2012

An Acoustic Account of the Allophonic Realization of /T/

Amber King
St. Cloud State University

Etienne Koffi
St. Cloud State University

Follow this and additional works at: https://repository.stcloudstate.edu/stcloud_ling

Part of the Applied Linguistics Commons

Recommended Citation
Available at: https://repository.stcloudstate.edu/stcloud_ling/vol1/iss1/12

This Article is brought to you for free and open access by the Repository at St. Cloud State. It has been accepted for inclusion in Linguistic Portfolios by an authorized editor of the Repository at St. Cloud State. For more information, please contact rsweelbaum@stcloudstate.edu.
1.0 Introduction

This paper is a laboratory phonology account of the different pronunciations of the phoneme /t/. Laboratory phonology is a relatively new analytical tool that is being used to validate and verify claims made by phonologists about the pronunciation of sounds. It is customary for phonologists to predict on the basis of auditory impressions and intuition alone that allophones exist for such and such phonemes. An allophone is defined as different realizations of the same phoneme based on the environments in which it occurs. For instance, it has been proposed that the phoneme /t/ has anywhere from four to eight allophones in General American English (GAE). To verify this claim Amber, one of the co-author of this paper recorded herself saying the words <still>, <Tim>, <kit>, <bitter>, <kitten>, <winter>, <fruition>, <furniture>, and <listen>. Her pronunciations are analyzed using Praat, an online acoustic phonetic software.

2.0 Overview of English Consonants and their Acoustic Behaviors

Fromkin et al. (2011) lists the main consonants found GAE in the chart below:

<table>
<thead>
<tr>
<th>Manner of Articulation</th>
<th>Place of Articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bilabial</td>
</tr>
<tr>
<td>Stop (oral)</td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>p</td>
</tr>
<tr>
<td>voiced</td>
<td></td>
</tr>
<tr>
<td>Nasal (voiced)</td>
<td>m</td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>f</td>
</tr>
<tr>
<td>voiced</td>
<td>v</td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>z</td>
</tr>
<tr>
<td>Glide</td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>m</td>
</tr>
<tr>
<td>voiced</td>
<td>w</td>
</tr>
<tr>
<td>Liquid (voiced)</td>
<td></td>
</tr>
<tr>
<td>(central)</td>
<td></td>
</tr>
<tr>
<td>(lateral)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: English Consonants (Fromkin, Rodman, & Hyam, 2011, p. 245)

Phonetically, consonants can be described with three main features: glottal state, place of articulation, and manner of articulation. Most acoustic accounts focus on the manner of articulation because it shows up better on spectrograms and waveforms. Ladefoged (2003, p. 159) writes that “acoustic phonetics analysis is not the best place to find out about different
places of articulation.” Below are some broad acoustic characteristics of classes of consonants according to their manner of articulation:

1. Stops /p, b, t, d, k, g/, on a spectrograph, are illustrated by what is called a stop gap; which is a weakening of sound. (Yavas, 2006). Voiced stops have a voice bar, which is a dark bar found at low frequencies (below 250 Hz) in the spectrogram, and a release burst that is shown by a prominent vertical line. Voiceless stops are characterized by aspiration which is indicated by a small frication noise that usually has a VOT duration of around 30 ms.

2. Nasals /m, n, ŋ/ perform very much like vowels but the main distinction is that nasals are not as dark as vowels are. They have a faint and low F1 frequency and a visible F3 around 2,500 Hz. Because of the smaller amount of energy produced by nasals, the F2 is normally not visible. (Yavas, 2006).

3. Fricatives /f, v, θ, ð, s, z, h/ are characterized on a spectrogram by a very unclear, scribbly pattern that has irregular vertical and horizontal lines. For sibilants such as [ʃ], this frication noise usually has an intensity of 58 – 68 dB. This particular fricative also usually appears between 2,000 and 6,500 Hz. (Yavas, 2006).

4. Affricates /ʧ/ and /ʤ/ appear as a combination of a stop and a fricative. They start with the stop-like closure and then end with friction noise. Yavas (2006) explains that though they share the same characteristics as the separate segments of a stop and a fricative, the duration of an affricate is shorter than the duration of a stop and a fricative.

5. Glides /w/ and /j/ perform very similarly to vowels on a spectrogram and are characterized by a low F1.

6. Liquids /l/ and /ɾ/ share some characteristics with stops and glides. Like stops, they are quite fast, but like glides they have a low F1. Unlike the glide /w/, liquids have a higher F3 that is barely visible in a spectrogram. (Yavas, 2006).

7. The flap [ɾ] differs from a stop in that it does not have a release burst.

8. The glottal stop [ʔ] is characterized by a lack of voicing but with a distinct line that is common to all other stops.

3.0 Data and Procedures

The data that serves as the basis for this paper comes from the recordings of Amber’s pronunciation of the words <still>, <Tim>, <kit>, <bitter>, <kitten>, <winter>, <fruition>, <furniture>, and <listen>. She read each word three times as naturally as possible in a quiet room and recorded them on an Olympus Digital Voice Recorder WS-710M that she borrowed from the university library. All in all, the analyses in this paper are based on 27 tokens (9 x 3). The measurements reported in each spectrogram represent the mean of three important acoustic correlates: pitch, intensity, and duration. These acoustic correlates are important for verifying the existence of the proposed allophones of /t/.

4.0 The Basic Phoneme

Phonologists claim that the pronunciation [t] that comes the closest to the underlying phoneme is when it is preceded by [s] at the beginning of a word. Therefore, in this paper, it is assumed that the pronunciation of [t] in the word [stɪl] (<still>) is the basic phoneme from which
all other allophones are derived. We will follow established linguistic conventions by placing basic phonemes inside of slanted bars / / . The allophones are placed inside of square brackets [   ], and the orthographic representations are contained within angle brackets <  >. The acoustic characteristics of the basic phoneme in Amber’s pronunciation are as reported in Figure 1:

The presence of [t] in the spectrogram and waveform is indicated by a tiny rectangle. The spectrogram shows a clearly demarcated burst release towards which an arrow points. The aperiodic waves on the waveform and the dark striations towards the top of the burst release indicate that [t] is mildly aspirated. The pronunciation of the basic form of /t/ lasts 32 ms. Now that the basic acoustic facts of Amber’s pronunciation of the basic form of [t] are known, let’s see how she pronounces it in the eight remaining words where it occurs.

4.1 The Aspirated Allophone of /t/

It is believed that most native speakers of GAE aspirate /t/ when it occurs before a stressed vowel. Phonologists state this pronunciation rule formally as follows:

/t/ → [tʰ]/ [+vocalic, +stress]

Does Amber’s pronunciation bear this rule out? What information can we glean from Figure 2 to answer this question?
The area of the spectrogram and the waveform within the circle shows evidence of strong aspiration. First, the waveforms are aperiodic and very dark. Secondly, we see a lot of striations in the spectrogram. We also see a fairly strong burst release. The length of the VOT (voice onset timing) is almost twice as long for $[tʰ]$ (60 ms) than it is for $[t]$ (32 ms). VOT is the time lag between the burst release and the beginning of the vibration of the vocal cords. It is indicated on the spectrogram by the two-headed arrow. A longer VOT is evidence that aspiration has taken place. Moreover, the overall length of $[tʰ]$ underscores the fact that it is different from the $[t]$ in $<$still$>$. In fact, it is more than three times longer. Linguists represent the aspirated allophone of $[t]$ as $[tʰ]$.

### 4.2 The Unreleased Allophone of /t/

In General American English (GAE), the phoneme /t/ is not released when it occurs at the end of words. Unreleasing means that the tip of the tongue stays in the alveolar area and does not come down until the speaker has to pronounce another sound. The International Phonetic Alphabet’s (IPA) symbol for unreleased /t/ is $[t̚]$. The rule that helps predict this pronunciation is stated formally as follows:

$$/t/ \rightarrow [t̚] / ___#$$

The spectrographic image in Figure 3 shows that Amber’s pronunciation of /t/ in $<$kit$>$ is unreleased. The area contained inside the circle shows that there are two burst releases for $[t]$. Both are very faint, which means that $[t]$ is not pronounced forcefully at all. The duration of the release /t/ is very short, only 4 ms. It is also not very loud compared with all the other pronunciations of /t/. Its intensity (loudness) is only 70 dB.
Most phonological accounts of [t̚] in GAE do not mention that it is lightly aspirated. Thomas (2011, pp. 122-125) mentions it but only in relation with one dialect of British English. However, this phenomenon may be more widespread in GAE than has been reported. We see it also in Doug’s pronunciation of the past tense suffix discussed in this volume.

4.3 The Two Flapped Allophones of /t/ 

In some pronunciations of /t/, the tip of the tongue touches the alveolar ridge very quickly. People with no linguistic training equate this sound with a [d]. But it is not a [d] because when people are asked whether they pronounce it the same way as the two [d]s in <did>, they say no. Linguists use the term “flap” or “tap” to describe this allophone of /t/. The IPA symbol used to transcribe it is [ɾ]. This pronunciation of /t/ is very widespread, and is in many respects a hallmark of GAE. We hear it when /t/ occurs between two vowels, the first of which is stressed. We also hear it in the word <little> even though /t/ is immediately followed by an [l]. Phonologists propose the following rule to predict the environment in which flapping occurs:

\[ /t/ \rightarrow [\text{ɾ}]/ [+\text{voc, +stress}] \text{ ___ } [+\text{syllabic}] \]

Acoustically, Ogden (2009, p. 114) estimates the length of [ɾ] to be 30 – 40 ms long. Amber’s pronunciation of 32 ms falls within the expected range.
We see on the spectrogram that there is no burst release. This means that /t/ is no longer pronounced as a stop consonant when it occurs in this phonological environment.

The acoustic description of [ɾ] in Figure 4 is very similar to the one in Figure 5:

The acoustic measurements of the two [t]s are almost identical. Only a small detail on the spectrogram shows that the /t/ in <winter> is different from the /t/ in <bitter>. On the edge of the area inside the circle, there is a faint burst release. This marks the presence of the alveolar nasal [n]. This [n] causes [ɾ] to be nasalized. For this reason, Ogden (2009, p. 114) refers to this pronunciation of /t/ as a “nasalized tap.”
The rule predicting this pronunciation can be stated formally as

/\t/ \rightarrow [ɾ̃] / [+nasal, +alveolar] ___ [+voc]

Ogden notes that the tapped nasal is between 30 – 40 ms. Amber’s pronunciation has the same duration.

The flapped nasal is an emergent pronunciation in GAE. Younger generations of GAE speakers in Central Minnesota use it more consistently and frequently than older generations. When Koffi’s son was in fourth grade, he wrote a story about his best friend. He consistently spelled his friend’s name as <Hunner> as opposed to <Hunter>. The pronunciation of <winter> by television weather forecasters in Minnesota vacillates between [t] and [ɾ]. We encounter the tapped nasal in common words such as <interesting> (first <t> and <twenty> (second <t>).

4.4 The Glottalized Allophone of /\t/  

The phoneme /\t/ is said to have undergone glottalization when it is pronounced in the glottis, that is, when the vocal folds open and close suddenly. One can have a feel for pronunciation when one says [o  o]. Between the first [o] and the second [o] there is a little stopgap. This is so because the vocal folds open and close quickly. When many GAE speakers pronounce the /\t/ in the word <kitten>, <mountain>, <button>, <mitten>, <flatten>, etc. the glottis acts exactly the same way. Phoneticians use the IPA symbol [ʔ] to represent glottalization.

Figure 6 shows that there is a stopgap in Amber’s pronunciation, as indicated by the area inside the rectangle. The burst release shows that [ʔ] is a stop. The aperiodic waves on the waveform indicate that there is faint aspiration inside of the stopgap. The glottalized /\t/ has the same duration as [ɾ] and [ɾ̃]. They are all about 30 ms long.
Phonologists propose a series of interacting rules to account for the glottalization of /t/. The derivation below tries to account for how the various rules interact:

- **Underlying phonemic representation**
  
  /# k i t e n #/

- **Stress rule**
  
  i

- **Aspiration rule**
  
  kʰ

- **Schwa rule**
  
  ø

- **Deletion of schwa**
  
  ∅

- **Syllabification rule**
  
  η

- **Glottalization of /t/**
  
  ?

- **Surface Phonetic Realization**
  
  [kʰiʔn̩]

The phoneme /t/ is glottalized as [ʔ] when it occurs between two vowels, the first of which is stressed. The unstressed vowel is usually followed immediately by an [n]. That unstressed vowel is then reduced to schwa, and is later deleted. The deletion of the schwa causes the segments [t] and [n] to form a cluster. However [tn] is an unacceptable coda cluster in English because the Obligatory Contour Principle (OCP) forbids two alveolar stops from clustering together. Roca and Johnson (1999, p. 273) refer to this constraint as a “phonological incest.” The OCP states that “similar melodies are disfavored as constituent siblings.” The sounds [t] and [n] are “siblings” because they belong to the same place of articulation. GAE speakers avoid violating the OCP by turning [n] into a syllable. This then leads to the glottalization of /t/. This is how generative phonologists account for the pronunciation /t/ in <kitten>.

### 4.4 The Fricative Allophone of /t/

The phoneme /t/ is pronounced as [ʃ] in many English words. Amber’s pronunciation of <fruition> bears this out, as indicated in the spectrogram and the waveform by the area inside of the circle. The dark aperiodic waves on the waveform indicate that there is a lot of acoustic energy being produced. The striations in the upper frequency levels of the spectrogram point to the same reality. The fact that the acoustic energy is spread over a large area, and that there is no voicing bar at the bottom of the spectrogram mean that /t/ has turned into the voiceless palatal fricative [ʃ].
The phoneme /t/ is pronounced this way most often when it occurs between a stressed vowel and a high front vowel. The rule for this pronunciation is stated formally as follows:

\[ /t/ \rightarrow [ʃ] \ [\ +\text{voc}, \ +\text{stress}] \quad [\ +\text{voc}, \ +\text{front}, \ +\text{high}] \]

We encounter this pronunciation in words such as <creation>, <initial>, <caution>, <participation>, etc.

4.5 The Affricated Allophone of /t/

There are many common words such as <denture>, <picture>, <future>, <lecture>, <furniture>, etc. in which /t/ is pronounced as [tʃ]. The rule for this pronunciation can be stated straightforwardly as follows:

\[ /t/ \rightarrow [tʃ] \ [\ +\text{voc}, \ +\text{stress}] \quad [\ +\text{voc}, \ +\text{high}, \ +\text{back}, \ +\text{rhotic}] \]

The feature [+rhotic] has to do with the sound [ɹ] that is present in most of these words. The rule states that /t/ is pronounced [tʃ] when it occurs between two vowels, the first of which is stressed and the second of which has the features [+voc, +high, +back, +rhotic].

The spectrographic picture and the waveform confirm that Amber pronounces /t/ as [tʃ]. This sound is an affricate, that is, it is a combination of a stop and a fricative. This combination can be seen clearly in Amber’s data. The stop is inside a circle, and the fricative is inside of a rectangle. An arrow points to the burst release on the spectrogram.

Two important acoustic correlates make the fricated allophones [ʃ] and [tʃ] stand out. They both have relatively high pitch, 225 Hz and 255 Hz respectively. They also have the longest durations of all the allophones, 137 ms and 107 ms respectively.
It should be noted before leaving the topic of frication that there are a number of words for which the pronunciation of /t/ vacillates between [tf] and [ʃ] in Central Minnesota and in some parts of the US. Two such words are <presidential> and <question>. In Koffi’s linguistic classes, students are equally split as far as the pronunciation of /t/ in these two words is concerned. Some realize it as [tf] while prefer [ʃ]. Johnson (2012, p. 162) provides the following explanation for why this is so:

Many researchers have noted that it is difficult to measure the acoustic characteristics of fricatives (particularly coronal fricatives), because there may be several spectral peaks as in the production of [s], and from utterance to utterance one or the other of these may have greatest amplitude. It has also been noted that there may be substantial range of inter-speaker variability in the frequencies of the spectral peaks of fricatives.

### 4.6 The “Deleted” Allophone of /t/

The phoneme /t/ is often silent in words such as <listen>, <fasten>, and <soften>. Phonologists note that a number of rules must interact to make deletion possible. The derivation below is our best attempt to show how these rules interact:

<table>
<thead>
<tr>
<th>Underlying Phonemic Representation</th>
<th># l ɪ s t ɛ n # /</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress rule</td>
<td>i</td>
</tr>
<tr>
<td>Schwa rule</td>
<td>ə</td>
</tr>
<tr>
<td>Deletion rule</td>
<td>Ø</td>
</tr>
<tr>
<td>Syllabification rule</td>
<td>η</td>
</tr>
<tr>
<td>Surface Phonetic Realization</td>
<td>[lisŋ]</td>
</tr>
</tbody>
</table>

To see clearly what is going on in Amber’s pronunciation, let’s examine both the spectrogram and the waveform in Figure 9. It is sectioned off into three areas. The area within the first rectangle deals with the vowel [i]. The area inside of the circle deals with [s]. The area within the last rectangle corresponds to [ə]. The spectrogram and the waveform show that Amber does not pronounce the phoneme /t/ in <listen>. This is the reason why its pitch, intensity, and duration are labeled “undefined”.

https://repository.stcloudstate.edu/stcloud_ling/vol1/iss1/12
The spectrogram also shows that Amber pronounces the second vowel as a schwa. This is the reason why the word is transcribed phonetically as [lɪsən], not [lɪsn]. The presence of the schwa is clearly visible in the area inside of the second rectangle.

5.0 Conclusion

This paper has shown that Amber’s pronunciation bears witness to the fact that /t/ has eight different allophones. The aspirated allophone [tʰ], the allophones [ʃ] and [tʃ] are acoustically salient. Consequently, they should be easier to teach to nonnative speakers of English. However, the flapped allophone [ɾ], the nasalized flapped [ɾ̃], the glottalized [ʔ], and the unreleased [t̚] can be more challenging to teach because they lack perceptual prominence. Amber intends to write her MA thesis on this topic. The data used in this analysis will be expanded to include the pronunciation of /t/ by nonnative speakers. This will give us the chance to verify the reliability of the pedagogical implications made on the basis of Amber’s pronunciation.

ABOUT THE AUTHORS

Amber King is a graduate student in the MA TESL/Applied Linguistic program. She has spent most her life in Minnesota. She earned a BA degree from the Minnesota State University Mankato in Spanish. Her area of interest is the interface between acoustic phonetics and phonology.
Ettien Koffi is Professor of Linguistics. He teaches the linguistic courses in the MA TESL/Applied Linguistics at Saint Cloud State University. His research interest is at the interface between acoustic phonetics and phonology.

References


