William Holder – A Pioneer of Phonetics

Angelika Braun

University of Trier, Germany
brauna@uni-trier.de

Abstract: This paper looks at William Holder's "Elements of Speech" [1] from the perspective of modern phonetic and phonological literature. The structure of this presentation, first touching on general linguistic concepts, follows the three principal stages of speech production: respiration/air stream mechanism, phonation, and articulation. Holder's rendition of speech perception is discussed in conjunction with the clinical application of teaching a deaf-mute to speak. Furthermore, the passage on what is today called prosody will be reviewed. The mere fact that this topic is addressed in Holder's work is remarkable since suprasegmentals were not normally an issue at all in 17th century publications on speech. The extent and precision of the description in the "Elements of Speech" is truly remarkable. A comparison with the writings of Wolfgang von Kempelen, who is often referred to as the pioneer of modern day phonetic theory, shows that Holder knew much more about speech and hearing than Kempelen even though his writings have received much less attention.

1 Introduction

The history of what is today called phonetics has received relatively little attention. As Abercrombie [2: 1] puts it, "There is, in fact, a firm opinion in some quarters that before about 1830 there was no such thing as phonetics." However, while the term "phonetics" may not have been around at that time, the working principles of the speech and hearing mechanisms were certainly described in various contexts: anatomical, orthoepic, linguistic, and clinical, to name only a few. Thus, the writings of those early days of speech analysis can be assessed in view of present-day knowledge.

This contribution focuses on a 17th century publication by William Holder entitled Elements of Speech. An Essay of Inquiry into the Natural Production of Letters: with An Appendix Concerning Persons Deaf & Dumb. So far, the focus of the scientific discussion on Holder's writings has been on his musicological works and, as far as speech is concerned, writings which are related to teaching a deaf-mute speaker to talk [3, 4]. Some authors emphasize his contribution to general phonetic knowledge [5, 6, 7], but somehow the full scope of the modernism of his approach does not seem to have been recognized.

Instead, Wolfgang von Kempelen has been credited with pioneering work on the analysis and synthesis of the speech process: "Den Grundstein dazu [i.e. die für das Zustandekommen eines Lautes nötige subtile Arbeit in allen ihren Einzelheiten zu ergründen, A.B.]" legte 1791 v. Kempelen, der durch seine erstaunliche Analyse und Synthese des Sprechmechanismus diese neue Richtung der phonetischen Forschung eröffnete, eine Leistung, die, fachgeschichtlich betrachtet, noch immer nicht gebührend eingeschätzt wurde [...]"1 [8: 161].

1 The foundation for this (i.e. for establishing the subtle detail which is required to constitute what is called sound) was set by Wolfgang von Kempelen [9], who, by way of his amazing analysis and synthesis of the
William Holder was born in 1616 and studied at Pembroke Hall, Cambridge, where he received an M.A. and became a fellow in 1640. He received his Doctor of Divinity from Oxford in 1662 and was elected Fellow of the Royal Society in 1663. He is primarily remembered as a musicologist, having published "A Treatise on the Natural Grounds and Principles of Harmony" in 1694.

A controversy arose within the Royal Society about his claim to have taught a deaf-mute called Alexander Popham to speak in 1659. That is why his book, which the present contribution is based on, contains an appendix "Concerning Persons Deaf and Dumb", in which he describes the method employed. In order to better understand the relevance of this issue it has to be pointed out that teaching a congenitally deaf person to speak seems to have been a major issue in the 17th century which was addressed in several monographs (cf. e.g. Amman [10], Bulwer [11]). The case of Alexander Popham evidently became the subject of a bitter dispute within the Royal Society between Holder and John Wallis, possibly because Popham suffered a relapse and was subsequently treated by John Wallis, who then claimed responsibility for the success [3: ii]. Regardless, Holder as well as Wallis based their teaching concepts on their respective understanding of the speech production mechanism. Firth [13: 115] insinuates that the publication of Holder's work was sabotaged by his rivals, possibly because it was more "modern" than they deemed appropriate.

In antiquity, the Greek physician Galenus was one of the very few who gained detailed knowledge about voice and speech production by dissecting human cadavers [8: 201]. During the Middle Ages, a ban on autopsies was quite rigorously enforced by the Catholic Church. This prevented researchers from exploring the anatomical bases of speech production and resulted in rather gross ideas about them [8: 73; 202]. During the Renaissance, scholars like Michelangelo Buonarotti and Leonardo da Vinci (secretly) performed autopsies in Italy and thus gained insight into the anatomy and physiology of the voice and speech production mechanisms (cf. the da Vinci anatomical sketches and [15: 42].

Thus, 16th and 17th century scholars as e.g. Amman [16] and Holder had relatively little to go on for their descriptions of speech production. It is therefore commendable that Holder would make the distinction between respiration, phonation and articulation, which is common today but was by no means well-established in the 17th century. If we compare Holder's writings to other 17th or even 18th century literature (e.g. Christopher Cooper, Alexander Hume, Petrus Montanis, or Francis Lodwick, cf. [5: 2; 12: 357]), we cannot but admire his insights and
the "modern" approach to the subject matter (see also [5]). This is an asset which is emphasized by Firth [14] who states that "[h]is descriptions of the sounds are quite good even judged by present-day standards" (p. 115). The following summary of Holder's work is organized along the lines of a present-day textbook of phonetics.

4 Holder on Letters and Sounds (and the Concept of Phonemes)

At first glance, it seems that – very much in keeping with other writings of the time – Holder confounded the concepts of letters and speech sounds. This then would have constituted one of the few notions in Holder's writings which are no longer generally acceptable today. Abercrombie [5: 317], however, points out the ambiguity of the term "letter" and traces it back to the 17th century: "'An Element of Speech, which is commonly call'd a Letter, hath a double signification' – it could be something you hear or something you see" [17: 1].

Holder does in fact draw a clear distinction between "letters" on the one hand and "symbols" or "characters" on the other (p. 12): "[...] the incongruous pronunciations of several Letters, as they lie described to the Eye by Symbols or Characters of the Alphabet of several Languages, which indeed ought to be only one [...]"(ibid.). This translates his "letters" into constituting what are today called "sounds" (or even phonemes) and his "symbols" or "characters" constituting what are today called "letters". He later talks about "Signs Audible" and "Signs Visible" and calls for alphabets in which "[c]haracters or Signs written were exactly accommodated to Speech" (p. 108), which makes it clear that Holder was well aware of the difference between sounds and letters.

Further evidence for a phonemic perspective can be found later in his book, when Holder writes: "There is so much space between a and e, that there may be a vowel inserted between them, and a fit character for it may be e, and perhaps some Languages may have a distinct use of such a vowel" (p. 81; emphasis mine, A.B.). The concept of distinct use of vowels, which is language-specific, clearly points to the phonemic principle. Incidentally, the Ash-symbol (A-E ligature) [æ] has made its way into the International Phonetic Alphabet, denoting precisely the vowel quality which Holder described.

5 Holder on Speech Production

Today, it seems so perfectly natural to talk about speech production in terms of respiration/air stream mechanisms, phonation, and articulation ([18, 19, and 20], to cite only a few) that we have become oblivious of the fact that this descriptive framework seems to have been lost in the 18th and early 19th centuries. It would seem that only in the 19th century did the knowledge about speech production fully return [15: 202; 21: 6].

5.1 The Source-Filter Theory of Speech Production

This theoretical concept is one of the true basics in speech production which can be found in any current textbook. This model is generally attributed to Gunnar Fant [22] who described the speech signal as the complex output of laryngeal activity (source) and supralaryngeal processes (filter).

Interestingly enough, Holder draws a similar distinction: He talks about "material" and "formal" causes of what he calls letters (i.e. sounds, see above): "Their Matter is various; viz.
Breath, or voice, *i.e.* Breath vocalized by the operation of the *Larynx*. Their *Form* is constituted by the Motions and Figures of the Organs of speech, affecting the Breath or Voice with a peculiar sound, by which each Letter is discriminated" (p. 64).

This distinction between matter and form, where "matter" refers to laryngeal activity and "form" to supralaryngeal activity, clearly reflects the distinction between source and filter as outlined in Fant [22]. The description provided by Holder is, of course, not based on acoustical data, but its analytical precision is still remarkable.

### 5.2 Respiration

Holder limits his remarks on the role of the lungs to a short description: "The *Lungs* are as *Bellows*, which supply a force of Breath: the *Aspera Arteria* is as the nose of *Bellows*, or as a channel in the sound Board of an Organ, to collect and conveigh the Breath, and somewhat more, by a power of contracting and dilating it self, which those have not" (pp. 22-23). This can be taken to imply that Holder is exclusively considering an egressive pulmonic airstream mechanism. This is in keeping with the focus on the sounds of English, which phonologically makes exclusive use of egressive pulmonic speech sounds. The description of the lungs as bellows has prevailed to this day (cf. Raphael et al. [23: 56).

### 5.3 Phonation

Holder is obviously well aware of the fact that the process of voicing happens in the larynx. However, he fails to recognize the nature of the various tissues involved in voice production. Instead, he considers "a vibration of [...] Cartilaginous Bodies" to "form[s] that Breath, into a Vocal sound or Voice [...]" (p. 23). Holder thus was oblivious to the essential role of the vocal folds in voice production. Subbiah [7], in expressly commending the fact that Holder "knew that the vocal chords were 'cartilaginous bodies''" (p. 174), fails to notice this inaccuracy in his rendition of the phonation process. – His description of the vocal folds as being cartilaginous, which is, of course, partly, but not entirely accurate, does not keep Holder from focussing on the sole difference between, e.g., /p, t, k/ and /b, d, g/ as one between "Articulations of Breath [...] and Articulations of Voice, or Breath vocalized" (p. 38). He goes into quite a bit of detail, characterizing /p/ as "wholly *Mute*, because it is nothing but Breath stopt" (pp. 38-39) and /b/ as accompanied by "a *murmuring* sound of the Voice, formed in the *Larynx*, and passing till it be stopt by the Appulse of the Lips" (p. 39).

### 5.4 Articulation

Holder's definition of "Articulation" is a very modern one: "By *Articulation* I mean a peculiar Motion and Figure of some parts belonging to the Mouth between the Throat and Lips, whereof some are more ease to be discerned and described [...]. Most [difficult] of all [are] the Vowels, where there are peculiar Figures of the Cavity of the Mouth between the Tongue and the Arch of the Palate [...]" (p. 27). He then proceeds to classify vowels as opening sounds and consonants as sounds involving some kind of obstruction:

---

3 For a detailed discussion of these concepts cf. Abercrombie [21].
4 Abercrombie [21] equates the notion of "cartilaginous Bodies" with "the sides of the larynx" (p. 4) and thus seems to accept it as an adequate description of voice production.
"[...] That in all Vowels the passage of the mouth is open and free, without any appulse\(^5\) of an Organ of Speech to another: But in all Consonants, there is an Appulse of the Organs, sometimes (if you abstract the Consonants from the Vowels) wholly precluding all sound; and in all of them, more or less, checking and abating it" (p. 29). This could well be a description contained in a 21\(^{st}\) century textbook.

5.4.1 Consonants

Holder classifies the consonants by
- articulation of voice vs. articulation of breath (i.e., voiced vs. voiceless [p. 23])
- appulse (plenary/occluse/close vs. partial/pervious [p. 36])
- place of articulation (pp. 37-39).

This, of course, reminds the reader of Abercrombie's three-term labels [18: 52] and the classification of consonants in basically every textbook of our time. – In this section, Holder discusses the distinction between nasal and oral consonants in a highly specified way which is totally acceptable from a present day point of view. First, he mentions the various places of articulation and the sounds thereby made. He then proceeds to specify the difference between (oral) stops and nasals:

"Thus the same Articulation; if of Breath, makes one letter; if of Breath vocalized, or voice, another; If of voice Nasall (i.e. when the Uvula is opened, and the voice passeth into the Mouth, and is there Articulated, and at the same time hath a free passage through the Nose) then it makes another; and lastly, if of Breath Nasal, then another" (pp. 33-34).

These categories may well be interpreted within the framework of distinctive feature theory, vocal vs. spirital corresponding to +/- voiced, and naso-vocal vs. naso-spirital corresponding to +/- nasal.

The remaining descriptions are non-binary and look very similar to those provided by Ladefoged [19] or Catford [20]. The following places of articulation are mentioned: labial, labiадental, linguadental, gingival, and palatick. These roughly correspond to the modern-day labels bilabial, labiodental, dental, alveolar, and velar. The degrees of closure are labelled close appulse (stops), pervious appulse (fricatives), free (approximant), and jarring (tap/trill). Interestingly enough, Holder's consonant chart (Figure 2, see below) lists the full range of speech sounds, filling all of the cells, amounting to a total of 36. He then proceeds to exclude those speech sounds "that prove not graceful, nor easie to be pronounced, viz. 2 Spiritals, 9 Naso-Spiritals, and 6 Naso-Vocals, in all 17; there will remain 19 Consonants, proper for use according to the design of Letters" (p. 66; cf. also p. 59). Thus, Holder assumes 19 consonant phonemes.

From today's perspective it must be noted that /h/, /j/, /w/, and /Ɂ/ are not listed in the chart. A closer look reveals, however, that Holder addresses all except /j/ in the text. The latter is, however, listed alongside [ʃ] in his figure on page 96 (cf. Fig. 2). As concerns the remaining phonemes, Holder defines the glottal stop as "one stop, which may [sic] be made in the Larynx, of Breath, before it comes to the Tongue and Palat" (p. 60). He also gives a reason for not listing it as a phoneme 'letter': "[...] I thought it not worthy to be inserted amongst the Letters" (p. 71).

\(^5\) Meaning "closure" or "narrowing". Holder later introduces "close appulse" to designate plosives and "pervious appulse" to designate fricatives (p. 62).
in that it is applied to Breath immediately as it passeth through the Aspera Arteria, and not to Breath or Voice Oral or Nasal (pp. 72-73).

As far as /h/ is concerned, Holder remarks that "[...] H is onely a Guttural Aspiration, i.e. a more forcible impulse of Breath from the Lungs, applied when we please, before or after other Letters" (p. 67). He concludes that it "[...] cannot properly be called a Letter, according to that description we have made of Letters; yet in that it causes a sensible, and not incommodious Discrimination of Sound, it ought to be annexed to the Alphabet (pp. 68-69).

The phoneme /w/ is mentioned only in passing, i.e. in conjunction with the pronunciation of /h/: "In WHAT, WHICH, and the like, H is pronounced before W. and so of right ought to be written" (p. 72). This evidently refers to the voiceless labial-velar fricative [ʍ].

Figure 1: Consonant chart (Holder, p. 62)

5.4.2 Vowels

Holder's description of the vowel system is more difficult to interpret because it refers to the Early Modern English state of affairs. The pronunciation of the vowel phonemes (and possibly also the phoneme system) were quite different then from what they are now. In an attempt to put Holder's description of vowels into perspective, various sources on the pronunciation of his time were consulted [24, 25, 26].

Holder, when naming the vowels of his time, pursues a similar path as did John Wells [27] with his lexical sets more than 300 years later by citing what he considers to be typical representations of the respective vowels. The following table represents the lexical items as listed by Holder, his notation, and the inferred IPA symbols.
Table 1: Lexical items as listed and transcribed by Holder

<table>
<thead>
<tr>
<th>Lexical Item</th>
<th>Holder's notation</th>
<th>IPA notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td>α</td>
<td>αː/ɔː</td>
</tr>
<tr>
<td>Fate</td>
<td>ä</td>
<td>æː/ɛː</td>
</tr>
<tr>
<td>Seal</td>
<td>ĕ</td>
<td>eː</td>
</tr>
<tr>
<td>Eel</td>
<td>i</td>
<td>iː</td>
</tr>
<tr>
<td>Cole</td>
<td>ò</td>
<td>oː</td>
</tr>
<tr>
<td>Fool</td>
<td>òo</td>
<td>uː</td>
</tr>
<tr>
<td>Rule</td>
<td>u</td>
<td>iuː</td>
</tr>
<tr>
<td>Two</td>
<td>χ</td>
<td>oː</td>
</tr>
<tr>
<td>Folly</td>
<td>ò</td>
<td>o/ɔ</td>
</tr>
<tr>
<td>Fat</td>
<td>ò</td>
<td>æ</td>
</tr>
<tr>
<td>Sell</td>
<td>ò</td>
<td>e/ɛ</td>
</tr>
<tr>
<td>Ill</td>
<td>ò</td>
<td>i</td>
</tr>
<tr>
<td>Full</td>
<td>ò</td>
<td>ð</td>
</tr>
</tbody>
</table>

As previously mentioned (cf. chapter 4 above), Holder proposes to insert a vowel æ between a and e. The examples which are given may not be exhaustive from a present-day perspective, but they represent the state of affairs in the second half of the 17th century.

Holder defines vowels as sounds "made by a free passage of Breath Vocalized [i.e. voicing, A.B.] through the cavity of the Mouth, without any appulse of the organs; the said cavity's being differently shaped by the postures of the Throat, Tongue and Lips, some or more of them, but chiefly the Tongue" (p. 80). He also recognizes that, owing to the continuous nature of vowel articulation, the number of phonetic vowels is unlimited, but that the number of vowel phonemes into which these articulations can be grouped is much smaller: "As to the Number of Vowels, they, being differenced by the shape of the cavity of the mouth, may be reckon'd very many, if small differences be allowed. But those which are remarkably distinguished [emphasis mine, A.B.] [...] may be reduced to these Eight [...]" (pp. 80-81). This clearly points to a phonological approach.

According to Holder, vowels are more difficult to describe than consonants for the following reasons: "The Articulations, that is, the Motions and Postures of the Organs in framing the Vowels, are more difficultly discerned, than those of the Consonants; because in the Consonants, the Appulse is more manifest to the sense of Touching, but in the Vowels it is [...] hard to discern the Figures made by the Motions of the Tongue, (inclining onely toward the Palat, and not touching it)" (p. 82). This description of the problem would easily pass for a 21st century textbook rendition. Still, his concept of vowel production is by far less clear and by far less accurate than that of consonant features.

Some notions contained in Holder's book correspond to features which are used today: For instance, his awareness of the dimensions "tongue height", which he calls "aperture" and "tongue position", which he calls "situation" can be inferred from his characterization of /i/ as "being the closest and forwardest" vowel (p. 89). But that scheme is not pursued in a systematic way, the only indication given by Holder is ordering "the Series of the Vowels according to their degrees of aperture, and recess towards the Larynx" (p. 90). Thus, he does not draw a

---

*No symbol given.*
clear separation between the two. He distinguishes four classes of vowels, two of which represent places of articulation and two describe the role of labiality (cf. Figure 2).

Figure 2: Chart of consonants and vowels with respect to voicing and nasality (Holder, p.98)

The places of articulation for vowels are "guttural" and "palatic", which would translate into front and back in modern terms, and the assignments of the different vowels to these places correspond to the present-day understanding of the distribution. As far as labiality is concerned, the description is less clear: Whereas in his (o), which is also [o:] in terms of IPA, he mentions the role of the lips as "drawing them a little rounder, [which] helps to accomplish the pronunciation of it" but he still concludes that this "is not enough to denominate it a Labial Vowel, because it receives not its Articulation from the lips" (p. 86). On the other hand, he lists \( \mathbf{u} \) and \( \mathbf{u} \) as "peculiar, [in] that they are framed by a double motion of Organs, that of the Lip, added to that of the Tong; and yet either or them is a single Letter, and not two, because the motions are at the same time, and not successive [...]" (p. 88). This clearly denotes these vowels as labial. Some of his further remarks, however, are much less clear: "Thus \( \mathbf{u} \) will be onely \( \mathbf{i} \) Labial, and \( \mathbf{u} \) will be \( \mathbf{oo} \) Labial, that is, by adding that motion of the under-Lip, \( \mathbf{i} \) will become \( \mathbf{u} \), and \( \mathbf{oo} \) will become \( \mathbf{u} \)" (p. 90). It seems unlikely that this description is explicable by the vowel pronunciation of his time; instead, a more likely cause for this rendition is a misconception of the interplay between tongue position and the role of vowel rounding in general.

On the other hand, it is remarkable that Holder recognizes the physiological and linguistic possibility of nasal and voiceless vowels, but he does not seem to make the connection to their linguistic use in any particular language such as French or Polish. Instead, he classifies them as "uneasie and unpleasant" (p. 98).
It is evident from Figure 3 that Holder ventures to incorporate vowels and consonants into one single chart. The vowels in this chart are distinguished from other sounds by what would today be called the feature +/- approximant [28]. This means that to some extent at least he applies descriptive parameters to consonants and vowels alike, once again a very modern concept of looking at speech sounds.

In addition to the vowels, Holder also mentions the notion of diphthongs. However, he does not regard the combination of e and o with a as "proper Diphthongs"(p. 94), whereas the i in stile "seems to be such a Diphthong [...] composed of a.i or e.i. and not a simple Original Vowel" (p. 95). These observations are difficult to interpret. It could be taken to imply that he does not recognize /r/-vocalization as a "proper diphthong", but this is not made explicit in his writings.

5.4.3 Syllable Structure and Prosody

In conjunction with the concept of syllable, Holder comments on phonotactics. He states that there has to be "one Vowel in every Syllable, for varieties sake, sometimes preceding, sometimes following, and sometimes interposed between the Consonants" (p. 92). He fails to note, however, that while this holds true for English, it is by no means a language universal.

When it comes to what are today considered prosodic features, Holder distinguishes between "emphasis" (intensity and/or duration) and "accent" (pitch): "There some other Accidents besides those spoken of before, which have an Influence in varying the Sound of Languages, as Accent and Emphasis; which though now much confounded seem to have been formerly more distinguished. Accent, as in the Greek names and usage, seems to have regarded the Tune of
the voice; the *Acute* accent raising the Voice in some certain Syllables to a higher, *i.e.* more acute Pitch or Tone, and the *Grave* depressing it lower, and both having some *Emphasis, i.e. more vigorous pronunciation*" (pp. 98-99).

He proceeds to discuss between-language differences in this respect between French and Ancient Greek. This, once again, is a very modern approach to comparative phonetics.

### 6 Holder on Speech Perception and Clinical Issues

In the "APPENDIX Concerning Persons DEAF AND DUMB", Holder touches upon the physiology of speech perception. He is well aware of the close interaction between speaking and hearing and of the fact that congenitally deaf individuals who do not receive treatment will not learn to speak: "[...] the Tong and the Ear, Speaking and Hearing, hold a correspondence, by which we learn to imitate the Sound of Speech, and understand the meaning of it" (p. 114). "[...] they who want that Sense of Discipline (Hearing) are also by consequence deprived of Speech [...]" (p. 115).

His understanding of the peripheral hearing mechanism seems somewhat limited. He regards a lax tympanic membrane as the principal cause of hearing loss/deafness. This in turn is caused by either a malfunction of the incus and malleus or laxness of the m. tensor tympani. "And I am of opinion, that the most frequent cause of Deafness is to be attributed to the Laxness of the *Tympanum*, when it has lost its Brace or Tension by some irregularity in the Figure of those Bones, or defect in that Muscle" (p. 113). His explanation for bone conduction is rather odd: "Now that which I would infer, is, That in those [i.e. deaf individuals, A.B.] generally the Auditory Nerve is sound, and by a branch of the same Nerve, that goes between the Ear and the Palat of the Mouth, they can make a shift to hear themselves, though their outward Ear be stopt by the Laxe Membrane to all Sounds, that come that way" (p. 129).

Holder evidently had intuitive knowledge of how to distinguish between peripheral and central causes of hearing loss: He asked the patient to hold a lute string between his teeth and, when he perceived sound originating from it, concluded that the auditory nerve was not damaged (p. 160).

The following remark points to Holder being well aware of the role of redundancy in speech perception: "Any *Equivocal* word spoken *alone*, cannot be determined to any one certain Sense and Signification by him that hears it; of which there are numerous examples in every Language: Yet the same word in *Connexion* of Speech, as part of a sentence, is understood as easily as any other" (p. 122).

### 7 Conclusion

At a time when awareness of the processes of speech production and speech perception were generally quite limited, individual scientists like William Holder display in their publications an amazing degree of "modern" thinking both with respect to general concepts like the source-filter model of speech production or the phoneme and the phonetic detail in the description of sound production. The degree of precision in the description of consonants as well as suprasegmentals well exceeds that of vowels. This contribution attempts to exemplify the import of his writings on general phonetics from today's perspective. The full scope of his
findings – especially with respect to the teaching of deaf-mute patients – has yet to be analyzed in detail.

8 References


