Understanding the Sounds of English: A Case for Media Broadcasters and Announcers

Shehu Sidi Ibrahim
Department of English and Literary Studies
Faculty of Arts
Usmanu Danfodiyo University, Sokoto, Nigeria.
Email: ccd379@yahoo.com

Abubakar Atiku Alkali
Department of English and Literary Studies
Faculty of Arts
Usmanu Danfodiyo University, Sokoto, Nigeria.
Email: alkali.abubakar@udusok.edu.ng
DOI: 10.56201/rjmcit.v10.no5.2024.pg85.92

Abstract
The most important form of communication is perhaps the spoken form. The naturalness of speech makes it to be the most effective means of communication. It is the most used medium, and it is particularly important because it assures the use of the sensitive organs in various degrees simultaneously. In everyday conversation, according to Crystal (1988), people speak about five to six syllables a second – around 300 a minute. This is an average, of course. Some people are naturally fast, and others naturally slow in their manner of speech. And speed varies greatly depending on the context. When reading aloud, the average is much lower – around 250 syllables per minute (spm). Reading the news on radio or television may produce even slower speeds, of around 200 spm. By contrast, in the middle of an exciting story, in intimate surroundings, a speaker can easily reach speeds of 500spm – though not usually for more than a few seconds at a time. Whether or not the validity of these findings could be extended to non-native speakers is an issue waiting to be investigated.

Keywords: *Sounds, Consonants, Vowels, Television, Radio, Broadcast*

1. Introduction

In all languages of the world, a non-native speaker is mostly identified from the way and manner he/she pronounces words in a target language. It is important therefore to strive to pronounce words, phrases and sentences correctly so as to be easily and fully understood. English is one of the languages in which it is difficult for non-native speakers to achieve expertise for a number of reasons. One of these is the absence of one-to-one correspondence between spelling and pronunciation of many English words. Another reason is the occurrence in English of sounds

which are seldom found in other (particularly Nigerian) languages. Other reasons given by scholars include learning English at a later age after proficiency must have been attained in the native language as well as being taught by non-native teachers who are themselves not very fluent in English.

English comprises 44 (forty-four) sound segments. Out of these, 24 (twenty-four) are consonant sounds while 20 (twenty) are vowel sounds. This contrasts with the English alphabet in which there are altogether 26 (twenty-six) letters of which 21 (twenty-one) are consonants whereas 5 (five) are vowels. The production of human speech is made possible by the air breathed in and out. The air inhaled is forced out through the vocal cords and the mouth for oral sounds such as /p/, /b/, /t/, etc. or through the nose for nasal sounds like /m/, /n/ and /ŋ/. When the vocal cords are spread apart such that air coming from the lungs passes unimpeded, 'voiceless' sounds are produced, for example /t/, /θ/, /s/, etc. On the other hand, when the vocal cords are drawn together in a manner that air has to force its way through, 'voiced' sounds are produced for example, /b/, /d/, /g/, etc.

Contrary to what some scholars may suggest, there is no part of a human being which is specifically designed for talking. The parts of the body which produce the sounds of language are incidentally useful for this purpose, but they all have other duties to perform which from the biological point of view, are older and more important duties that are related to, for example, breathing, chewing, swallowing, smelling, and other such activities. In the words of Sapir (1921), "Physiologically speech is an overlaid function, or, to be more precise, a group of overlaid functions, nervous and muscular, that has come into being and is maintained for very different ends than its own".

The organs which we use for talking are basically the same, and work in the same way, in all human beings. No racial differences in their structure or in their manner of acting have ever been established. It follows from this that everybody who is free from abnormalities is capable of pronouncing anything. Thus, the unpronounceable sounds which some languages are alleged to possess are myths (Abercrombie 1967).

More than half of the human body, from the head to the abdomen, is needed for the production of spoken language. There are three groups, or systems, of bodily organs which are brought into cooperation for this purpose: one group lies in the trunk, one in the throat, and one in the head, and they are usually known respectively as the respiratory system (the lungs, the muscles, the bronchial tubes and the windpipe or trachea), the phonatory system (the larynx or voice box), and the articulatory system (the nose, the lips and the mouth and its contents including especially the teeth and tongue).

2. **Use of English in the Media**

With the unprecedented proliferation of broadcast outfits in Nigeria, both radio and television in the last decade vis-à-vis the dearth of qualified broadcasters and announcers, the need for trained personnel to handle these outfits would not have been more pressing. Monitoring news broadcast

on some of these media outfits leaves one with the impression that untrained or poorly trained newscasters and announcers are being hired to perform such important tasks.
In broadcast performance situations, it is essential to pronounce words correctly. This involves two things: determining what the accepted form is and then making sure that this accepted form is produced right on the air. Broadcasters and announcers must realise the importance of proper articulation of words because since our voice is inevitably carried over a long distance and through many electrical devices, tendency is that what might be heard perfectly well in a face-to-face interaction, could be distorted by the time it reaches the audience.
Mispronounced words succeed only in alienating listeners and causing the announcer/broadcaster and the station to lose credibility. Broadcasters are therefore expected to pronounce words correctly. The causes of mispronunciation are many from simply unfamiliarity with the correct pronunciation to serious speech impediments.

The first step in becoming conscious of one's pronunciation is to record oneself and analyse one's work. Although it may be a painful experience for one to critique oneself after every recording, the benefits accruing from this exercise for an announcer/ broadcaster wanting to succeed in his/her chosen profession are enormous. A recording of a newspaper article of modest length may be selected for the practice and in doing the critique the following points may be considered:

- Are the endings of the words being dropped?
- Are there indistinct or omitted consonants within words?
- Are words being slurred together?
- Are there obvious cases of first language interference?
- To what extent are spelling/pronunciations prominent?
- Are proper nouns and articles properly pronounced?

In the next section of the paper, brief explanations of basic English sounds in respect to their place and manner of production or articulation are given.

3. Vowels

Traditionally, as Lyons (1981) puts it, vowels have been specified in terms of the position of the highest point of the tongue and the position of the lips. In the table below, they are divided into seven *short* and five *long vowels.* An alternative way of organising them is according to where (in the mouth) they are produced. This method allows us to describe them as *front, central* and *back.* We can qualify them further by how high the tongue and lower jaw are when we make these vowel sounds, and by whether our lips are rounded or spread, and finally by whether they are short or long. This scheme shows the following arrangement:

3.1 Front vowels

- /i/ -cream, seen (long high front spread vowel)

- /ɪ/ - bit, silly (short high front spread vowel).
- /ɛ/ - bet, head (short mid front spread vowel); this may also be shown by the symbol /e/.
- /æ/ - cat, dad (short low front spread vowel); this may also be shown by /a/.

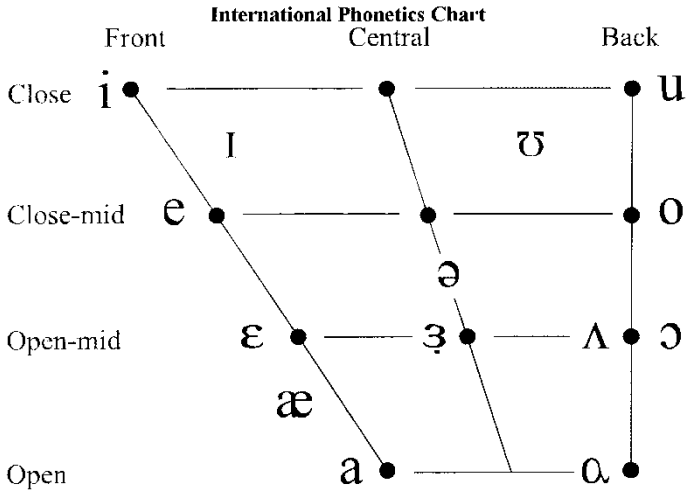
3.2 Central vowels

- /ɜ:/ - burn, firm (long mid-central spread vowel); this may also be shown by the symbol /ə:/.
- /ə/ - about, clever (short mid-central spread vowel); this is sometimes known as *schwa*, or the neutral vowel sound – it never occurs in a stressed position.
- /ʌ/ - cut, nut (short low front spread vowel); this vowel is quite uncommon among speakers in the Midlands and further north in Britain.

3.3 Back vowels

- /u:/ - wool, glue (long high back rounded vowel)
- /ʊ/ - put, soot (short high back rounded vowel); also shown by /u/
- /ɔ:/ - corn, call (long mid back rounded vowel) also shown by /o:/
- /ɒ/ - dog, rotten (short low back rounded vowel) also shown by /o/
- /ɑ:/ - hard, far (long low back spread vowel)

The vowels can also be arranged in a table or even depict them against a cross-section of the human mouth. Here is an example from the IPA chart (revised 2005.)



Note fro els while /u/, /o/, /ɔ/ and /ʒ/ are represen epresented on the chart is /ɑ:/.

4. Consonants

In the formation of consonants, the air stream through the vocal tract is obstructed in some way. Consonants can be classified according to the place and manner of this obstruction. Some of the possible places and manner of articulation are indicated below:

- *Glottal articulation*: - articulation by the glottis. This is used for one consonant in English. This is /h/ in initial position in *house* or *hope*.

- *Velar articulation* – this is done with the back of the tongue against the velum. It is used for initial hard /g/ (as in *go*) and for final /ŋ/ (as in *sing*).

- *Palatal articulation* – this is done with the front of the tongue on the hard palate. It is used for /ʤ/ (as in *jam*) and for /ʃ/ (as in *sheep* or *sugar*).

- *Alveolar articulation* – this is done with the tongue blade on the alveolar ridge. It is used for /t/ (as in *teeth*), /d/ (as in *death*) /z/ (as in *zebra*) /n/ (as in *no*) and /l/ (as in *light*).

- *Dental articulation* – this is done with the tip of the tongue on the back of the upper front teeth. It is used for /θ/ (as in *think*) and /ð/ (as in *that*). This is one form of articulation that we can observe and feel ourselves doing.

- *Labio-dental articulation* – this is done with the lower lip and upper front teeth. It is used for /v/ and /f/ (as in *vampire* and *fire*).

- *Labial articulation* – this is done with the lips for /b/ (as in *boat*) and /m/ (as in *most*). The two lips are used (as in English) this is bilabial articulation.

This scheme gives us a different arrangement into stop (or plosive) consonants, affricates, fricatives, nasal consonants, laterals and approximants.

- Stop consonants (so-called because the airflow is stopped) or plosive consonants (because it is subsequently released, causing an outrush of air and a burst of sound) are:
  - Bilabial voiced /b/ (as in *boat*) and voiceless /p/ (as in *post*)
  - Alveolar voiced /d/ (as in *dad*) and voiceless /t/ (as in *tap*)
  - Velar voiced /g/ (as in *golf*) and voiceless /k/ (as in *cow*)
- Affricates are a kind of stop consonants, where the expelled air causes friction rather than plosion. They are palatal /ʧ/ (as in *cheat*) and /ʤ/ (as in *jam*).
- Fricatives come from restricting, but not completely stopping the airflow. The air passes through a narrow space and the sound arises from the friction this produces. They come in voiced and unvoiced pairs:
  - Labio-dental voiced /v/ (as in *vole*) and unvoiced /f/ (as in *foal*)
  - Dental voiced /ð/ (as in *those*) and unvoiced /θ/ (as in *thick*)
  - Alveolar voiced /z/ (as in *zest*) and unvoiced /s/ (as in *sent*)
  - Palatal voiced /ʒ/ as in the middle of *leisure*) and unvoiced /ʃ/ (as at the end of *trash*)
- Nasal consonants involve closing the articulators but lowering the uvula, which normally closes off the route to the nose, through which the air escapes. There are three nasal

consonants in English:

- Bilabial /m/ (as in mine)
- Alveolar /n/ (as in nine)
- Velar /ŋ/ (as at the end of sing)

- Lateral consonants allow the air to escape at the sides of the tongue. In English there is only one such sound, which is alveolar /l/ (as at the start of *lamp*)
- Approximants do not impede the flow of air. They are all voiced but are counted as consonants chiefly because of how they function in syllables. They are:
  - Bilabial /w/ (as in *water*)
  - Alveolar /r/ (as in *road*)
  - Palatal /j/ (as in *yet*)

The following chart adopted from Jones (2003) shows a summary of the manner and place of articulation of each of the consonant sounds

MANNER	**PLACE**							
	Bilabial	Labiodental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					ʧ ʤ			
Fricative		f v	θ ð	s z	ʃ ʒ			h
Nasal	m			n			ŋ	
Lateral				l				
Approximant	w			r		j		

It should be noted that English allows a cluster of up to three (3) consonants in syllable-initial position (usually called Onset) and up to four (4) consonants in syllable-final position (usually called Coda). It should be noted further that though the clustered consonants are realised together, the sound of each individual segment could be clearly discerned. It is again worthy of note that it is not every occurrence of more than one consonant word initially, medially or finally that counts as a cluster; hence, the double <p> in *apple*, the double <n> in *beginning*, the double <g> in *beggar*, the two consonant sequences in words such as *psychology, knife, pneumonia, sword, physics, mnemonic*, etc. are not consonant clusters because they are realised as single sound units.

The foregoing explanation is also relevant to the digraphs such as <th>, >sh>, <ph>, <kn>, <gn>, <gh>, etc. In words like *three, shine, philosophy, knife, gnash, laugh*, etc. the digraphs are pronounced as single sounds rather than as consonant clusters. The following summarises English syllable requirements:

Onsets of English syllables according to Hammond (1994, p.53) may have the following structure: “the onset can contain between zero and three positions”. Zero means a syllable

without an onset, such as one that begins with a vowel, e.g. 'air' [eə]. From one to three is illustrated thus:
C – a single consonant followed by a vowel, e.g. 'key' [ki].
CC – a cluster comprising two consonants followed by a vowel, e.g. 'cry' [kɹɑɪ].
CCC – a cluster of three consonants, followed by a vowel, e.g. 'spray' [spɹeɪ].
The above schema is systematic because the phonotactics of English allows in the first instance, any consonant as the onset of a syllable to be followed by a vowel except /ŋ/ and /ʒ/ (Gimson, 1970). It should however be noted that in some English words of French origin such as 'genre', the onset of the first syllable begins with /ʒ/.
In a CCV structure, the fricative /h/, the affricates /ʧ/ and /ʤ/, as well as the sonorants /m, n, ŋ, I, r, w/ are not allowed in the initial position, whereas syllables of the CCCV type must always begin with /s/ (Hammond, 1999). It should be added however that in the later sequence, the next consonant after /s/ must be one of the voiceless plosives /p, t, k/ to be followed by either /l/ or /r/. Hammond (1999) further explains that all the consonants of English can appear in the syllable coda position. Except /h/ It should be noted however, that in Received Pronunciation (RP), which is a non-rhotic variety of English, in addition to /h/, /r, w, j/ do not also occur in coda positions. Words ending in these consonants such as 'car' [kɑ:], 'star' [stɑ:], 'window' [wɪndəʊ], 'follow' [fɒləʊ], 'money' [mʌnɪ] and 'my' [mɑɪ] only have these consonants in the orthographic forms of the words. Harris (1994, p.53) observes that the coda in English can have from zero to four consonants, with the condition that word-medial codas are restricted to two consonants. Codas may thus be characterised as follows:
C – a single consonant, e.g. 'cat' [kæ**t**]. Any consonant except /h, r, w, j/.
CC – two consonants, e.g. 'crank' [kɹæ**ŋk**]. The second consonant must be an obstruent.
CCC – three consonants, e.g. 'prompt' [pɹɒ**mpt**]. The last two consonants must be obstruent.
CCCC – four consonants, e.g. 'prompts' [pɹɒ**mpts**]. The last (fourth) consonant must be /s/.

5. Conclusion

Having come this far, it is important to state that sounds can be looked at in different ways. On the one hand, we can think of them as individual segments that follow one another in the chain of speech like beads on a string; this is the view implied by normal phonetic transcription, which is the most useful way of recording sounds for practical purposes. On the other hand, we can think of each sound, not as an indivisible unit, but rather as a bundle of features. Each feature is the consequence of an articulator movement of the speech organs, has characteristic acoustical properties that can be recorded by instruments designed for the purpose and produces its own special effect on the ear of the listener. This, in the view of this paper, should be the main concern of individuals whose professional demands include communicating to Nigerians through the medium of television.

References

Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh University Press.

Crystal, D. (1988) *The English Language*. Penguin Group.

Gimson, A. (1970). *An introduction to the pronunciation of English*. 3rd edition. Edward Arnold.

Hammond, M. (1999). *The phonology of English: A prosodic Optimality-Theoretic approach*. Oxford University Press.

Harris, J. (1994). *English sound structure*. Blackwell.

Jones, D (2003) *Cambridge English pronouncing dictionary*, 16th ed. P. Roach, J. Hartman and J. Setter (eds.) Cambridge: Cambridge University Press.

Lyons, J. (1981) *Language and Linguistics*: *An introduction*. Cambridge University Press.

Sapir, E. (1921) *Language*: *An introduction to the study of speech*. Harcourt, Brace.