Linguistics 601 *Introduction to Linguistic Analysis*
Winter, 2010
Robert Levine
11 January

Phonetics, part 1:The vocal tract as an acoustic mechanism

The one concrete, physically measurable aspect of human linguistic capability is the system of speech sounds which provides, in each language, the raw materials for constructing meaningful units. Since speech sounds *are* sounds, they have the same origin as all other sounds: in *pressure waves* arising from some initial impact on the air in a certain region of space, much like the symmetrical wave which is created when a stone is thrown into a volume of water. The impact temporarily creates a wave front of increased density, which propagates outward in decreasing strength until it finally disappears at some distance from the source. When this wave front reaches a human ear, the energy communicated to the auditory system ultimately results in a particular sensation which depends on the amount of energy carried by the wave, its characteristic frequency and the component spectrum of that frequency, which is what a sound spectrograph measures. In all these respects, speech sounds are like other sounds in nature.

What differentiates speech sounds from others is the role the former play in the organization of natural language. Speakers of each language select a relatively small number of sounds producible by the vocal organs to construct higher-order objects—that is, meaning-bearing elements—from; or, more precisly, to supply the pronunciation of such elements. There is in fact quite range of sounds to choose from, made available by physical gestures in the *vocal tract*—a term which includes the lungs, esophagus, glottis, tongue, the whole mouth cavity, the nasal passages, and the lips.

1 Air stream mechanisms

Crucial to the production of any speech sound at all is the existence of a stream of air which can serve as the impact source for a compression wave. There are four mechanisms which can provide this airstream in the human vocal tract.

1.1 Egressive mechanisms

So-called egressive mechanisms produce an air stream within the vocal tract which exists the body and creates a pressure wave in the surrounding air.

1.2 The pulmonic source

The lungs are the primary source of the airflow which creates teh necessary impact on the surrounding air. The diaphragm and intercostal muscles jointly provide plenty of pressure on the lungs, and the volume of air that the lungs can hold guarantees a relatively long lifespan for any sound made this way, and the process by which the lungs empty and fill is already in place as part of normal breathing, so nothing new has to be introduced. For these reasons, it makes a lot of sense that in every language in the world, the lungs not the main source of airflow for the production of speech sounds.

1.3 The glottalic source

Things get a lot more complex when we look at the use of the glottis, or larynx, or vocal cords, as the source of air flow. The process here is quite specialized: no other use is made of this mechanism other than the production of speech sounds. The sequence of events in generating this air flow can be summarized as follows:

i. the vocal cords shut completely, while simultanteously a closure is created in some other part of the vocal tract higher up;
ii. internal pressure is applied to the closed vocal cords, forcing them upwards, and thus creating pressure on the air trapped between the closed glottis on the bottom and the point of closure at the top;
iii. the upper closure is suddenly, released;
iv. the 'tube' of previously trapped air is expelled from the mouth, colliding with the surrounding air and generating a compression wave.

In the early years of the previous century, investigators were struck by the sharpness of such sounds; they were typically written p!, t! and so on in Amerindian phonetic transcription systems. We write these as p', t', etc., for reasons that will become clear when we start looking at specific IPA symbols for certain sounds.

1.4 Ingressive mechanisms

Ingressive mechanism pull air *inwards*. The effect is the same—creation of a shock wave in the surrounding air—but the point of origin of the compression effect is inside the vocal tract, rather than outside it. There are two physical sequences which create this kind of airflow:

1.4.1 Glottalic

Just as there is an egressive glottalic mechanism, there is an ingressive glottalic sequence of vocal tract action, which is in some respects the reverse of what was described above for sounds like p', though the mechanism is a good deal more complicated in the case of glottalic ingressives:

i. the vocal cords shut completely, while simultanteously a closure is created in some other part of the vocal tract higher up;
ii. internal pressure is applied to the closed vocal cords, forcing them downwards, and thus creating pressure on the air trapped between the closed glottis on the bottom and the point of closure at the top;
iii. the lower closure is suddenly, released—i.e., the glottis is opened;
iv. the 'tube' of previously trapped air is drawn through the vocal cords, causing them to vibrate; the upper closure is released, and typically air from the lung then flows outward.

1.4.2 Velaric, aka clicking

The velaric mechanicism involve a closure created by the tongue, which is then lowered, creating a lower-density region in the mouth between the tongue root and the point of closure higher up. The upper closure is abruptly released, causing the air above the tongue closure to move rapidly into the low-pressure region originally created, giving rise to a very sharp clicking kind of sound.

2 Articulators

The articulators are the parts of the vocal tract which can move, to create points of closure in the vocal trace. The engineering reason for creating such closures is the resulting change in the size and shape of the vocal tract through which air flows once the closure is released, giving rise to quite different acoustic 'signatures' perceived by our auditory system as different sounds. Think of the valves and stops on wind instruments and you'll see the point: a (near-)continuous flow of air, interrupted momentarily by closures in the mouth and elsewhere changing the countours of the resonating chambers through which the air moves, with each change alterning the properties of the acoustic wave in specific respects that register in our neural interpretation centers as different speech sounds.

2.1 The tongue

The tongue is by far the most versatile and finely-tuned articulator we possess. The concentration of taste-receptor neural buncless at the front of the tongue is matched by a very high degree of motor control in the same are. This means that we have close to pinpoint accuracy in moving the very tip of the tongue, and a small region immediately back from it, to particular places in the mouth to create many phonetic distinctions in a relatively small region of space. The front of the tongue contacts the teeth, the ridge just behind the from incisors, the hard palatte behind this ridge, and so on; but, due to the shape of the tongue and the limits of muscle control, the middle and rear of the tongue come increasingly into play as the tongue's targets of contact at the top of the mouth move further back.

2.2 The lips

The mobility of the lips makes them well-suited for articulation, but their range of movement is limited by physiology to the very front of the vocal tract: they can contact each other, and they can contact the front teeth. But they can also move forward (as when you try to blow air onto your hand in a fairly narrow stream) to extend the size (and significantly alter the shape) of the oral resonating chamber—the volume of space in the mouth that the airstream bounces around in en route to its encounter with the air mass outside the mouth.

2.3 The vocal cords

High up in the throat sits the glottis (or larynx), a rather complex structure comprising sheets of cartilage and muscle tissue which regulates the space between two thick folds of epithelial tissue—the same kind of stuff our skin is made of. The glottix functions to in effect guard the trachea, the tubular membrane connecting the mouth and lungs; by shutting completely (closing the vocal cords tightly together) it is possible to keep liquids from entering the lungs; thus the vocal cords close when you're holding your breath under water. But a side benefit of the larynx is the possibility of using the vocal cords to add certain phonetic properties to the speech signal, creating the possibility of additional sound distinctions. The principal property along these lines is called *voicing*: when the vocal cords are partially closed, they can be induced to vibrated under the pressure of air from the the lungs whose movement they partially block. You can see how this work (roughly, anyway) by taking a sheet of thin paper stretching it tightly in from of you mouth so that your breath goes just over the top of it, and then blowing on it. Nothing will happen at first; but if you gradually release the tension you're exerting on the sheet, a point will come when the top of the paper begins to vibrate rapidly back and forth. This is what happens with the vocal cords, and because we actually have rather fine control over this component of our respiratory anatomy, it's possible for human beings to modify and fine-tune the region of the larynx in which airflow-triggered vibration takes place. We can, for example, get a kind of reduced or partial voicing, technically labeled 'laryngealization', in certain classes of sounds, where the laryngealized versions are treated, phonologically, as distinct from their normally voiced corresponding sounds.

3 Places of articulation

The articulators are the parts that move; but where do they move *to?*. In the case of the vocal cords, that's easy: they move towards and away from each other, and that's the end of the story. In the case of the lips, as pointed out earlier, the range of motion is also quite limited: towards or away from each other, or to the teeth. The tongue, in contrast, has a great deal of freedom of motion, and can move to and from quite range of targets, at each of which it can impose either a complete or a partial blockage of the air stream. In either case, by controlling the size and shape of the volume in which air is confined prior to its release (resulting in the formation of a pressure wave with a distinctive structure), the tongue plays a major role in creating the distinctive acoustic signatures that we recognize, via our interpretive wetware and software, as those of distinct sounds. The following comments sketch some of the specific targets of the tongue's movement in creating these closures.

3.1 The teeth, and more

The tip of the tongue contacts the edge of the upper central incisors (more familiarly referred to as the front teeth), and often doesn't stop there, but actually goes *past* the teeth, so that what comes in contact with the edge of the incisors isn't the tip, but the broader section immediatly behind it. If you experiment with this point of contact, you'll quickly discover that it's rather difficult to completely block the flow of air past the tongue in this far-forward position, or even when the tongue tip is placed on the inner surface of the upper front teeth, just short of their cutting edge.

3.2 The alveolar ridge

In contrast, it's quite easy to create a complete closure by placing the tip of the tongue against the *alveolar ridge*, a slight fold in the upper part of the mouth just behind the upper central incisors. In this position, enough of the tongue is raised, and tightly enough, to prevent virtually any air to pass between the mouth and the outside air, regardless of the airstream mechanism.

3.3 The hard palate

Behind the alveolar ridge, the roof of the mouth curves upward as a hard, bony concave shape whose surface is covered by a thin membrane of tissue. This region is called the *hard palate*, and while the very front part of the tongue can make contact with it, there are few if any languages displaying speech sounds based on such contact. Instead, the part of the tongue used to create closures in this part of the mouth is the *laminal* region, just forward of the midpoint of the tongue, colloquially lableled the 'blade'. Here the tongue has far fewer surface nerve ending (hence much less sensitivity) along with a good deal less accuracy in placement.

3.4 The velar continuum

If you run your finger along the roof of you mouth (and are able to suppress your gag reflex for long enough), you'll feel an abrupt change where the hard palate ends, replaced by a soft, fleshy area that runs quite far back towards the throat. This 'soft palate' is actually not part of the same skeletal structure as the hard palate; it is instead a flap of muscle, covered by a layer of membranous tissue similar to the lining of the throat, which can move downwards to uncover a passageway into the nasal cavity. The middle and back part of the tongue rises to make contact with the velum, and, while this section of the tongue is considerably reduced in flexibility from the frontmost part, there is still a range of contact points available to it, from the very front of the velum, right at the border with the hard palate, to the very back part, just before the structure called the uvula.

3.5 The uvula

The uvula pretty much marks the end of the mouth (as we usually think of it) as the oral part of the vocal tract. Any further back and you're into to the throat. Sounds made with a uvula contact are reserved for the very backmost part of the tongue, and are abundantly attested, in an almost exuberant manner, in native languages of the American Pacific northwest, the Pacific coast and far western prairies of Canada, and Alaska. By the way, a caution to those of you who are interested in supplementing your reading and coursework by consulting online sources which are only weakly vetted, such as Wikipedia: their entry on the uvula contains the comment that

> *The uvula functions in tandem with the back of the throat, the palate, and air coming up from the lungs to create a number of guttural and other sounds. Uvular consonants are not found in most*

> *dialects of English, though they are found in many Semitic, Caucasian, and Turkic languages, as well as several languages of Western Europe such as German, French and a few Celtic languages. Certain African languages[which?] use the uvula to produce click consonants as well, though other than that, uvular consonants are fairly uncommon in Sub-Saharan Africa.*

No mention is made here of the Salishan family of aboriginal languages, or of Tlingit, a Na-Dene language of Alaska, which between them have probably more uvula consonants of various types—some of them unknown, as far as I know, anywhere else on planet Earth—than the rest of the world's language combined. You just have to be very critical about what you take away from publically constructed knowledge bases such as Wiki. To an untrained ear, uvular sounds give the impression of what English speakers would regard as 'k-like' sounds, except that there's something wrong that's hard to identify. What's 'wrong' is just that the sound is made further back—for some speakers of languages with these sounds, *way* further back—than the velar sounds, e.g., [k], are made. What you often find is that the velar and uvular sounds are treated in a very similar fashion: the possible modifications and variants of the former are also displayed by the latter. More on this when we get down to specific descriptions of classes of speech sounds. . .

3.6 The pharynx

With the pharynx, we in effect take our leave of the mouth as a source of speech sounds. The pharynx is just the back of the upper throat, and the only way contact can be made with the pharynx is by moving the tongue root—the base point of attachment to the hyoid bone (itself a rather curious structure, the only bone with no attachment to any other bone, but instead secured and stabilized by a system of ligaments)—backwards. This produces a constriction of the throat whose effects on the flow of air from the lungs are clearly recognizable, even though it's not obvious to the untrained ear just where those effects are coming from.

4 What's next...

What I've presented here covers the principle components of speech sound articulation—the bits that interact to give rise to the wide range of sounds we observe in human speech across the world's languages. What we have yet to do is examine the way in which these elements are brought together to create, first, the speech sounds of American English, and second, a selection of possibilities from elsewhere in the world that illustrate how small differences in the use of these physiological mechanisms can produce seemingly very exotic, difficult-to-recognize phonetic material (thus, the famously 'difficult' Welsh *ll*, appearing in scary-looking words such as *twllwch* and many Welsh placenames beginning with *Llan-*, is nothing more than the analogue of English [l] along the same lines that [s] is the analogue of [z]: the difference in each case is solely whether or not the vocal cords are vibrating). In the next 'handout', I'll give you a working subclassification of English sounds, based on *where* and *how* they're made, that will make clear the systematic nature of the English speech sound inventory. With luck, I'll be able to get this posted before the weekend is over...