Preparation for discussion: Voice onset time (VOT)

The difference between a voiced and a voiceless sound is that during a voiced sound, the vocal folds are vibrating, but during a voiceless sound, the vocal folds are not vibrating.

There is another property by which voiceless sounds can differ: they can be \textit{aspirated} (accompanied by an [h]-like puff of air) or \textit{unaspirated}. Aspirated sounds are transcribed with a superscript $[^h]$ placed after the symbol for the unaspirated voiceless sound. (Compare the English words \textit{pie}, \textit{spy}, \textit{nap}; can you hear the difference between aspirated [p$^h$] and unaspirated [p]? Which word(s) make(s) a Kleenex move?)

The difference between voiced, voiceless unaspirated, and voiceless aspirated sounds is a matter of the \textit{relative timing of oral and glottal closures/openings}.

Prepare for class discussion by reading through this handout and making sure you understand the diagrams. Also, try to answer the questions in sections (3) and (4); these are somewhat difficult, but see what you can do.

1. Schematic diagrams of movements in the vocal tract

We can use the following convention to represent the various structures in the vocal tract that can open and close:

\begin{center}
\begin{tikzpicture}
  \draw[<->,line width=2] (0,0) -- (2,0);
  \draw[<->,line width=2] (0,0) -- (0,1);
  \draw[<->,line width=2] (0,1) -- (0,2);
  \draw[<->,line width=2] (0,2) -- (2,2);
  \draw (1,0) -- (1,1);
  \draw (1,0) -- (2,0);
  \draw (1,1) -- (2,1);
  \draw (1,2) -- (2,2);
\end{tikzpicture}
\end{center}

So if we wanted to represent the sequence [mɑ], we could show that first the lips are closed and the velum is open, and then the lips are open and the velum is closed. We can also show that the vocal folds are vibrating (\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash) rather than open or tightly closed.

\begin{center}
\begin{tikzpicture}
  \draw (0,0) -- (2,0);
  \draw (0,0) -- (0,1);
  \draw (0,1) -- (2,1);
  \draw (0,1) -- (0,2);
  \draw (0,2) -- (2,2);
  \draw (1,0) -- (1,1);
  \draw (1,0) -- (1,2);
  \draw (1,2) -- (2,2);
  \draw (1,1) -- (1,2);
\end{tikzpicture}
\end{center}

Lips: \\
Velum: \\
Glottis: \textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash\textbackslash

Note that the lips open at roughly the same time as the velum closes.

- What would we have instead of [mɑ] if the velum stayed open when the lips opened?
- What would we have instead of [mɑ] if the lips never opened?
2. **Voice Onset Time (VOT)**

Both voiceless unaspirated stops like [p] and voiceless aspirated stops like [pʰ] are produced with the vocal folds spread apart so that they can not vibrate. The way these two kinds of stops differ is in the **amount of time** between the **release of the stop closure** and the **onset of vocal-fold vibration** — this is known as the **voice onset time (VOT)**.

A. **Voiceless aspirated stops** (as in English *pa* [pʰɑː]) look like this:

```
Oral closure (lips for [p]) \____________________________

Glottis \\\\\\\\\\\\\\\\\\______________

Time (ms) .....................0......*............... (* = onset of voicing)
```

Notice that there is an interval of time during which the oral closure is open but the glottis is still open too. This is the time during which aspiration occurs: note that the state of the vocal tract is just like that for [h].

We say that voiceless aspirated stops have a **positive VOT**, because voicing starts *after* the oral closure is released. (Think of a timeline, with “zero” at the point of release of the oral closure.)

B. **Voiceless unaspirated stops** (as in English *spa* [spɑː] or French *pas* [pa]) look like this:

```
Oral closure (lips for [p]) \____________________________

Glottis \\\\\\\\\\\\\\\\\\______________

Time (ms) .....................0..................................
```

We say that voiceless unaspirated stops have a **VOT of (very close to) zero**, because voicing starts just about when the oral closure is released. (In real life, many stops that we would classify as voiceless unaspirated stops have a **very small positive VOT** — voicing starts a few milliseconds after the closure release.)
C. Voiced stops

For a stop to be literally voiced, there has to be voicing (vocal-fold vibration) during the stop closure.

A voiced stop (as in English about [əˈbɔːt] or French bas [ba]) looks like this:

```
Oral closure               \_____________
(lips for [b])             \_____________
Glottis                   \//\//\//\//\//\//\//\//\//\//\//\//
Time (ms)                 *...................0........................... (* = onset of voicing)
```

We say that voiced stops have a negative VOT. (Do you see why it's negative?)

3. The truth about “voiced and voiceless stops” in English

For “voiced stops” in certain contexts in English, if we pay careful attention to them or measure them instrumentally, we find that they do not have the negative VOT of true voiced stops — there is no actual voicing during the stop closure. Instead, they have a VOT of approximately zero, making them phonetically voiceless unaspirated stops like those in French, Hindi, or Thai.

Contexts where English “voiced stops” are likely to be phonetically realized as voiceless unaspirated stops are at the beginning of a phrase or following a voiceless sound. When English “voiced stops” are preceded by voiced sounds, they are more likely to be phonetically voiced — that is, the voicing is more likely to continue through the stop closure.

So why do some linguists call the English stops in situations like bout, door, go “voiced” when most of the time they’re voiceless unaspirated? Two reasons.

• First of all, English speakers consider these stops to be cognitively (phonologically) the same sound as the sounds in about, adore, ago, which usually are truly voiced stops.

• Secondly, the vocal folds in the stops in bout = [p] and about = [b] (etc.) are held in basically the same position — a position that is compatible with voicing. So the stop in bout is not “actively” voiceless. It just ends up being voiceless because not enough of an effort is made to expand the size of the oral cavity to compensate for the buildup in pressure that occurs when air is pushed through the glottis during voicing. In words like pout and apple, the [p(ʰ)] is usually produced with the vocal folds actively spread apart to prevent voicing.

There are also linguists who argue that the primary distinction between the “voiceless” and “voiced” stops of English should be seen, not as voicing, but as aspiration. (On this view, the unaspirated stops become secondarily voiced in certain contexts.)

Here are some examples of what often happens in English with stops, using “p” as a representative of voiceless stops and “b” as a representative of voiced stops.
Narrow transcription: The English sounds “p” and “b” in sample contexts

<table>
<thead>
<tr>
<th></th>
<th>Word-initial</th>
<th>After vowel</th>
<th>After [s] in same syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>English “b”</td>
<td>[p] bout</td>
<td>[b] about</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>[b] able</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[pʰ] apple</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

So sometimes the contrast between “b” and “p” in English is between a voiceless unaspirated and a voiceless aspirated stop, and sometimes it is between a voiced stop and a voiceless (aspirated or unaspirated) stop. In a consonant cluster preceded by [s], as in spout, there is only one kind of stop that is possible in English: voiceless unaspirated.

To think about:

Jimi Hendrix lyric: ‘Scuse me while I kiss the sky
Misheard as: ‘Scuse me while I kiss this guy

Why?

4. VOT beyond stops

Discussions of VOT typically focus on stops, because it is stops that are most likely to have multiple VOT categories. But we can extend this idea to other classes of consonants.

To think about:

- **Fricatives and affricates** can also be voiced or voiceless. How could we measure the extent of voicing in a fricative or affricate? What landmarks would we need to use in our duration measurements?
- **Affricates** can be aspirated. (So can fricatives, although this is much more rare.) If we are looking for evidence of positive VOT in an affricate or a fricative, what landmarks should we use in our duration measurements?