12 Analysing phonetic and phonological variation on the segmental level

ULRIKE GUT

1 Introduction

Differences in pronunciation between different regional, national and social accents of languages are immediately striking even to non-linguists. While a small number of these differences are purely incidental, for example the fact that the second vowel in *tomato* is pronounced /a/ in Southern Standard British English but /eɪ/ in Standard American English, most differences between accents of any language are systematic. This systematic variation can be analysed with reference to a large number of phonological domains and structures and different phonological and phonetic processes. On the segmental level (see Gut, Chapter 13, this volume, for variation on the suprasegmental level), these include:

the phoneme inventory (vowels and consonants),
the phonetic realization of vowels and consonants, and
the phonotactic distribution of phonemes.

In recent years, many research methods for exploring phonological and phonetic variation on the segmental level have become established. Moreover, a number of new methods have become available, some of which have not yet been applied to this area. In the following sections, both the widely used and some new methods will be presented and discussed. Section 2 is concerned with the analysis of vowels in terms of phoneme inventory (Section 2.1), their phonetic realization (Section 2.2) and their articulation (Section 2.3). Section 3 evaluates methods of studying consonants: the consonant inventory (Section 3.1), their phonetic realization (Section 3.2) and their articulation (Section 3.3).

2 Analysis of vowels: inventory and distribution, realization, articulation

Vowels are speech sounds that are produced with no audible obstruction of the airstream (cf. e.g. Giegerich 1992: 12). The vowels of a language are very susceptible to changes, as both the history of many languages

 https://doi.org/10.1017/CBO9780511792519.016

and the synchronic variation within many languages demonstrate. The vowel inventory of a language consists of those vocalic sounds that have phonemic status, i.e. that are contrastive or meaning-distinguishing. The linguistic concept of a phoneme is based on the idea that speakers have mental representations of the contrastive speech sounds of their language – they are assumed to form part of their linguistic knowledge. There is unfortunately no direct way of studying mental representations of speakers. Speakers' linguistic knowledge thus has to be deduced from analyses of vowels in actual speech production. Taking English vowels as an example, the method of auditory analysis is described in Section 2.1. Acoustic measurements of variation in the phonetic realization of vowels are discussed in Section 2.2. Section 2.3 presents recent methods of measuring vowel articulations.

2.1 Analysis of vowel inventories and distribution: auditory analysis

The current standard method for the description of the vowel inventory of a variety of English involves the standard lexical sets developed by Wells (1982: 122ff.). Each lexical set consists of a number of English words with a shared vowel. It is represented by an unambiguous monosyllabic keyword printed in small capitals. Thus, the FLEECE standard lexical set comprises words such as *creep*, *sleeve*, *key* and *people*, which all share the vowel /i/. This vowel is therefore often referred to as the FLEECE vowel. The keywords were chosen to end in either a voiceless alveolar or a dental consonant in order to minimize the effects of the adjacent consonants on vowel articulation (Wells 1982: 123). An analysis of British and American English yielded 24 lexical sets for vowels in stressed syllables. These are complemented by three sets for vowels in unstressed syllables. Table 12.1 lists these 27 lexical sets with their respective keyword and some example words.

In a typical set-up of data collection for an auditory analysis of vowel inventories, speakers are recorded reading a list of all 27 keywords printed on a piece of paper or presented to them on a computer screen (see Section 2.2 for a description of how to make good recordings). Since the reading of word lists represents a fairly rare speaking style, often a reading passage is given to the speakers as well. The standard text used for English is 'The North Wind and the Sun' (see Text 1); it has been translated into many other languages for the study of their vowel inventories (see the International Phonetic Association's *Handbook of the International Phonetic Association* 1999).

Text 1. Standard text for the analysis of the vowel inventory of a variety of English.

> The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he

could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

Table 12.1. *The standard lexical sets for the analysis of vowel inventories with keyword and example words*

Lexical set	Example words
KIT	*ship, rip, dim*
DRESS	*step, ebb, stem*
TRAP	*bad, cab, ham*
LOT	*stop, rob, swan*
STRUT	*cub, rub, hum*
FOOT	*full, look, could*
BATH	*staff, clasp, dance*
CLOTH	*cough, long, gone*
NURSE	*hurt, term, work*
FLEECE	*creep, sleeve, key*
FACE	*weight, rein, steak*
PALM	*calm, bra, father*
THOUGHT	*taut, hawk, broad*
GOAT	*soap, soul, home*
GOOSE	*who, group, few*
PRICE	*ripe, tribe, aisle*
CHOICE	*boy, void, coin*
MOUTH	*pouch, noun, how*
NEAR	*beer, clear, fierce*
SQUARE	*care, air, tear*
START	*far, sharp, farm*
NORTH	*war, storm, for*
FORCE	*floor, coarse, ore*
CURE	*poor, tour, fury*
happY	*copy, city, penny*
lettER	*paper, offer, anchor*
commA	*quota, panda, saga*

The standard procedure is then to identify the vowels in the recordings by transcribing them with the symbols of the International Phonetic Alphabet (IPA).[1] The IPA was first issued by the International Phonetic Association[2] in

[1] When transcribing on a computer it might be advisable to use SAMPA, the computer-readable version of the IPA (Wells *et al*. 1992). IPA Unicode fonts are available at various places, e.g. at www.wazu.jp/gallery/Fonts_IPA.html and www.unc.edu/~jlsmith/ipa-fonts.html.

[2] www.arts.gla.ac.uk/IPA/ipa.html.

1886 and has since then undergone several revisions (the last one was completed in 2005). It contains transcription symbols for all distinctive speech sounds that occur in any language of the world. In addition, it offers transcription symbols for fine phonetic details and prosodic features, the so-called diacritics (see Section 2.2). In the IPA vowel chart (Figure 12.1), the symbols are arranged roughly according to the articulation of vowels. On the vertical axis, the vertical position of the tongue and lower jaw is represented, ranging from close (or high) to open (or low). The horizontal axis refers to the specific part of the tongue that is active during articulation, comprising the front, central and back part. When a position in the IPA vowel chart is filled by two symbols, the vowel on the right is produced with rounded lips and the one on the left with spread lips.

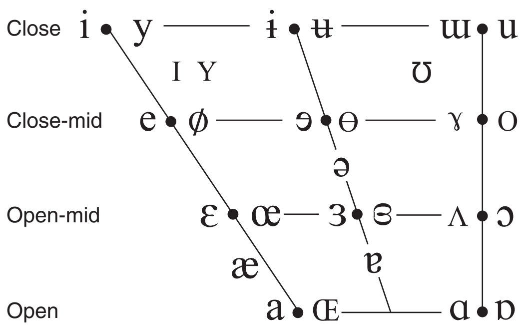

Figure 12.1. *The IPA transcription symbols for the cardinal vowels. Reprinted with permission from the International Phonetic Association. Copyright 2005 by International Phonetic Association*

Note that the placements of the vowel symbols in the vowel chart in Figure 12.1 are idealized. The positions of tongue and jaw that are indicated by the arrangement of the symbols are extreme positions which are rarely reached in real articulation. These idealized vowels are referred to as cardinal vowels. Their primary function is to serve as reference values for phoneticians and phonologists who wish to describe the vowel inventory of a language. Unfortunately, the difference between the front and the central part of the tongue or a mid-closed and a mid-open mouth, for example, is neither absolute nor clear-cut. In other words, vowels have no exact articulatory boundaries, and differences in vowel articulation are gradual. Consequently, there remains a subjective factor in any auditory analysis of vowels: agreement between different transcribers transcribing the same recording and the reliability of transcriptions vary between low and satisfactory, depending on the amount of training transcribers had and on the transcribing conditions (e.g. Wester *et al*. 2001; Gut and Bayerl 2004). Moreover, even transcribers who had extensive training are influenced by the phonetic context in their transcription of vowels (Nairn and Hurford 1995).

Based on the auditory method, differences in the distribution between vowels in different accents of English have been analysed. For example, Schneider (2007: 76) reports that in the Caribbean accents of English the NEAR and

SQUARE vowels are merged, i.e. these two lexical sets are pronounced with an identical vowel. In Standard Southern British English (SSBE), LOT is pronounced with a /ɒ/, whereas in General American (GA) this lexical set has the vowel /ɑ/. Variation of the pronunciation of STRUT is especially high across accents of English: it is pronounced /ʌ/ in SSBE, /ʊ/ in many Northern British accents, /ɔ/ in Nigerian English and /ɑ/ in Singapore English (see e.g. Mesthrie and Bhatt 2008: 121).

In most accents of English such as SSBE, GA, Australian English and Canadian English, the distribution of vowels furthermore differs distinctly depending on their degree of stress. While in stressed syllables all full vowels, i.e. vowels produced at the periphery of the vowel chart presented in Figure 12.1, can occur, the only vowels that appear in unstressed syllables in these varieties are the central 'reduced' vowels /ə/ and /ɪ/. In some accents of English, conversely, this distribution of vowels is different: Setter (2003), for example, reports that many syllables that usually contain a reduced vowel in British English are produced with a full vowel in Hong Kong English.

In summary, the auditory analysis of the vowel inventory of a language is a well established standard method (see the *Handbook of the International Phonetic Association* 1999 for details). Compared to other methods, it is relatively fast and requires only some basic recording equipment. Auditory analyses, however, require training, in which the correct associations of sounds and IPA symbols have to be learned. Yet, it appears that even very experienced transcribers never reach absolute agreement and reliability. For an auditory analysis of a particular vowel inventory, it is necessary to record a sufficiently large number of informants in order to avoid documenting individual speakers' idiosyncrasies. This might be especially important for the analysis of non-native varieties of a language that typically have a high inter-speaker variability. One drawback of the method of auditory analysis is that the choice of the appropriate IPA vowel symbol can sometimes be difficult, and agreement between linguists can be low. It is impossible to prove whether this method really succeeds in describing speakers' mental representations of contrastive vowels. Some phoneticians in fact doubt whether the phoneme is a useful concept at all and whether speakers really need representations of phonemes for speech production and perception. Descriptions of the vowel inventory of all the languages in the world are published regularly in the *Journal of the International Phonetic Association* (JIPA). Comparative descriptions of the variation of vowel inventories across varieties of English based on the method of auditory analysis can be found in Wells (1982), Schneider *et al*. (2004), Schneider (2007) and Mesthrie and Bhatt (2008).

2.2 Acoustic analysis of vowel realizations

Acoustic analyses are carried out in order to explore fine-grained details of the phonetic properties of vowels. In acoustic terms, vowels consist of periodic sound waves with different frequencies at different intensities. The lowest frequency is the one that listeners perceive as pitch (see Gut, Chapter 13, this volume); the higher frequencies, which are called harmonics or overtones, provide the particular quality of the vowel. Of the sine waves with different frequencies that are produced by the vocal fold vibration in the larynx, some are enhanced and others are dampened by the resonating properties of the vocal tract. It is the shape and position of the articulatory organs and the size of the vocal tract that influence the quality of vowels: the position of the tongue and jaw change the size and shape of the oral cavity, and this in turn influences which of the frequencies produced by vocal fold vibration are enhanced in intensity and which are dampened. Those bands of frequencies that have very high intensity (= energy) are called formants. Thus, formants are the most important acoustic cues for vowels – they distinguish the different vowels of a language. By the same token, they precisely reflect differences in articulation between individual vowels.

With the help of a spectrogram it is possible to see and measure the formants of vowels. Speech analysis software packages such as *Praat*,[3] *SpeechAnalyzer*[4] and *Wavesurfer*[5] create spectrograms from sound recordings, in which the individual frequencies of each speech sound and their corresponding intensities are made visible. Figure 12.2 shows the spectrogram of some vowels that was created by *Praat*, the most widely used and very powerful speech analysis software. The horizontal axis shows time, the vertical axis depicts the different frequencies, and the dark horizontal bands indicate the greater relative intensity of particular frequencies, i.e. the formants, in each vowel. Most of the vowels in Figure 12.2 have three distinguishable formants.

Most speech analysis software packages offer an automatic analysis of formants that allows a measurement of the actual frequencies of the formants.

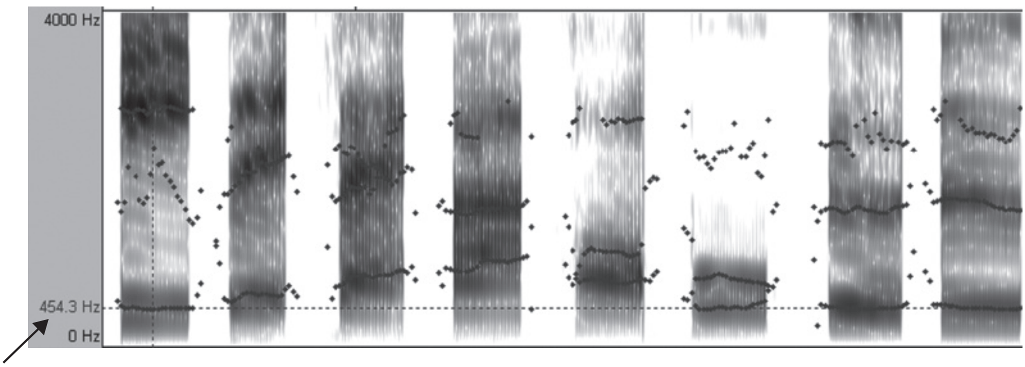

Figure 12.2. *Spectrogram of the vowels [i], [ɪ], [e], [æ], [ɒ], [ɔ], [ʊ] and [u] with indication of the formants*

[3] Downloadable at www.fon.hum.uva.nl/praat.
[4] Downloadable at www.sil.org/computing/sa/index.htm.
[5] Downloadable at www.speech.kth.se/wavesurfer.

The exact value of each formant frequency at each point in time is displayed in the unit Hertz (Hz). The position of each of the formants of the vowels shown in Figure 12.2 is indicated by the dotted lines that are superimposed on the spectrogram. The formant with the lowest frequency is called F1, the second one F2, the third one F3 and so on. For the differentiation of vowels in a particular accent it is usually sufficient to refer to the first two or three formants. For measuring for instance the F1 of the first vowel [i], one only has to click on the dotted formant line and the value will be displayed (in our case 454.3 Hz; see Figure 12.2 on the left). In *Praat*, each formant value is even calculated when pressing the keys F1, F2 etc. Figure 12.2 illustrates that the height of F1 is correlated with tongue height. The lower the tongue is during the production of the vowel, the higher is F1. The relationship between vowel type and the second formant is more complicated: there is some correlation between back vowels and a relatively low F2. Furthermore, F2 is also considerably affected by whether the vowel is produced with rounded or spread lips.

When measuring formant height in individual vowels, it is important to choose an appropriate place for this, since the sounds preceding and following a vowel influence the shape of the formants. It is customary to select the 'stable part' at the mid-point of the vowel for formant measurements. The beginning and end point of a vowel are usually determined with reference to formant movements, which can be seen in the spectrogram. The standard procedure (e.g. Peterson and Lehiste 1960) is to indicate the beginning of a vowel at the beginning of a stable formant structure, especially at the onset of F1, and to indicate the end at the end of a stable formant structure, especially at the end of F2. This method is unfortunately far from straightforward when the vowel in question is very short. When comparing formants in vowels produced by different speakers it has to be borne in mind that, due to physiological differences between speakers, two vowels that sound the same are likely to have different formants. Quantitative comparisons across different speakers are thus usually carried out by applying normalization (see Syrdal and Gopal 1986 for an overview of the different normalization techniques).

The average formant frequencies of vowels can be taken to plot an acoustic map, in which the vowels are positioned according to the frequency of their first two formants. Figure 12.3 shows the formant chart for American English vowels. The values of F1 are displayed in inverse order on the vertical axis, whereas the values of F2 are displayed on the horizontal axis, again in inverse order. Thus, in a slightly unusual arrangement, the zero values of both axes appear in the top right-hand corner. This arrangement shows immediately that the vowels are positioned in such an acoustic map similarly to how they are usually arranged in the IPA vowel chart (see Figure 12.1), where vowels are placed based on their articulatory properties. Thus, the formants of vowels demonstrate a clear relationship to the traditional articulatory description of vowels. Tongue height is inversely correlated with height of F1. F2 varies with tongue retraction (back and front position) and lip rounding.

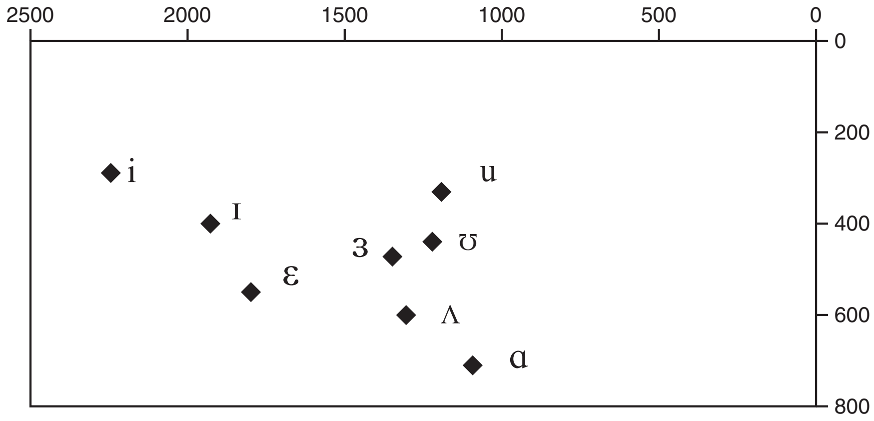

Figure 12.3. *Acoustic map for some American English vowels. F2 values on top, F1 values on the right. Reprinted with permission from Gut (2009: 154)*

Many descriptions of the realization of vowels in different accents of English are based on the acoustic measurement of formants. The average formant values for both Standard British English and Standard American English can, for example, be taken from Kent and Read (2002: 111f. and 122). Rosenfelder (2007) analyses the raising of /aɪ/ and /aʊ/ in Canadian English, and Watt and Tillotson (2001) describe fronting in Bradford English. Recent changes in Received Pronunciation such as the lowering of /æ/ and the fronting and loss of rounding of /u/ are reported by Hawkins and Midgley (2005). By the same token, Harrington, Palethorpe and Watson (2000) document similar changes in the Queen's English over the past 50 years. A description of the Northern Cities Shift based on acoustic measurements can be found in Labov (1994). Based on an acoustic analysis, Low, Grabe and Nolan (2001) show that reduced vowels in Singapore English are produced more peripherally in the F1/F2 formant space than in British English, i.e. they are centralized to a lesser extent.

Other phonetic properties of vowels that can be measured acoustically are duration and intensity of the formants. When the beginning and end points of a vowel are marked (see above) in a speech analysis program, its exact duration will be displayed. The measurement of vowel durations is often employed for a comparison of vowel reduction and speech rhythm across different accents of English (see Gut, Chapter 13, this volume). Furthermore, the relative intensities of specific vowel formants are measured to calculate spectral tilt, which has been shown to play a significant role in the production of stress in English (e.g. Campbell and Beckman 1997; Okobi 2006).

The method of an acoustic analysis of vowels, in contrast to an auditory analysis, yields exact quantitative results. Furthermore, again unlike the auditory analysis, fine phonetic details in the variation of vowel productions can be measured with it (see Kent and Read 2002, chapters 3 to 6, Davenport and

Hannahs 2005, chapter 5, Yavaş 2006, chapter 5, Clark, Yallop and Fletcher 2007, chapter 7, and Gut 2009b, chapter 5 for details). On the downside, this methodology is far more time-consuming than an auditory analysis and requires both technical skills and a minimum understanding of the acoustic properties of vowels. Moreover, in order to carry out reliable measurements of the acoustic properties of speech, it is necessary to obtain high-quality speech recordings. The basic prerequisite for such a recording is a suitable recording environment with minimal background noise, ideally a sound-treated room that excludes all external noise. Both the recording device and the microphone that are used need to be sensitive enough to capture all acoustic cues that are perceptually important. When recording with a digital device, a sampling rate of at least 22 kHz, ideally 48 kHz, should be chosen. Another decision that has to be made regarding quantization refers to the number of separate amplitude levels that are represented on a computer. They are stored in binary digits (bits) so that quantization is expressed in bits. Most researchers encode speech signals with 12 or 16 bits.

2.3 Articulatory measurements of vowel production

The articulation of vowels and especially the actions of the tongue and lips during vowel production can also be examined with articulatory measurements. Electromagnetic mid-sagittal articulography, or EMA, allows the analysis of articulatory movements within the oral cavity with high time resolution (see Hoole and Nguyen 1999 for a description of this method and its use in studies on coarticulation). For an EMA study, small transducers, which are connected to an amplifier with wires, are placed on a speaker's tongue, jaw, lips and nose. The speaker additionally sits inside a helmet-shaped structure in which the transducers create a magnetic field (see Figure 12.4). The measurement of the alternating currents from the transducers allows the calculation of the degree and velocity of tongue and lip movements during articulation. Due to the expenses and the restriction of articulation experienced by participants in these experiments, only few speakers can be analysed and the type of speech that can be elicited is fairly restricted.

Another method of analysing tongue movements is ultrasound imaging (see e.g. Stone 2005). The transducer, which is placed below the chin, emits ultra-high frequency sound waves that are reflected by the tongue and oral cavity. The received reflections (or echoes) are then converted into an electrical signal and sent to a computer, which reconstructs them into a 2D image. Although the speaker's head needs to be stabilized, this method is far less invasive than EMA and allows the analysis of a wide range of speech data.

The nasalization of vowels, which has been claimed to be characteristic of certain Caribbean and Pacific accents of English (Schneider 2007: 76), can also be studied with articulatory methods. A nasometer measures nasal airflow, i.e. the amount of breath leaving through the nose during articulation (see Figure 12.5). For this, a speaker places his or her nose above a metal plate. Small microphones

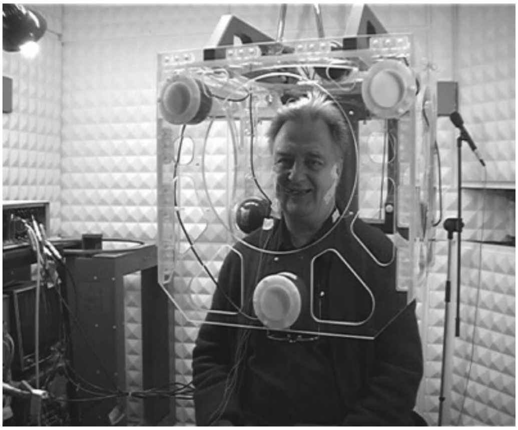

Figure 12.4. *A participant in an EMA experiment. Reprinted with permission from Draxler (2008: 73)*

are then either fitted to the speaker's head or attached to poles on this plate. The upper microphone records air flowing from the nose, whereas the lower one records air flowing from the oral cavity. This enables researchers to measure the relationship between both types of airflow.

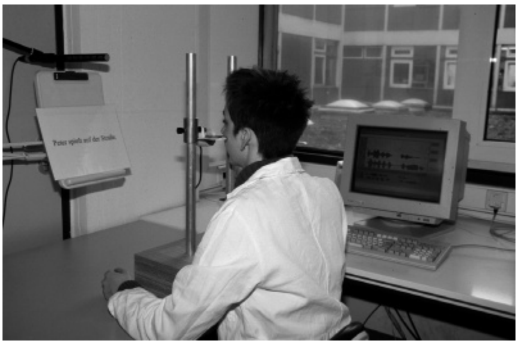

Figure 12.5. *A speaker on a nasometer*

Clearly, methods of studying vowel articulation require special equipment that is usually only found in phonetics laboratories or medical units. The EMA method, in addition, is very expensive and requires intensive training for the correct interpretation of the results. Moreover, extensive training is required for the interpretation of the imaging data. These methods do, however, offer unique possibilities of measuring articulatory processes.

3 Analysis of consonants: inventory and distribution, realization and articulation

Consonants are generally defined as speech sounds that are produced with some audible obstruction of the airstream (see e.g. Giegerich 1992: 8). The consonant inventory of languages seems to be less susceptible to change than the vowel inventory, as both diachronic and synchronic comparisons show. Remarkable variation in the consonant system between accents of languages, however, can be found in terms of their distribution (see Section 3.1). The methods for the investigation of the phonetic realization of consonants are described in Section 3.2. In articulatory terms, consonants can be described according to their place and manner of articulation. Articulatory measurements of these are presented in Section 3.3.

3.1 Consonant inventory and distribution

As in the case of vowels, the study of the consonant inventory of a language aims at describing the mental representation of contrastive consonants that speakers of this language or accent are presumed to have. The established method, as in the study of the vowel inventory, is to record speakers reading either a list of words or a reading passage (often also 'The North Wind and the Sun'). The recordings are then analysed auditorily and transcribed by using the IPA symbols for consonants. Figure 12.6 presents the IPA symbols for all pulmonic consonants, i.e. consonants produced with an egressive pulmonic airstream. The chart can be read in the following way: each row represents a different manner of articulation, ranging from plosive to lateral approximant. Each column refers to a different place of articulation: reading from left to right, from the lips (bilabial) to the larynx (glottal). Each cell of the chart thus represents a particular combination of a manner and a place of articulation. When two symbols appear in one cell, the one on the left is the symbol for the voiceless consonant and the one on the right is the symbol for its voiced counterpart. Cells that are shaded grey signify combinations of place and manner of articulation that are physiologically impossible. Empty cells hold possible combinations of place and manner of articulation, but no such sound has yet been discovered in the languages of the world.

	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	ʙ			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			

Figure 12.6. *The IPA transcription symbols for the pulmonic consonants. Reprinted with permission from the International Phonetic Association. Copyright 2005 by International Phonetic Association*

Phonotactic analyses are concerned with the combinations of consonants that are allowed in syllable onsets and codas in a particular language. For example, in English, in non-rhotic accents such as SSBE, /ɹ/ does not occur in the syllable coda, whereas in rhotic accents such as GA it does. Similarly, while in SSBE the onset clusters /nj/, /tj/ and /lj/ occur, they do not occur in GA. In some varieties of English, coda consonant clusters tend to be avoided on a large scale (see e.g. Gut 2005b for Singapore English). Accents of English can also differ in the distribution of allophones of certain consonants. While in Irish English, /l/ is generally produced as a 'clear' /l/ with an alveolar place of articulation, in SSBE a velarized 'dark' /l/ is produced in the syllable coda and GA, finally, has dark(er) /l/ in all positions.

Like the auditory analysis of vowels, the auditory analysis of consonants is comparatively fast and standardized, although there is no standard text for the recordings. In order to avoid documenting individual speakers' idiosyncrasies, a sufficiently large number of speakers should be recorded. The auditory analysis method and transcription require some training, but perfect agreement between transcribers will never be reached. Wesenick and Kipp (1996), for example, investigated the reliability of the transcription of consonants: The pairwise agreement calculated for the ten transcribers was, on average, 94.8%. The lowest agreement was found for stops (89.9%), and the highest for nasals (97.5%).

3.2 Realization of consonants: auditory and acoustic analysis methods

The phonetic realization of consonants can be investigated both auditorily and acoustically. In an auditory analysis, one of the IPA diacritic symbols (see Figure 12.7) is added to the transcription of a consonant in order to indicate additional articulatory details of its production. For example, the transcription [kæp̚tn] describes that the /p/ in the word *captain* is unreleased, and the transcription [tʰi] shows that the /t/ in *tea* is articulated with aspiration. By the same token, the velarized 'dark' allophone of /l/ is transcribed as [ɫ]. Based on this method, Wells (1982: 74) claims that some accents in Scotland and Northern England do not have aspirated /p,t,k/. By the same token, Mesthrie and Bhatt (2008: 128) report that /n/ is retroflex before /ɖ/ and /ʈ/ in Indian English.

Variation in the phonetic realization of consonants can also be measured acoustically. The aspiration of plosives, for example, is measured by calculating the voice onset time (VOT), the time interval between the release of the consonant closure and the beginning of voicing. Figure 12.8 illustrates the measurement of the VOT of the /k/ in *came*. The release of the consonant is clearly visible in the waveform, as is the beginning of the voicing of the following vowel.

For British and American English voiced plosives, the burst typically occurs between 20 ms before and 20 ms after voicing begins (Kent and Read 2002: 151).

◌̥	Voiceless	n̥ d̥	◌̤	Breathy voiced	b̤ a̤	◌̪	Dental	t̪ d̪
◌̬	Voiced	s̬ t̬	◌̰	Creaky voiced	b̰ a̰	◌̺	Apical	t̺ d̺
ʰ	Aspirated	tʰ dʰ	◌̼	Linguolabial	t̼ d̼	◌̻	Laminal	t̻ d̻
◌̹	More rounded	ɔ̹	ʷ	Labialized	tʷ dʷ	◌̃	Nasalized	ẽ
◌̜	Less rounded	ɔ̜	ʲ	Palatalized	tʲ dʲ	ⁿ	Nasal release	dⁿ
◌̟	Advanced	u̟	ˠ	Velarized	tˠ dˠ	ˡ	Lateral release	dˡ
◌̠	Retracted	e̠	ˤ	Pharyngealized	tˤ dˤ	◌̚	No audible release	d̚
◌̈	Centralized	ë	◌̴	Velarized or pharyngealized	ɫ			
◌̽	Mid-centralized	e̽	◌̝	Raised	e̝	(ɹ̝ = voiced alveolar fricative)		
◌̩	Syllabic	n̩	◌̞	Lowered	e̞	(β̞ = voiced bilabial approximant)		
◌̯	Non-syllabic	e̯	◌̘	Advanced Tongue Root	e̘			
˞	Rhoticity	ɚ a˞	◌̙	Retracted Tongue Root	e̙			

Figure 12.7. *The IPA transcription symbols for phonetic details: the diacritics. Reprinted with permission from the International Phonetic Association. Copyright 2005 by International Phonetic Association*

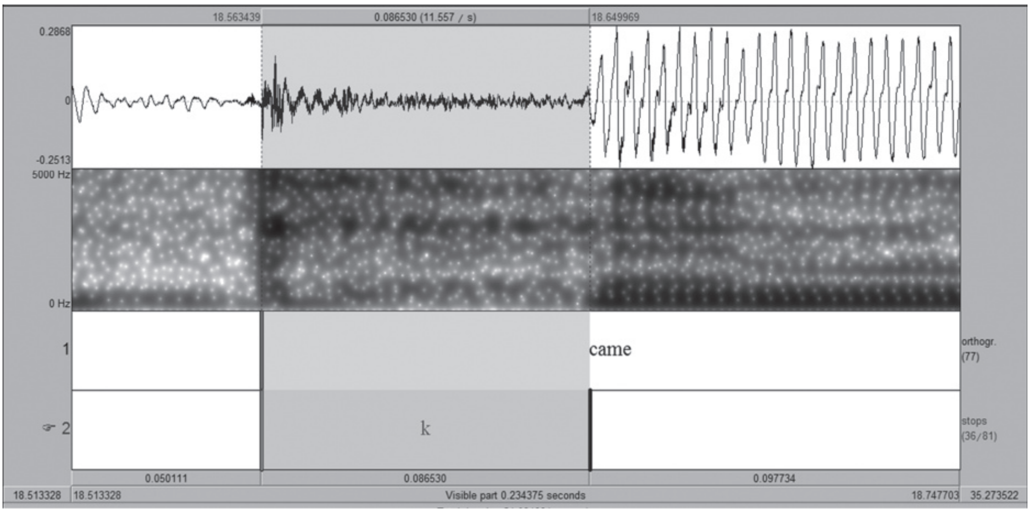

Figure 12.8. *Measurement of the VOT of the /k/ in* came

In other words, when producing [ba], [da] or [ga], speakers usually release the blockage of the airstream for the [b], [d] or [g] between 20 ms before and 20 ms after their vocal folds start vibrating. The corresponding VOT values are thus −20 ms to +20 ms. For voiceless plosives in British and American English, the typical VOT ranges between +40 and +80 ms. Aspirated plosives can have a VOT of up to +120 ms, which means that there can be a 120 ms interval filled with friction between the release of the airstream obstruction and the beginning of voicing for the vowel in words like *pat*, *tack* and *cap* (Kent and Read 2002: 151). VOT values for voiceless plosives in languages differ considerably, and it has been shown that for example Spanish and Japanese learners of English produce very different VOT values in English (e.g. Schmidt and Flege 1996; Riney and Tagaki 1999).

A comparison of auditory and acoustic methods of studying the realization of consonants shows that each method has its strengths and weaknesses. While an auditory analysis requires far less time than an acoustic analysis, it yields less reliable results. The acoustic measurements are objective and quantitative but call for high quality recordings and labour-intensive measurements. While for an auditory analysis training in the use of the IPA symbols and diacritics is necessary, the acoustic analysis presupposes some knowledge of the acoustic properties of consonants (see e.g. Davenport and Hannahs 2005, chapter 5; Yavaş 2006, chapter 5; Clark, Yallop and Fletcher 2007, chapter 7). For both approaches, ideally recordings that comprise many different speaking styles should be analysed because phoneme realizations vary with speaking style and speech rate. It might be a good strategy for both the analysis of vowels and consonants to first carry out an auditory analysis to identify those aspects and areas that promise to benefit most from a subsequent acoustic analysis.

3.3 Measurement of consonant articulation

Some aspects of the articulation of consonants can be measured in a fairly direct way. For example, vocal fold activity can be examined with a laryngograph that measures the degree of voicing during the articulation of speech sounds. For this, two electrodes are placed on a speaker's throat on each side of the thyroid cartilage, and a weak electrical current is passed between them. The strength of the current shows the degree of contact between the vocal folds, which is displayed in a waveform. For directly observing vocal fold activity an endoscope can be used. An endoscope consists of a tube that is fitted with a light source and is connected to a recording device. It is inserted into a speaker's mouth through the nose and held directly above the larynx, where the vocal fold activity is captured with high-speed digital imaging (see Hirose 1988). This method unfortunately is fairly invasive and uncomfortable for the participant. Moreover, it is limited in its power to observe natural articulatory gestures, since no speech sounds can be produced that involve significant tongue movement.

It is furthermore possible to measure the contact of the tongue with the roof of the mouth during consonant articulation. This is done with the help of an electropalatograph, which consists of a thin artificial palate that is fitted with a large number of electrodes. The artificial palate is about 1.5 mm thick and, after having been modelled exactly after the individual shape of the speaker's hard palate, is fitted over it. The electrodes on the artificial palate fire when they are touched by the tongue, so that the position and degree of tongue contact can be measured with the help of an attached computer. Palatographic studies have shown that the patterns of articulatory movements can differ enormously across speakers producing the same sound. Figure 12.9 shows the place and degree of contact of the tongue with the palate for two speakers articulating [t] and [d]. The upper part of the picture shows the region of the alveolar ridge, the bottom part the region of the hard palate. Black areas indicate strong contact, while grey areas were only lightly touched by the tongue, and white areas not at all. The left-hand palatogram shows that the speaker's tongue has strong contact with both the alveolar ridge and the two sides of the palate during the articulation of [t] and [d], whereas the second speaker, whose palatogram can be seen on the right, articulates those sounds predominately with tongue contact at the alveolar ridge.

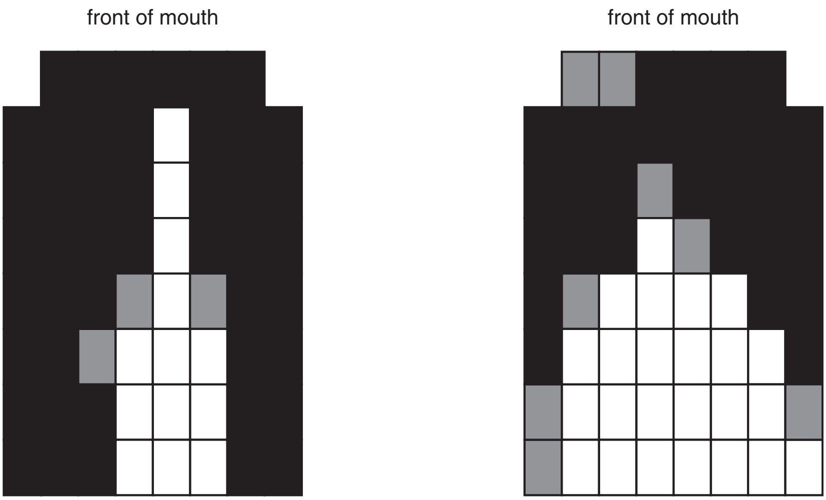

Figure 12.9. *Palatograms of [t] and [d] for two speakers. Reprinted with permission from Gut (2009: 47)*

Based on a palatographic study, Price (1981) measured the duration of tongue-palate closure in /t/ and /d/ produced by American English speakers. By the same token, Fang (2008) demonstrates the exact place of articulation of three sibilants in Mandarin Chinese. These measurements of articulation are the only ones that capture articulatory details of consonants fairly directly. However, for these methods specialist equipment is required that is usually only found in phonetics laboratories or medical research institutions.

Analysing phonetic and phonological variation on the segmental level	
Pros and potentials	Cons and caveats
• auditory analysis and transcription requires little technical support	• auditory analysis and transcription requires intensive training; lack of agreement between transcribers possible
• acoustic analysis captures fine-grained details and is very precise	• technical prerequisites, skills, knowledge of physical processes and time required for acoustic analysis
• measurements of articulatory movements possible	• articulatory obstruction by articulatory measurements; expensive technical equipment and specialized knowledge for interpretation of results required

Further reading

Clark, John, Yallop, Colin and Fletcher, Janet 2007. *An introduction to phonetics and phonology*. 3rd edn. Oxford: Blackwell.

Davenport, Mike and Hannahs, Stephen 2005. *Introducing phonetics and phonology*. 2nd edn. London: Hodder Arnold.

Gut, Ulrike 2009b. *Introduction to English phonetics and phonology*. Frankfurt: Peter Lang.

International Phonetic Association 1999. *Handbook of the International Phonetic Association*. Cambridge University Press.

Johnson, Keith 1997. *Acoustic and auditory phonetics*. Oxford: Blackwell.

Kent, Raymond and Read, Charles 2002. *The acoustic analysis of speech*. 2nd edn. Albany: Delmar, Thompson Learning.

Wells, John 1982. *Accents of English*. Cambridge University Press.