

# PHONETIC NOTATION SYSTEMS

Jo Verhoeven & A. Hirson

City, University of London, Division of Language and Communication Science,  
Phonetics Laboratory

This is a slightly updated version of an article which was published as: Verhoeven, J., Hirson, A. (2009). Phonetic notation systems. IN J. Verschuren (Ed.) Handbook of Pragmatics, vol. 13, pp.1-14.

## 1. INTRODUCTION

Phonetic notation systems have long been used as an essential tool in the phonetic description of languages of the world and also more generally in other branches of the linguistic sciences. In this article, one of the most widely used phonetic transcription systems will be discussed in detail, i.e. the alphabet of the International Phonetic Association. This alphabetic system is firmly rooted in a strong international phonetic tradition which emerged in the last decades of the 1800s. Before going into this transcription system, however, it may be useful to define the subject area of phonetics, which is done elegantly in Abercrombie (1967) by making a rigorous distinction between the concepts of language and medium. Distinguishing language from medium is distinguishing patterns from their material embodiment. "Language, we could say, is form, while the medium is substance" (Abercrombie 1967: 1). In other words, language refers to the underlying abstract patterns and rules of the language, while words, language refers to the underlying abstract patterns and rules of the language, while the medium acts as its vehicle, i.e. it externalizes language in some form mediating between 'speaker' and 'listener'. Linguistics is the discipline that studies language, while phonetics can be appropriately regarded as the study of the medium of spoken language: "The study of the medium of spoken language, in all its aspects and all its varieties, constitutes the subject of phonetics. Phonetics is concerned with the medium as used in speaking all human languages (...), and as used in all styles of speech (whether supposed to be good or bad, normal or abnormal)" (Abercrombie 1967: 2). Abercrombie classifies the different types of media on the basis of the reception activity involved. In spoken language the medium is aural i.e. perceived by the listener's ear. The production activity involved is the use of the articulators to produce speech sounds to be perceived by a listener. In written language the medium is visual or perceived by the eye of a reader. In this case, the medium is the product of fine co-ordinated movements of hands and fingers to produce shapes on paper. Finally, there is the tactile medium perceived via the sense of touch of the receiver. A good example of this is the Braille system in which specific patterns of dots are embossed on paper and 'read' by feeling these patterns. At its most general level, the term notation refers to the conversion process of one medium into another. Phonetic notation specifically transforms the aural medium into a visual or tactile medium to provide an accurate and permanent record of the ways in which speech sounds are produced in languages (articulation). Furthermore, such representation also indicates what these speech sounds sound like (perception). Phonetic notation thus informs both about the speech production and the speech perception

processes involved in the aural medium.

## 2. HISTORICAL TAXONOMIC PERSPECTIVE

Many attempts have been made in the past to develop phonetic transcription systems. A basic distinction that is traditionally made is between alphabetic and analphabetic notation systems. In analphabetic systems, individual speech sounds are represented as strings of symbols in which each symbol explicitly describes a specific articulatory feature of the sound involved. The best-known examples of such phonetic transcription systems are those developed by the Danish linguist and co-founder of the International Phonetic Association Otto Jespersen (1889) and the American phonetician Kenneth Lee Pike (1943). Jespersen's transcription system uses combinations of Greek and Roman letters, numerals, italics, heavy type and, subscript and superscript letters. The Greek letters represent the active articulators involved. Active articulators are those that move during speech production. In Jespersen's system  $\alpha$  refers to the lips,  $\beta$  to the tip of the tongue,  $\gamma$  to the upper surface of the tongue,  $\delta$  to the velum or soft palate including the uvula,  $\epsilon$  to the larynx containing the vocal folds and  $\zeta$  to the organs of respiration. The Roman letters refer to the passive articulators: these articulators are relatively passive and primarily capture place of articulation: labial (a-c), interdental (d), dental (e), alveolar (t), postalveolar (g), palatal (h, i), velar (j), uvular (k), pharyngeal (l). Doubling a letter or bold print refers to a broader place of articulation rather than a single point, while two simultaneous places of articulation (double articulations) are represented by two letters separated by a comma. The numerals, finally, give a conventional indication of the degree of constriction of the vocal tract created by the active and passive articulator: 0 = complete closure, 1 = narrow approximation etc. These symbols can be used to create relatively compact but intractable formulas such as  $\alpha^7 \beta \gamma \delta \epsilon \zeta$  to denote the vowel in English 'all'.

Pike's analphabetic transcription system developed in *Phonetics* (Pike 1943) provides even more detail about speech articulation than Jespersen's and looks even more Byzantine by today's standards. His 'functional analphabetic symbolism' consists of Roman and italic letters in upper and lower case. The italic symbols specify the broad categories of the articulatory mechanisms involved, while the Roman headings express subclassifications. In his notational system the transcription of the voiceless alveolar stop [t] requires 34 symbols: M<sub>all</sub>D<sub>e</sub>C<sub>V</sub>v<sub>el</sub>cA<sub>pp</sub>a<sub>t</sub>d<sub>tl</sub>n<sub>r</sub>a<sub>n</sub>s<sub>f</sub>s<sub>i</sub>F<sub>S</sub>s. In this 'formula', the energy-producing mechanism of the sound is for instance specified by the symbols following left-most M: it indicates an airstream mechanism (a) in which the initiator (I) is the lungs (l). In addition, the direction of the airstream (D) is egressive (e).

Both systems described above have in common that the symbols are arranged horizontally with a front-to-back articulation organisation represented from left-to-right. They graphically illustrate that providing a truly exhaustive articulatory description of speech sounds is impossible: Pike has argued for instance that even his own degree of articulatory detail is "suggestive but by no means exhaustive" (Pike 1943: 153). Whether complete or incomplete, they are highly impractical in everyday use. In Pike's system for example, the transcription of single speech segments such as the alveolar nasal [n] and the half-close back

vowel [o] require 79 and 88 symbols respectively. It is not hard to imagine the explosion of information involved in the transcription of simple words, let alone longer stretches of speech. This is probably the reason why the practical implementation of these transcription systems never took off.

Whereas alphabetic notation systems represent speech sounds by a sequence of symbols, alphabetic transcription systems aim to represent individual speech sounds by single symbols as much as possible. Within such notation systems, a distinction is generally made between Roman-based systems and non-Roman-based (or iconic) systems. In iconic systems the symbols are not chosen arbitrarily, but they are intended to somehow resemble what they represent. This can be illustrated with reference to Alexander Melville Bell's *Visible Speech* which dates back to 1867. In this system the symbol O denotes open glottis which "is pictorial of the expanded breath-channel in the throat" (Bell 1867: 46). This symbol is contrasted with I which stands for "glottis closed for vocal fold vibration". Furthermore, all the consonant symbols are similar in general shape (a variant of C), but the orientation of the opening differs depending on the place of articulation involved. One of the advantages of iconic systems is that they are quite explicit about the articulatory dimensions involved, but their main disadvantage is that they can be difficult to read since many of the symbols are very similar in appearance. Furthermore, they can be difficult to remember unless frequently used. Abercrombie (1967) considered their biggest disadvantage to be that the symbols are often motivated by a particular theory of phonetic description, although arguably this objection may be directed at any system for phonetic transcription.

Roman-based notation systems differ from iconic systems in that the symbols chosen to represent speech sounds are chosen arbitrarily. Probably the most widely-used and most generally-accepted system of Roman-based phonetic notation was developed by The International Phonetic Association or IPA, an organisation which was inaugurated in 1886 as The Phonetic Teachers' Association by a group of French language teachers.

### **3. THE IPA ALPHABET**

The phonetic transcription system of the International Phonetic Association is an alphabetic notation system. This means that it uses a set of symbols which are based on Roman orthography, supplemented in various ways by elements from other symbol sets which are often modified slightly to harmonise with the Roman-based phonetic symbols. In such alphabetic systems, each phonetic symbol represents a composite set of articulatory features. E.g. [p] represents a 'voiceless bilabial plosive' powered by a pulmonic egressive airstream, i.e. information about any particular sound relates to the airflow, whether the vocal folds are vibrating (voiced/voiceless), the place of articulation, and the type (or manner) of articulation. The original design principles underlying this notation system were stated in *Principles* (1948):

1. There should be a separate letter for each distinctive sound; that is, for each sound which, being used instead of another, in the same language, can change the meaning of a word.

2. When any sound is found in several languages, the same sign should be used in all. This applies also to very similar shades of sound.
3. The alphabet should consist as much as possible of the ordinary letters of the roman alphabet, as few new letters as possible being used.
4. In assigning values to the roman letters, international usage should decide.
5. The new letters should be suggestive of the sounds they represent, by their resemblance to the old ones.
6. Diacritic marks should be avoided, being trying for the eyes and troublesome to write. (Principles 1948: 3 of cover).

Although the IPA system is a system for phonetic notation, it is clearly rooted in a phonological perspective: its most basic principle is the relevance of phonological distinctiveness in assigning separate symbols to sounds which function contrastively in languages of the world.

### 3.1 CONSONANT SYMBOLS OF THE IPA

The most up-to-date version of the IPA consists of 79 discrete ‘alphabetic’ phonetic symbols for consonants. The biggest set consists of 69 symbols which refer to pulmonic consonants, i.e. consonants which are powered by a pulmonic egressive airstream mechanism. In these sounds, the airstream is initiated by the respiratory system (the lungs). The relatively large number of symbols in this set as compared to other sets is a natural reflection of the fact that 94% of the sounds in the languages of the world are made on a pulmonic egressive airstream mechanism. These symbols are presented on various charts. In the main chart in Figure 1, the consonants are ordered in cells resulting from the intersection of rows and columns: the rows represent the different ‘manners’ of articulation, while the columns represent the different ‘places’ of articulation. The manners are ordered from more constricted articulations (upper rows) to more open articulations (lower rows).

CONSONANTS (PULMONIC) © 2018 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill				r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

**Figure 1:** IPA chart with the symbols for pulmonic consonants

The sounds involving the strongest constriction of the vocal tract are plosives, nasals, trills and taps/flaps. From an articulatory point of view, all these sounds involve at least in part a stricture of complete closure in the vocal tract so that the airstream is prevented from escaping via the mouth. In a plosive such as [d],

the tip of the tongue touches the roof of the mouth just behind the upper teeth. This traps the airstream inside the mouth so that pressure builds up behind the occlusion. Upon release of the contact between the articulators, the air escapes with a brief explosive burst which is characteristic for the perceptual impression of these sounds. All the languages in the world described to date (100%) have plosives.

The articulation of nasals such as [n] is similar to that of the plosive [d] in that there is a complete obstruction of the airstream in the mouth created by the tongue tip touching the roof of the mouth just behind the upper teeth. This prevents the airstream from escaping via the mouth. In nasals, however, the velum is lowered so that the airstream is directed into the nose along which it escapes. 96% of the languages of the world have nasal speech segments.

In trills such as [r] the tongue tip trills against the passive articulator and as such they are characterized by a sequence of intermittent complete closures of the vocal tract which result from aerodynamic conditions similar to vocal fold vibration. Trills occur in 36% of the languages of the world.

Taps or flaps are characterized by a ballistic action of the tongue. In taps, the tip of the tongue is raised from its neutral position and is thrown in a ballistic movement against the roof of the mouth to make a very short closure which interrupts the airstream. In flaps, the tip of the tongue is curled back inside the mouth and then thrown forward back into its neutral position. During this movement, the tongue flaps against the roof of the mouth to momentarily obstruct the airstream. Both taps and flaps are different from stops in terms of the ballistics involved and the inherently short duration of the closure. These sounds occur in 35% of the languages of the world.

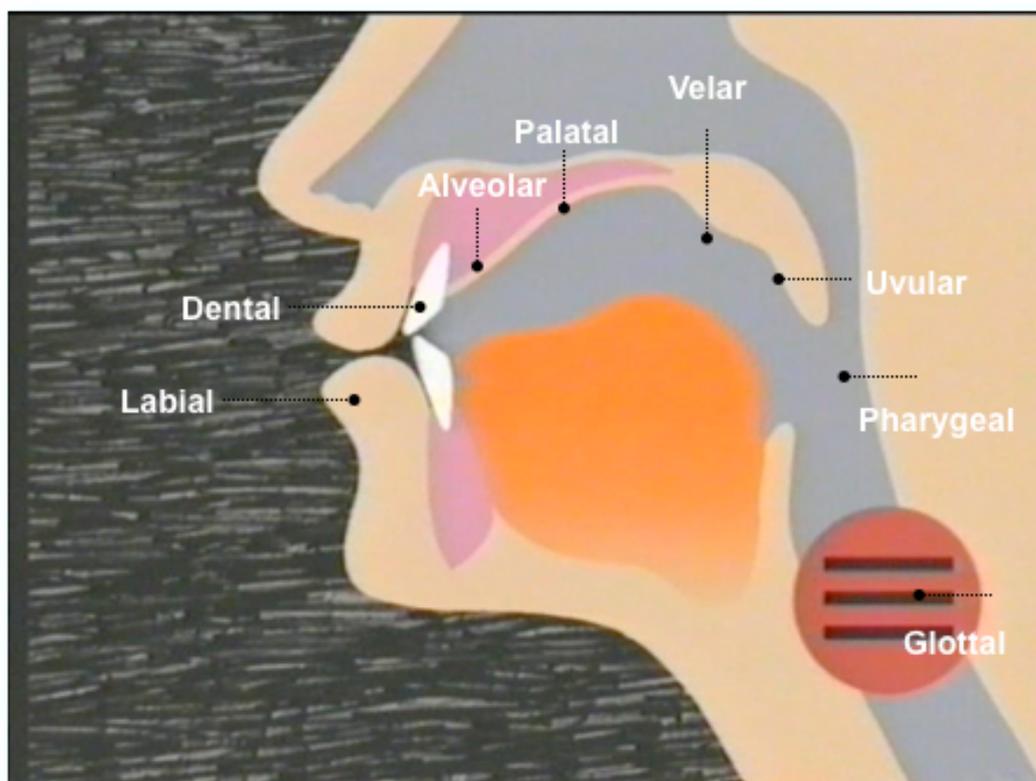
Fricatives such as [f], [s] and [ʃ] are made on a stricture of close (i.e. incomplete) approximation: this means that the vocal tract is narrowed to a considerable degree by the articulators involved. During the articulation of these sounds, the airstream is forced through the remaining narrow gap, which causes an audible hiss. 93 % of the languages in the world have fricatives. In lateral fricatives, the gap is located sideways so that the air escapes laterally.

The approximants are produced on a relatively open stricture of the IPA places of articulation: the gap between the articulators is wider than in fricatives, which results in a minimal obstruction of the airstream so that no turbulence or hiss is generated. 96% of the languages in the world have approximants. This IPA category is further subdivided into central and lateral approximants. The lateral approximants such as [l] are characterized by placing the tongue as an obstacle against the medial part of the roof of the mouth so that the air is forced to flow around it.

In passing, it should be noted that some of the IPA terminology referring to manner of articulation clearly reflects a perceptual approach to phonetic classification, rather than a purely articulatory one. The term 'plosive' for example elegantly captures the perceptual impression of explosion of these sounds upon release of the articulators and in this sense it is different from a

term like 'stop' which refers more to the effect on the airstream of complete vocal tract constriction. The effect of this perceptual orientation is that the IPA classification sometimes highlights the perceptual differences between sound classes rather than articulatory similarities. This is not always useful and may obscure interesting similarities between sound categories. For example, the IPA system treats plosives and nasals as two quite distinct sound categories. There is no doubt that the perceptual difference between these categories is quite substantial. Nevertheless, such classification obscures the fact that from an articulatory point of view plosives and nasals are actually very similar in that both are made on a stricture of complete closure in the oral cavity (Laver 1994).

Apart from the different constriction types, the IPA distinguishes between 11 places of articulation which are ordered in Figure 1 from front articulations such as bilabial (to the left) to back articulations such as glottal (to the right). These places of articulation are defined by specifying the passive articulator involved and are illustrated in Figure 2:



**Figure 2:** illustration of the IPA place of articulation labels. The individual labels are explained in Table 1, in which the percentages refer to the frequency of occurrence of the corresponding sounds in languages of the world.

**Table 1:** description of the IPA places of articulation. Occurrence indicates the percentage of 451 UPSID languages with this place of articulation.

Place	Description	Occurrence
Bilabial or labial	The articulators are the lower and upper lip for sounds like [b] and [m] in 'bat' and 'mat'.	99.78 %
Labiodental	The lower lip approaches the upper teeth, for sounds like [v] in 'veal'.	45%
Dental	The tongue tip approaches the upper teeth as for [ð] in 'this'.	35.03 %
Alveolar	The tongue blade approaches the bony ridge just behind the upper teeth. This creates sounds like [t] as in 'tell'	63.6 %
Postalveolar	The tongue blade moves towards the area between the alveolar ridge and the hard palate. This gives rise to sounds like [ʃ] in 'shield'.	64.08 %
Retroflex	The tip of the tongue is curled back and is raised towards the hard palate. The tongue typically adopts a concave shape.	20.8 %
Palatal	The mid section of the tongue is lifted towards the hard palate. The tongue typically adopts a convex shape.	89.0 %
Velar	The back of the tongue approaches the velum as for [g] in 'go'.	99.6 %
Uvular	The back of the tongue is raised towards the uvula, such as in a uvular trill [R]	18.0 %
Pharyngeal	The tongue root moves towards the pharyngeal wall	4.21 %
Glottal	Involves a constriction of the glottis (e.g. glottal stop [ʔ])	77.50 %

In terms of the use of IPA symbols, it can be noted that for all manners of articulation except fricatives, there are no separate symbols for dental and postalveolar places of articulation and as such these sounds are left underspecified: However, the full specifications can easily be derived from the alveolar symbols by means of diacritics (see further).

A third organisational principle underlying this main consonant chart in Figure 1 is that within each cell of the table, voiceless segments with the same manner and place of articulation are conventionally positioned left, while voiced segments appear on the right. In broad outline, voiceless sounds are produced without vibration of the vocal folds, while voiced sounds do have vocal fold vibration. It is known that all languages of the world use voicing distinctions linguistically, but the proportion of voiceless and voiced speech sounds in sound inventories may vary substantially. Catford (1977) has calculated the voiced-to-voiceless ratio for various languages. In European languages the number of voiced speech sounds tends to be substantially bigger than the number of voiceless sounds, giving rise to relatively high voiced-to-voiceless ratios (French: 3.55; English: 2.57), while the reverse is true for an Asian language such as

Canton Chinese (0.70).

Finally, it should be mentioned that the chart in Figure 1 contains a significant number of shaded and empty cells. The shaded cells on the chart represent sound categories which are deemed physiologically impossible, while the empty slots refer to sound classes which do not have separate symbols. However, in many cases the symbols for these sounds can be easily derived by adding diacritics to existing symbols. A good example of this is the bilabial approximant which occurs in intervocalic positions in Spanish: the symbol for this sound can be derived by using the alphabetic symbol for the voiced bilabial fricative [β] and adding a diacritic to indicate that the manner of articulation is more open, i.e. approximant. Hence, the symbol [β̞] is obtained.

Besides the symbols for pulmonic consonants, the IPA has 11 symbols to represent consonants produced on airstream mechanisms other than pulmonic: these are listed in a separate table which is illustrated in Figure 3:

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
◌͡ Bilabial	ɓ Bilabial	' Examples:
◌̄͡ Dental	ɗ Dental/alveolar	p' Bilabial
◌̄͡! (Post)alveolar	ɟ Palatal	t' Dental/alveolar
◌̄͡≠ Palatoalveolar	ɠ Velar	k' Velar
◌̄͡   Alveolar lateral	ʄ Uvular	s' Alveolar fricative

**Figure 3:** IPA chart with the symbols for non-pulmonic consonants

Ejectives are powered by a laryngeal (glottalic) egressive airstream mechanism. The initiator of this airstream mechanism is the larynx: the vocal folds are adducted to close the glottis and the larynx as a whole is pulled upwards in a piston-like action. This creates positive pressure above the vocal folds and the air flows outward. Ejective sounds are indicated by adding an apostrophe to the corresponding pulmonic symbol. It should be noted that ejectives need a rather tight constriction of the vocal tract so that only plosives, fricatives and affricates can be changed into ejectives by means of a laryngeal airstream mechanism. Ejectives are by definition voiceless. 15.08% of the world's languages have ejective plosives, 13.08 % have ejective affricates and only 2.22% have ejective fricatives.

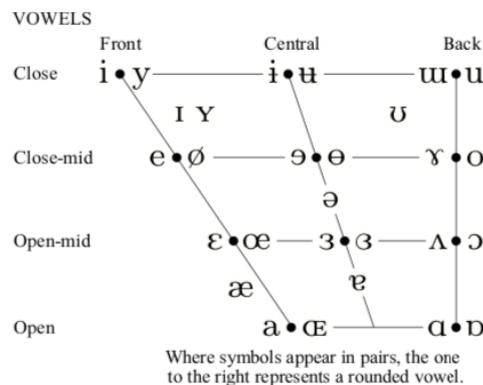
Voiced implosives are produced on a combined airstream mechanism, i.e. laryngeal ingressive and pulmonic egressive. As far as the laryngeal ingressive component is concerned, the vocal folds are adducted and the larynx as a whole is pulled downwards. This creates negative pressure above the vocal folds and air flows into the oral cavity to compensate the resulting pressure difference with atmospheric pressure. The voicing arises from the fact that the vocal folds are not firmly pressed together and as a result of the lowering larynx over a

static column of subglottal air, a small amount of air may leak upward causing a few cycles of vocal fold vibration. The symbols for voiced implosives all have an upward-pointing right hook extension. Voiced implosives occur in 11.31% of the UPSID languages: of the 51 languages with voiced implosives, 38 are spoken on the African continent. Curiously the IPA does not include separate symbols for voiceless implosives. Nevertheless, they do occur contrastively with voiced implosives and voiced pulmonic plosives in languages such as Seereer-Siin in Senegal (McLaughlin 2005), and Lendu and Ngiti in the Democratic Republic of Congo (Goyvaerts 1988; Demolin et al. 2002). Although in the spirit of the IPA principles, their phonemic contrastiveness would merit separate symbols such as [ɓ̥], [ɗ̥], [ɠ̥] and [ɣ̥], currently these sounds are represented adding a voiceless diacritic to the corresponding symbols for voiced implosives: [ɓ], [ɗ], [ɠ] and [ɣ].

Finally, clicks are produced on a lingual airstream mechanism: the tongue makes two simultaneous points of contact with the roof of the mouth. The back of the tongue makes contact with the velum, while the blade of the tongue touches the alveolar ridge, the postalveolar region or the hard palate. Between these two occlusions, a small amount of air is trapped. Subsequently, the back of the tongue is drawn backward while maintaining contact with the velum: this rarefies the air between the two closures. When the anterior closure is released, the air implodes into the oral cavity to compensate the pressure difference. The non-linguistic use of clicks is widespread, but linguistically these sounds are confined to Khoisan and Bantu languages of the Southern African continent. The IPA symbols for clicks are underspecified in that they represent the different places of articulation only. This means that further specifications are necessary to indicate whether a click is voiceless, voiced, nasal etc. This can be achieved by adding a superscript before the click symbol as in [k̰] (voiceless), [k̠] (voiced) or [k̠̰] (nasal).

### 3.2 VOWEL SYMBOLS IN THE IPA

Vowel sounds can be defined as sounds with a stricture of open approximation which constitute the nuclei of syllables. The IPA distinguishes 28 vowel symbols presented on a trapezoid vowel chart which can be regarded as a schematic representation of the vocal tract. The IPA vowel chart is illustrated in Figure 4:



**Figure 4:** Illustration of the IPA vowel symbols plotted on a vowel chart

All these vowels are oral in that the velum is raised during their production so

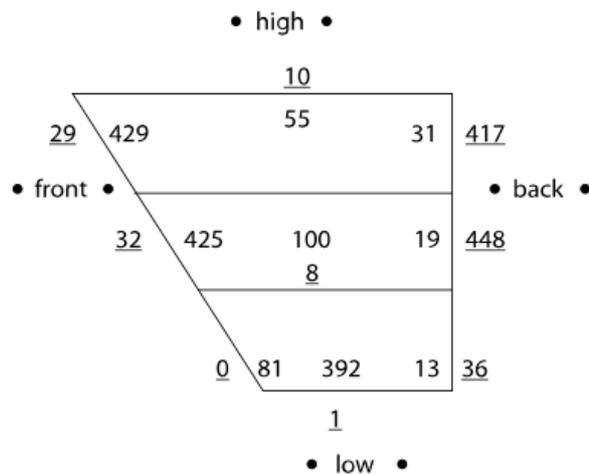
that the airstream escapes via the oral cavity only. The vertical dimension of the chart represents tongue height: vowels with a high tongue position (close vowels) are positioned at the top of the chart, while vowels with a low tongue position (open vowels) are placed at the bottom. In effect, this dimension is somewhat equivalent to the manner of articulation in consonants.

The horizontal dimension of the vowel chart represents the location of the highest point of the tongue in the vocal tract: just as in the consonant chart, the vowel chart also has a left-to-right orientation in terms of place of articulation. Vowels articulated at the front of the oral cavity are located left, while the back vowels appear to the right with the central vowels featuring in between. In Figure 4 it can be seen that most of the vowel symbols are plotted in pairs: the left symbol of each pair represents the unrounded vowel, while the right symbol refers to the rounded vowel (i.e. produced with rounded lips) with the same degree of opening and backness.

The precise perceptual qualities of these vowels may be rather difficult to deal with by non-phoneticians. As a result, textbooks on phonetic notation often present descriptions of their perceptual qualities by referring to vowels in specific languages. This gives descriptions such as [i] 'as in Fr. si; Ger. wie. With value more remote from cardinal in Eng. see' or [ɛ] 'as in Northern Eng. Pronunciation of pen, get; Fr. mettre (short), maître (long); Ger. Bett; Ital. pesca (peach), era (...)' (Principles, 1948: 8).

It should be pointed out though that such descriptions are only to be considered as rough approximations. More correctly, the qualities of at least 16 of these vowels have to be appropriately regarded as Cardinal Vowels: their qualities are not based on the vowels of any existing language. They were created by Daniel Jones as a set of fixed peripheral reference points with respect to which the vowels in all languages of the world can be described. In creating these vowel qualities, Jones used well-defined articulatory and perceptual criteria, the detailed discussion of which is beyond the scope of this article. The interested reader is however referred to e.g. Abercrombie (1967) for an insightful description of these design principles. Suffice it to say that their perceptual qualities can only be accurately illustrated and learnt by going back to the original recordings of these vowels as they were spoken by Daniel Jones himself. These recordings are now easily accessible on the internet.

The number of qualitatively different oral vowels in languages of the world varies considerably and ranges between 3 (in Arabic) and 24 (in !Xu, Khoisan). The vast majority of languages have vowel systems with between 5 and 7 vowels. The frequency of occurrence of vowels representing the various articulatory dimensions discussed above is given in Figure 5.



**Figure 5:** Frequencies of the different vowels in languages of the world out of a total of 2,549 UPSID vowels. Underscored figures refer to rounded vowels (Adapted from Crystal 1997: 169)

From Figure 5, it can be seen that the various articulatory dimensions in vowels are not distributed equally: front vowels tend to be unrounded, while the vast majority of back vowels are rounded. Rounding in the front vowels and unrounding of back vowels does occur, but is relatively uncommon in languages of the world. Furthermore, it can be seen that low vowels tend to be central.

### 3.3 EXTENDING THE ROMAN-BASED SYMBOL STOCK

From the discussion so far it seems clear that at least from an English perspective some of the symbols in Figure 1 are quite familiar from standard orthography: on the upper symbol row [p, b, t, d, c, k, g, q] are Roman letters which are familiar to users of Roman-based orthographies.

This stock of roman-based phonetic symbols has been extended by the IPA in various ways: besides significant import from other orthographies and symbol sets, several typographical changes and modifications to existing symbols have been used. A first extension has been achieved by importing symbols from other alphabets such as an older extended Roman alphabet for English. Examples of this are the vowel symbols [æ] and [œ], which are still used today in Scandinavian orthographies and occasionally also appear in present-day English spelling (MacMahon 1996). Other symbols have been taken from the Greek alphabet such as a.o. [φ] (Phi), [β] (Beta), [θ] (Theta), and [χ] (Chi) to represent several of the fricative sounds.

A second extension has been the importation of symbols from other non-orthographic symbol sets. A good example of this is the symbol for the first sound in the English word 'shield' which is represented as [ʃ]: this symbol is better known as the integral symbol in mathematics. In phonetics it is referred to as an 'esh'.

A third way of obtaining additional phonetic symbols is the implementation of typographical changes to existing symbols. A small number of symbols (i.e. 6) has been created by small-capitalizing lower case Latin letters: [G], [N] and [R]

are used to represent respectively the voiced stop, nasal and trill with a uvular place of articulation, while [B] represents the bilabial trill. Furthermore, [L] refers to the voiced velar lateral approximant and [H] to the voiceless epiglottal fricative.

Some symbols have been arrived at by turning an existing symbol upside down: this is illustrated by the symbol for the uvular fricative [Ɂ] and the voiceless labial-velar fricative [ɸ].

Other symbols have undergone more complicated changes such as mirror-imaging an existing symbol and then turning it upside down, such as [ɹ], [ə] and [ɐ].

A final principle that has been a quite productive source of new symbols is typographically modifying existing roman symbols. All the symbols referring to a retroflex 'place of articulation' have been obtained by adding a downward right-hook extension to the corresponding symbol for an alveolar place of articulation: thus the set [ɖ], [ɳ], [ɽ], [ʐ], [ɻ] and [ɺ] is derived from [d], [n], [r], [z], [l] and [l].

Something very similar has been achieved for the voiced implosive symbols which have been derived from existing symbols by adding an upward right-hook extension, thus creating the series [ɓ], [ɗ], [ɟ], [ɠ] and [ɣ].

It should be mentioned that such typographical modifications have imported some degree of iconicity in the IPA alphabet in that a recurring aspect of a set of symbols represents a specific meaning extending over a whole set of symbols. But whatever the modification or adaptation that has occurred, the IPA has always tried to harmonise the appearance of the phonetic symbols to blend well with the Roman-based symbol set, which are all set in Roman type, never in italic type.

### **3.4 DIACRITICS**

Diacritics are extra markings which can be associated with main sound symbols to capture finer shades of pronunciation. The IPA alphabet contains 31 such diacritics and their usage is illustrated in Figure 6:

DIACRITICS Some diacritics may be placed above a symbol with a descender, e.g.  $\underset{\circ}{\eta}$

◦ Voiceless	$\underset{\circ}{n}$ $\underset{\circ}{d}$	.. Breathy voiced	$\underset{..}{b}$ $\underset{..}{a}$	┘ Dental	$\underset{\text{┘}}{t}$ $\underset{\text{┘}}{d}$
◌ Voiced	$\underset{\text{◌}}{s}$ $\underset{\text{◌}}{t}$	~ Creaky voiced	$\underset{\sim}{b}$ $\underset{\sim}{a}$	┘ Apical	$\underset{\text{┘}}{t}$ $\underset{\text{┘}}{d}$
h Aspirated	$t^h$ $d^h$	~ Linguolabial	$\underset{\sim}{t}$ $\underset{\sim}{d}$	┘ Laminal	$\underset{\text{┘}}{t}$ $\underset{\text{┘}}{d}$
◌ More rounded	$\underset{\text{◌}}{\text{ɔ}}$	ᵂ Labialized	$t^w$ $d^w$	◌ Nasalized	$\underset{\text{◌}}{e}$
◌ Less rounded	$\underset{\text{◌}}{\text{ɔ}}$	j Palatalized	$t^j$ $d^j$	ⁿ Nasal release	$d^n$
◌ Advanced	$\underset{\text{◌}}{u}$	ʲ Velarized	$t^y$ $d^y$	ˀ Lateral release	$d^l$
◌ Retracted	$\underset{\text{◌}}{e}$	ˠ Pharyngealized	$t^ɣ$ $d^ɣ$	ˀ No audible release	$d^ɹ$
◌ Centralized	$\underset{\text{◌}}{e}$	~ Velarized or pharyngealized	$\underset{\sim}{t}$		
˘ Mid-centralized	$\underset{\text{˘}}{e}$	˥ Raised	$\underset{\text{˥}}{e}$ ( $\underset{\text{˥}}{ɹ}$ = voiced alveolar fricative )		
◌ Syllabic	$\underset{\text{◌}}{n}$	˥ Lowered	$\underset{\text{˥}}{e}$ ( $\underset{\text{˥}}{\beta}$ = voiced bilabial approximant )		
◌ Non-syllabic	$\underset{\text{◌}}{e}$	◌ Advanced Tongue Root	$\underset{\text{◌}}{e}$		
◌ Rhoticity	$\underset{\text{◌}}{\text{ɹ}}$ $\underset{\text{◌}}{\text{ɻ}}$	◌ Retracted Tongue Root	$\underset{\text{◌}}{e}$		

**Figure 6:** IPA chart illustrating the use of diacritics

It is possible to distinguish several subsets of diacritics. The most extensive set of diacritics functions to describe modifications of the main sound segment in either place or manner of articulation. Examples of place changes are advancement [a<sup>◌</sup>] or retraction [a<sup>◌</sup>] of place of articulation, while manner changes are exemplified by raising [e̥] or lowering [e̞]: in [e̥] the vowel is pronounced with a higher tongue position than for Cardinal Vowel [e], while in [e̞] the tongue is positioned lower.

A second set of diacritics enables a more detailed specification of phonation characteristics. These can be used to specify that a normally voiced speech segment is devoiced such as in [n̥] or whether a normally voiceless sound appears voiced as in [t̚].

The third set relates to the release of consonants to capture processes like aspiration of plosives [p<sup>h</sup>], nasal release [d<sup>◌</sup>], lateral release [d<sup>l</sup>] or the absence of any audible release [d<sup>ˀ</sup>].

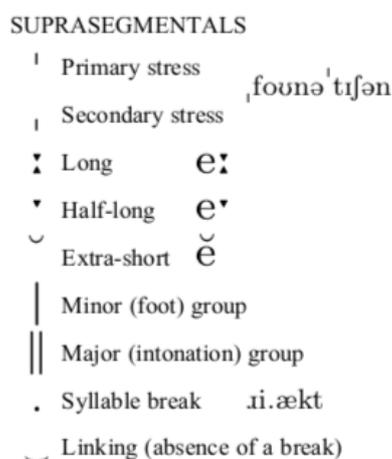
The fourth set allows the transcription of phenomena which normally arise from co-articulation processes such as lip rounding, secondary articulations (labialization [b<sup>◌</sup>], palatalization [ɟ<sup>◌</sup>], velarization [ɫ<sup>◌</sup>] and pharyngealization [ɣ<sup>◌</sup>]), nasalization [ẽ] and rhotacization [ɻ].

Finally, two further diacritics specify the syllabicity of speech segments: they can be used to indicate whether a normally non-syllabic speech segment (consonant) occurs syllabically [n̩] or that a normally syllabic sound (vowel) is used non-syllabically [e̯].

Diacritics are generally, but not always, placed below the symbols for the main sound segments, but may appear as superscripts positioned to the right of or above a symbol. However, it is perfectly acceptable to place a diacritic above the sound symbol if this contributes to improving the readability of the transcriptions. This applies particularly to sound symbols which significantly descend below the writing line such as for instance the palatal [ɲ] and velar nasal [ŋ]. In order to avoid clutter, the voicelessness of such segments can be indicated by placing the diacritic above the main sound symbol as in [ɲ̥] and [ŋ̥].

It has to be kept in mind, however, that potential placement conflicts have to be avoided, especially when transcribing nasalisation ([ẽ̃]) and creaky voice ([ḛ̃]), and breathy voice ([ẽ̤]) and centralisation ([ẽ̠]). In these instances the position of the diacritic above or below the main phonetic symbol does matter.

In addition to the diacritics mentioned above, there are a further 9 diacritics and symbols relating to the suprasegmental aspects of speech. These are illustrated in Figure 7:



**Figure 7:** List of the IPA suprasegmental diacritics

Some of these diacritics refer to word accent placement (primary and secondary stress) and the delineation of intonation groups, while others specify segment length.

Finally, the IPA has 14 diacritics describing aspects of lexical tone in tone languages and pitch movements associated with word accents. These are illustrated in Figure 8:

TONES AND WORD ACCENTS	
LEVEL	CONTOUR
ě or ǀ Extra high	ě or / Rising
é ǁ High	ê \ Falling
ē ǂ Mid	ě ↗ High rising
è ǃ Low	ě ↘ Low rising
ë Ǆ Extra low	ě ↙ Rising-falling
↓ Downstep	↗ Global rise
↑ Upstep	↘ Global fall

**Figure 8:** List of the IPA tone and word-accent diacritics

### 3.5 IPA IDENTIFICATION NUMBERS

It is useful to point out that each IPA symbol and diacritic has been allocated a unique three digit identification number. This permits unambiguous reference to the symbols in e.g. email and in communication with publishers of manuscripts containing phonetic fonts. The codes for consonants begin with 1 (e.g. [p] = 102) and the numbers referring to vowels begin with 3 (e.g. [i] = 301). Segmental diacritics begin with 4 (e.g. palatalisation = 421), while the suprasegmental diacritics begin with 5 (e.g. long duration = 503). In the numerical notation, the diacritics are associated with the main sound segments by means of the addition sign: e.g. [p] = 102 + 421.

### 4. EXTENSIONS OF THE IPA

Besides the 'traditional' IPA symbols, this phonetic notation system has been given various extensions. The first important extension of the IPA was introduced in 1990 by the International Association of Clinical Linguists and Phoneticians in Clinical Phonetics and Linguistics (Duckworth et al. 1990). In this extension a number of new symbols have been introduced to transcribe typical features in disordered speech. An example of this is the velopharyngeal fricative [f̞] which is common in cleft palate speech or the bidental percussive [ʔ]. Moreover, a significant set of diacritics has been added to describe substantial detail relating to various aspects of airstream mechanisms, phonation, articulation and timing. For example, the devoicing process of typically voiced sounds can be specified with great detail: partial devoicing [ç̚], initial partial devoicing [ç̚], final partial devoicing [ç̚]. A striking feature of Extended IPA is the use of bracket notations such as parentheses for silent articulations (speech without sound), double parentheses for unintelligible stretches of speech and curly brackets for detailed prosodic notation.

The second extension is known as SAMPA which is short for Speech Assessment Methods Phonetic Alphabet (Wells 1997). This is a machine-readable phonetic alphabet developed in the late 1980s by an international group of phoneticians. The aim of this project was to provide a standard machine-readable encoding of phonetic symbols. The underlying principle was to take over as many symbols from the IPA (such as ordinary p, b, t, d, k, g), but new symbols had to be

introduced for IPA symbols which are not machine readable such as IPA [ə] and IPA [ɑ]: these are transcribed in SAMPA as '@' and 'A' respectively.

Whereas SAMPA was originally restricted to provide machine readable symbols for the sounds in a restricted number of European languages, X-SAMPA was developed by John Wells as an extension of SAMPA which covers the whole range of IPA symbols. In this system the voiced bilabial implosive [ɓ] for example is represented by 'b\_<' and the voiced glottal fricative [ɦ] by 'h\'. Although unicode solves some of the problems that SAMPA and X-SAMPA aimed to address, both are still frequently used in computational phonetics and speech technology applications.

Finally, it may also be interesting to note that the IPA alphabet is available in a braille version, the original of which was developed by Merrick and Potthoff (1932). Conceptually, in this notation system the aural medium is converted into a tactile medium rather than a visual one. The Braille IPA chart does not have the familiar arrangement of the IPA symbols in terms of manner and place of articulation as intersections of columns and rows: since the Braille symbols are significantly bigger than the print versions, the conventional arrangement of phonetic symbols would yield a page which is too big for practical purposes (Wells-Jensen 2005).

## **5. LEARNING TO USE THE IPA**

Learning to recognise the sounds of the world's languages and making accurate phonetic transcriptions are essential skills for students in many language-related disciplines. Nevertheless, acquiring these skills can be a challenging task especially in view of the increasing financial restrictions on academic institutions: traditional ear training classes are very labour intensive and consequently expensive to run. This is solved in part by Eartrainer (Verhoeven 2003; Verhoeven & Davey 2007), which is a user-friendly multimedia package developed at Antwerp University and City, University of London. This package provides a large stock of ear training exercises for students to use outside the classroom. On the basis of video recordings of an experienced phonetician's delivery of a wide variety of sounds in the world's languages, students engage in different types of ear training exercises with various degrees of difficulty. In the easier types of exercises, students have to provide the correct IPA symbols for descriptive phonetic labels or for individual sounds they hear. In the more demanding exercises, learners have to complete partial transcriptions of nonsense words or they have to transcribe whole words. The interface of the latter exercise type is illustrated in Figure 9:



**Figure 9:** Interface of Eartrainer illustrating the transcription module

In each case, Eartrainer provides instant feedback on students' performance and keeps track of their progress. Furthermore, the transcriptions are analysed in detail so that directed feedback can be given on the specific phonetic dimensions which individual users seem to struggle with.

## 6. USING THE IPA: PHONETIC TRANSCRIPTION

The task of phonetically transcribing a speech sample is approached somewhat differently by different branches of Linguistics and Phonetics. In common to all is the desire to commit to paper rather more detail about the speech material than can be conveyed by orthographic transcription. Phonetic notation is used in different ways depending upon its intended purpose. Thus, a so-called 'broad transcription' aims to capture the phonemic contrasts that are contained in speech. For example, the word 'water' in standard British English may be transcribed as /'wɔtə/. The representation of the basic sound contrasts in the word is clear when compared with the transcription of /'wɔmə/ ('warmer') or /'dɔtə/ ('daughter'), in which the replacement of a single contrastive sound (or 'phoneme') would result in a change in meaning. It should be noted that such broad or 'phonemic' transcription is usually presented between obliques. This may be contrasted with 'narrow phonetic' transcription in square brackets that aims to include details above and beyond that which is required to convey the meanings of the words.

In the word 'water' e.g., it may be relevant perhaps to indicate that the medial consonant is realised as a glottal stop ['wɔʔə] (as might be found in a London speaker), or that the initial sound is produced as a voiceless labial velar fricative

i.e. [ˈmɒtə], or perhaps that the sample derives from a rhotic speaker who uses an r-coloured final schwa and a medial tap [ɾ] rather than a plosive [ˈwɔrə]. In such ‘narrow’ phonetic transcription, the fine detail of pronunciation is included whereas in the phonemic transcription it is implicitly acknowledged that the use of [t] or [ʔ] does not change the intended meaning of the word. In short, therefore, the level of transcription is selected depending upon the task in hand – and in either event the IPA will provide the speech scientist with the necessary transcription tools.

It may be mentioned in passing that producing a transcription may be a painstakingly detailed job. One difficulty is that the sound medium carrying the speech is intrinsically transitory. To overcome this ephemeral characteristic of speech, the phonetic field worker will frequently make a sound recording, nowadays probably using solid state technology. When the speech is replayed, the sound of course remains ephemeral but at least it may be repeatedly reproduced, refreshing the impression in the listener’s auditory cortex.

Whether ‘broad’ or ‘narrow’, the transcription produced in this way is subjective, based as it is upon the interpretation, albeit skilled one, of the perceiver’s ear. In an attempt to make the process more objective, phoneticians (qua speech scientists) will transfer the speech recording to a computer, and examine sound spectrograms of the sound in order to test specific auditory impressions. To return to our example one last time, some corroboration of a rhotic pronunciation of the word ‘water’ might be empirical evidence of a somewhat lowered third resonance (or formant) in the final syllable. By means of such auditory and acoustic phonetic methods, the rich detail of speech may be written down and thoroughly underpinned by scientific methods.

### Notes

1. It should be noted that vertically arranged alphabetic systems have been developed as well. A good example of this is that of Thomas Wright Hill (1860). Such vertical arrangement seems even more impractical from a typographical point of view.
2. The frequency calculations of sound categories reported in this paper are based on the sound inventories of 451 languages, 317 of which were originally published in Maddieson (1984). User-friendly access to this database is now available on [http://web.phonetik.uni-frankfurt.de/upsid\\_find.html](http://web.phonetik.uni-frankfurt.de/upsid_find.html).

### References

**Abercrombie, D.**

**1967** *Elements of general phonetics*. Edinburgh University Press.

**Bell, A.M.**

**1867** *Visible speech: The science of universal alphabets; or self-interpreting physiological letters, for the writing of all languages in one alphabet*. Simpkin, Marshall & Co; N. Trübner & Co.

**Catford, I.**

**1977** *Fundamental problems in phonetics*. Edinburgh University Press.

**Crystal, D.**

**1997** *The Cambridge encyclopedia of language*. Cambridge University Press.

**Demolin, D., H. Ngonga-Ke-Mbembe & A. Soquet**

**2002** Phonetic characteristics of an unexploded palatal implosive in Hendo. *Journal of the International Phonetic Association* 32: 1–15.

**Duckworth, M., G. Allen, W. Hardcastle & M.J. Ball**

**1990** Extensions to the International Phonetic Alphabet for the transcription of atypical speech. *Clinical Linguistics and Phonetics* 4: 273–80.

**Goyvaerts, D.**

**1988** Glottalized consonants: a new dimension. *Belgian Journal of Linguistics* 3: 97–102.

**Jespersen, O.**

**1889** *The articulations of speech sounds represented by means of analphabetic symbols*. N.G. Elwert.

**Laver, J.**

**1994** *Principles of phonetics*. Cambridge University Press.

**Maddieson, I.**

**1984** *Patterns of sounds*. Cambridge University Press.

**Macmahon, M.K.**

**1996** Phonetic notation. In P.T. Daniels & W. Bright (eds.) *The world's writing systems*: 821–46. Oxford University Press.

**Merrick, P.W. & W. Potthoff**

**1932** *A Braille notation of the International Phonetic Alphabet*. The Royal National Institute for the Blind.

**Mclaughlin, F.**

**2005** Voiceless implosives in Seereer-Siin. *Journal of the International Phonetic Association* 35: 201–14.

**Pike, K.L.**

**1943** *Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds*. University of Michigan Press.

**Principles**

**1948** *The principles of the International Phonetic Association being a description of the International Phonetic Alphabet and the manner of using it, illustrated by texts in 51 languages*. University College London.

**Verhoeven, J.**

**2003** Ear training on line as an educational tool for the teaching of IPA transcription. In L. Polka, M. Sundara, M.J. Sole, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*: 3085–87.

**Verhoeven, J. & R. Davey**

**2007** A multimedia approach to eartraining and IPA transcription. In J. Maidment (ed.) *Proceedings of the Phonetics Teaching and Learning Conference*: 1–4.

**Wells, J.C.**

**1997** SAMPA computer readable phonetic alphabet. In D. Gibbon R. Moore & R. Winski (eds) *Handbook of Standards and Resources for Spoken Language Systems*: Part IV, section B. Mouton de Gruyter.

**Wells-Jensen, S.**

**2005** The Braille International Phonetic Alphabet and other options: the blind student in the phonetics classroom. *Journal of the International Phonetic*

*Association*: 35: 221–30.

**Wright Hill, T.**

**1860** *On the articulation of speech.*