds preliminary data on the acoustic differences between the dental and alveolar stops in IE.

Author's address:

Department of English
333 Fisk Hall
American University of Beirut
Beirut
Lebanon
nk114@aub.edu.lb

References

- Aschmann, Rick. 2015. *North American English dialects, based on pronunciation patterns*. <http://aschmann.net/AmEng/>
- Best, Catherine T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (eds.), *The development of speech perception: The transition from speech sounds to spoken words*, 167–224. Cambridge, MA: MIT Press.
- Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 171–204. Timonium, MD: York Press.
- Best, Catherine T. & Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20, 305–30.
- Best, Catherine T. & M. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Ocke-Schwen Bohn & Murray J. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*, 13–34. Amsterdam: John Benjamins.
- Boersma, Paul & David Weenink. 2011. *Praat: Doing phonetics by computer*, version 5.3.03. www.praat.org
- Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27, 207–9.
- Dart, Sarah N. 1991. Articulatory and acoustic qualities of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79.
- Dart, Sarah N. 1998. Comparing French and English coronal consonant articulation. *Journal of Phonetics* 26, 71–94.
- DiCano, C. T. 2012. Cross-linguistic perception of Itunyoso Trique tone. *Journal of Phonetics* 40(5), 672–88.
- van Dommelen, Wim A. & Olaf A. Husby. 2009. Perception of Norwegian word tones by Chinese and German listeners. In Michael A. Watkins, Andreia S. Rauber & Barbara O. Baptista (eds.), *Recent research in second language phonetics/phonology: Perception and production*, 308–21. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Donaldson, Tamsin. 1980. *Ngiyambaa, the language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- Dubois, Silvie & Barbara M. Horvath. 1998. Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change* 10, 245–61.
- Dupoux, E., P. Christophe, N. Sebastian Galles & J. Mehler. 1997. A distressing deafness in French. *Journal of Memory and Language* 36, 406–21.
- Edwards, Walter F. 2008. African American Vernacular English: Phonology. In Edgar W. Schneider (ed.), *Varieties of English*, vol. 2: *The Americas and the Caribbean*, 181–91. Berlin and New York: Mouton de Gruyter.
- Flege, James E. 1984. The effect of linguistic experience on Arabs' perception of the English /s/ vs. /z/ contrast. *Folia Linguist.* 18, 117–38.
- Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*, 233–77. Timonium, MD: York Press.
- Francis, A. L., V. Ciocca, L. Ma & K. Fenn. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36, 268–94.
- Frazier, Melissa. 2009. The production and perception of pitch and glottalization in Yucatec Maya. PhD thesis, University of North Carolina at Chapel Hill.

- Gordon, Matthew J. 2008. New York, Philadelphia, and other northern cities: Phonology. In Edgar W. Schneider (ed.), *Varieties of English*, vol. 2: *The Americas and the Caribbean*, 67–86. Berlin and New York: Mouton de Gruyter.
- Haowen, Jiang. 2010. Malayalam: A grammatical sketch and a text. Unpublished MS, Department of Linguistics, Rice University.
- Hayward, K. M., Y. A. Omar & M. Goesche. 1989. Dental and alveolar stops in KiMvita Swahili: An electropalatographic study. *African Languages and Cultures* 2, 51–72.
- Hazan, V. & S. Rosen. 1991. Individual variability in the perception of cues to place contrasts in initial stops. *Perception and Psychophysics* 49, 187–200.
- Heffner, R-M. S. 1940. A note on vowel length in American speech. *Language* 16 (1), 33–47.
- Hickey, Raymond. 1984. Coronal segments in Irish English. *Journal of Linguistics* 20, 233–50.
- Hickey, Raymond. 2008. Irish English: Phonology. In Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton *Handbook of varieties of English*, vol 1: *Phonology*, 71–105. Berlin: Mouton de Gruyter.
- House, Arthur S. 1961. On vowel duration in English. *Journal of the Acoustical Society of America* 33, 1174–8.
- House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25, 105–13.
- Hughes, E. J. & V. J. Leeding. 1971. The phonemes of Nunggubuyu. *Papers on the Languages of Australian Aboriginals (Australian Aboriginal Studies)* 38, 72–81.
- Huttar, George L. & Jean F. Kirton. 1981. Contrasts in Yanywa consonants. In Andrew B. Gonzalez & David Thomas (eds.), *Linguistics across continents: Studies in honor of Richard S. Pittman*. Manila: Summer Institute of Linguistics.
- Jongman, A., S. E. Blumstein & A. Lahiri. 1985. Acoustic properties for dental and alveolar stop consonants: A cross-language study. *Journal of Phonetics* 13, 235–51.
- Kallen, Jeffrey L. 2005. Internal and external factors in phonological convergence: The case of English /t/ lenition. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect change: Convergence and divergence in European languages*, 51–80. Cambridge: Cambridge University Press.
- Kallen, Jeffrey L. 2013. *Irish English*, vol. 2: *The Republic of Ireland*. Berlin: Mouton de Gruyter.
- Kondaurova, Maria & Alexander L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124(6), 3959–71.
- Kuhl, Patricia K. 1991. Human adults and human infants show a ‘perceptual magnet effect’ for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50, 93–107.
- Kuhl, Patricia K., Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens & Björn Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–8.
- Ladefoged, Peter. 2001. *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world’s languages*. Oxford: Blackwell.
- Lavoie, Lisa M. 2002. Subphonemic and suballophonic consonant variation: The role of the phoneme inventory. *ZAS Papers in Phonetics* 28, 39–54.
- Lisker, L. & A. S. Abramson. 1964. A cross-language study of voicing in initial stops. *Word* 20, 384–422.

- Lisker, L. & A. S. Abramson. 1970. The voicing dimension: Some experiments in comparative phonetics. In *Proceedings of the 6th International Congress of Phonetic Sciences*, 563–7. Prague.
- Maddieson, Ian. 1997. Phonetic universals. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 619–39. Oxford: Blackwell.
- Mann, Virginia A. 1986. Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English 'l' and 'r'. *Cognition* 24, 169–96.
- McAllister, R., J. Flege & T. Piske. 2002. The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30, 229–58.
- McGuire, G. & M. Babel. 2012. A cross-modal account for synchronic and diachronic patterns of /f/ and /θ/ in English. *Journal of Laboratory Phonology* 3, 251–72.
- Morphy, F. 1983. Djapu, a Yolngu dialect. In R. M. W. Dixon & Barry J. Blake (eds.), *Handbook of Australian languages*, vol. 3. Amsterdam: John Benjamins.
- Nagy, Naomi & Julie Roberts. 2008. New England: Phonology. In Edgar W. Schneider (ed.), *Varieties of English*, vol. 2: *The Americas and the Caribbean*, 52–66. Berlin and New York: Mouton de Gruyter.
- Ó hÚrdail, R. 1997. Hiberno English: Historical background and synchronic features and variation. In Hildegard L. C. Tristram (ed.), *The Celtic Englishes*, 180–200. Heidelberg: Winter.
- Pandeli, Helen, Joseph F. Eska & Joan Rahilly. 1997. Problems of phonetic transcription: The case of the Hiberno-English slit-t. *Journal of the International Phonetic Association* 27, 65–75.
- Peirce, J.W. 2007. PsychoPy – psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2), 8–13.
- Peterson, G. E. & I. Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693–703.
- Polka, Linda. 1989. The role of experience in speech perception: Evidence from cross-language studies with adults. PhD thesis, University of South Florida.
- Polka, Linda. 1991. Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America* 89, 2961–77.
- Pruitt, John S., James J. Jenkins & Winifred Strange. 2006. Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *Journal of the Acoustical Society of America* 119(3), 1684–96.
- R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. www.R-project.org
- 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.
- Remijsen, Bert & Caguor Adong Manyang. 2009. Luanyjang Dinka. *Journal of the International Phonetic Association* 39(1), 113–24.
- Sokolović-Perović, M. 2009. Voicing-conditioned vowel duration in Southern Serbian. *Newcastle Working Papers in Linguistics* 15, 126–37.
- Strange, Winifred. 1992. Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. In Yoh'ichi Tohkura, Eric Vatikiotis-Bateson & Yoshinori Sagisaka (eds.), *Speech perception, production, and linguistic structure*, 197–219. Tokyo: OHM.
- Strange, Winifred. 1995. Cross-language studies of speech perception: An historical review. In Winifred Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*, 3–45. Timonium, MD: York Press.

- Thomas, Erik R. 2008. Rural Southern white accents. In Edgar W. Schneider (ed.), *Varieties of English*, vol. 2: *The Americas and the Caribbean*, 87–114. Berlin and New York: Mouton de Gruyter.
- Wells, J. C. 1982. *Accents of English*, vol. 2: *The British Isles*. Cambridge: Cambridge University Press.
- Werker, Janet F. & Richard C. Tees. 1984. Cross-Language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49–63.
- Xu, Yisheng, Jackson T. Gandour & Alexander L. Francis. 2006. Effects of language experience and stimulus complexity on categorical perception of pitch direction. *Journal of the Acoustical Society of America* 120(2), 1063–74.

Appendix: Test words

| Initial voiceless Alveolar / Dental | Final voiceless Alveolar / Dental | Initial voiced Alveolar / Dental | Final voiced Alveolar / Dental |
|--|--------------------------------------|-------------------------------------|-----------------------------------|
| <i>tanks / thanks</i> | <i>boat / both</i> | <i>dare / there</i> | <i>bade / bathe</i> |
| <i>taught / thought</i> | <i>teat / teeth</i> | <i>den / then</i> | <i>breed / breathe</i> |
| <i>tin / thin</i> | <i>tent / tenth</i> | <i>die / thy</i> | <i>ride / writhe</i> |