1

Introduction to Australian English speech production

1.1 Introduction

Speech is the primary medium of communication for almost all human social groups. It is a highly complex skill that involves precise coordination of the structures of the upper and lower respiratory systems. The act of speaking is a uniquely human facility and although many animals have developed elaborate communication systems, none has the ability to formulate novel audible (or written) messages of the type and complexity available to humans.

When we speak, we impart a whole range of interesting information about ourselves over and above the intended linguistic message. Since speech is produced by coordinated movements of the lungs, the larynx and the mouth, including the lips, tongue and jaw, it also contains characteristics specific to the speaker and determined by an individual's physiology. These physiological effects are referred to as **extralinguistic features** and include clues to our size, gender, age and health – aspects that are not under our voluntary control yet are fundamental characteristics that allow our voices to be identified by our friends and acquaintances. Speech also has the potential to reveal something of our emotional or attitudinal state. Signs of emotion such as excitement or sadness, or of attitudes like boredom or enthusiasm, can be reflected in the way we speak, mainly through our patterns of pitch and loudness, and our rate of speaking. As this is ancillary to the linguistic message,

it is referred to as **paralinguistic** information. Additional characteristics of speech can also provide clues to our social group membership such as the regional, socioeconomic or educational groups that we either belong to or aspire to. The distinctive patterns of speech used by different social groups are referred to as accents. Everybody speaks with an accent, the characteristics of which are a consequence of the contact we've had with different people at different times in our lives. Accent also expresses the various identities that we bring to our social interactions. Humans are social beings and we typically adopt symbolic behaviours that are similar to those of the people with whom we share our lives. Speech is therefore an important expression of group membership and it plays a vital function as a symbolic representation of who we are. Our accent reflects our social and cultural history, including aspects of our heritage (for example, speaking English with a French accent) and regional affiliation (speaking Australian English with an Adelaide accent or a Melbourne accent). The accuracy of the social impressions we form about people based on their accent depends on how well we know the accent. Positive and negative opinions about accents are usually a response to the social connotations associated with an accent and not to the linguistic integrity of the accent itself. Each accent is equally capable of meeting the needs of its users.

During the pre-adolescent and adolescent years, when we are developing our sense of identity, it is our peers of similar age or slightly older who are particularly influential in shaping our accent. Accents are typically most fluid during this time but usually become established by the early twenties when a more stable sense of self and identity has developed. However, researchers (for example, Harrington 2006) have found that it is possible for accents to continue changing during adulthood if social/cultural situations or individual aspirations alter. It is important to acknowledge that accents are in constant flux and the description of the speech sounds of any particular accent provides a snapshot of time, place and context. **Synchronic variation** (that is, variation across the accent at one particular time in history) is present in all accents. This variation relates to social effects such as area of upbringing, gender and social group membership, but it also relates to situational effects such as stylistic variability based on levels of formality and the context of any interaction. **Diachronic variation** (that is, variation that occurs over time) is a reflection of linguistic change. Diachronic variation has its roots in synchronic variation because such variability within a speech community is the catalyst for both long- and short-term change.

Listeners to our spoken language are faced with a diverse and complex array of signals about who we are and what we intend to convey, whether deliberate or subconscious. All spoken social interaction involves so much more than simply sending and receiving linguistic messages, and includes the complex process of recognising, responding to and negotiating the

extralinguistic and paralinguistic signals that accompany any interpersonal exchange.

In this book we will focus on Australian English (AusE), one of the core global dialects of the English language. We define dialect as a variety of a language that has its own specific and unique characteristics of vocabulary (the words and their meanings), syntax (the grammatical conventions that govern the creation of words and how they are combined to form meaningful utterances) and phonology (the structure and function of the sound system including how sounds can be combined to form syllables, words and phrases, and how patterns of prominence and intonation influence the spoken output). Accent specifically relates to the phonological characteristics of the dialect. As the focus of this book is AusE speech patterns, we will be concentrating on aspects of phonology and therefore the AusE accent.

Accents vary according to:

- **Phonemic system:** This relates to the categories of contrastive sounds (phonemes) that are used to make words. AusE has the same system of sound contrasts (phonemic system) as Standard Southern British English (SSBE) and New Zealand English (NZE), but differs from many American English (AmE) dialects. For instance, AusE, SSBE and NZE share a contrast between the vowel sounds in words like 'palm' and 'bomb' whereas General AmE doesn't differentiate these two vowels.
- **Phonetics:** The specific detailed characteristics of the individual speech sounds and how the sounds vary in different contexts are phonetic effects. For example, the sound for the letter 'l' in AusE has a number of different variants that are selected according to phonetic context. The 'l' at the beginning of the word 'leaf' is different from the 'l' at the end of the word 'feel'. If you say these words carefully you should be able to hear and feel the difference between the two types of 'l'. This is a phonetic difference. In some accent varieties, such as for speakers from South-East London, the 'l' sound at the end of a word may be more like a vowel, sounding a little like the vowel in 'hood'. In Irish English the /l/ at the beginning of a word and the /l/ at the end of a word are not very different from each other. These are phonetic differences between the two dialects.
- **Phonotactics:** The restrictions on how sounds combine and how they are distributed in the construction of words are phonotactic effects. For instance, AusE does not have 'r' sounds before consonants or pauses in words like 'car' and 'card', but Scottish English does. This is because AusE is a non-rhotic dialect, where 'r' sounds only occur at the beginning of words or before vowel sounds as in 'red' and 'borrow'. In rhotic dialects, such as Scottish English and many varieties of AmE, 'r' sounds occur in all positions without the condition that they occur before vowels.

- **Lexical features:** The characteristics of the individual sounds that occur in particular words are lexical effects. For example, in many varieties of AusE the word 'dance' is pronounced with either the vowel in the word 'cat' or with the vowel in the word 'cart'. The typical pronunciation in NZE, or the AusE varieties spoken in South Australia, is a pronunciation with the vowel found in the word 'cart'.
- **Prosody:** The patterns of intonation, rhythm and prominence are referred to as prosody. For example, the stress pattern of the word 'finance' varies across accents with either the first or the second syllable receiving prominence. This is an example of difference in lexical stress. Prosody also relates to features that extend across entire utterances, including the shape of the pitch contour and its relationship to patterns of prominence. The pitch pattern can be thought of as the tune of an utterance, and the characteristics of tunes vary across different dialects. Many AusE speakers in certain social interactions, particularly when engaged in a narrative, use an intonation pattern sometimes referred to as Australian Questioning Intonation, High Rising Tune (HRT) or more recently as 'uptalk' (Fletcher et al. 2004, McGregor & Palethorpe 2008). HRT is a tune that occurs when statements are produced with rising intonation. Tunes of this type also occur in New Zealand, Belfast and Glasgow varieties of English.

Many of these features will be discussed throughout this book. Individual characteristics and details of the production of the vowel and consonant sounds will be detailed in Chapter 2. In Chapter 3, information about syllable structure, word stress and sentence stress will be given, and in Chapters 4 and 5 phonemic and phonetic transcription, the two different types of transcription that can be used to describe speech, will be explained. In these transcription chapters, an extensive set of exercises is available to assist you in developing competence in the different transcription techniques. All dialects in all languages display accent variation and AusE is no exception to this. AusE varies across speakers of different ages and genders from different regions and social/cultural groups, and it is beyond the scope of this book to detail this variation. For convenience we will therefore focus on Standard AusE as the dominant dialect in the community. Standard AusE in this sense simply refers to the accent type used by the majority of Australians. It does not refer to any notion of correctness or prestige, and is a term adopted to differentiate the dominant variety from others in the community such as Australian Aboriginal English and Ethnocultural varieties. As you read through the material, you may like to consider how your particular accent corresponds to, or differs from, the typical accent type being discussed.

1.2 Phoneme and allophone

There are two main forms of transcription that can be used to document speech sounds: phonemic and phonetic transcription. The goal of phonemic transcription is to represent phonemes. A **phoneme** is an abstract linguistic construct that exists in the mind of a speaker or listener. A phoneme is not an individual sound, but a class of sounds that can be categorised together. Each phoneme can therefore be considered a label to represent a category of sounds. The label encompasses all of the variants of a particular sound. For example, all the different 'p' sounds in the words 'pan', 'nap', 'span', 'spring', 'naps' and 'napkin' are categorised as the same sound by speakers of English, that is, the phoneme /p/. The actual spoken variants of a phoneme are known as **allophones**. Allophones are entities that exist in the acoustic stream of speech, and are often specifically conditioned by phonetic context (that is, the sounds that surround them). For example, the different variants of the phoneme /p/ that occur in 'pan' and 'span' are context-dependent allophones of that phoneme. If you observe these two sounds carefully, you will notice that the /p/ in 'pan' and the /p/ in 'span' are actually quite different from each other. In technical terms the /p/ in 'pan' is a voiceless aspirated bilabial stop (one that allows an explosive puff of air to be expelled from the lips) and the /p/ in 'span' is a voiceless unaspirated bilabial stop (one that doesn't have the same amount of explosive air release). You can feel the difference in the degree of air expelled for each type of /p/ by placing your hand in front of your lips while you say the two words. These two 'p' sounds are context-dependent allophones, as one occurs alone at the beginning of the word and the other occurs after the /s/ at the beginning of a word, but they remain members of the single phoneme /p/. As the differences between the different types of /p/ are not very important for speakers of English, we tend not to notice them. Speakers of other languages, such as Thai for instance, would notice this difference more readily as it is an important contrast in their language.

Alphabetic writing systems are based on phonemic principles and use orthographic symbols to represent individual phonemes so that words can be differentiated in writing. For example, in the words 'pat' and 'pad', speakers of English identify the final sound as the salient difference between the two words. Therefore, the letters 't' and 'd' are necessary in the spelling system to represent the two sounds and differentiate between the words. The words 'pat' and 'pad' represent a **minimal pair** as speakers agree that these two words contain the same sounds except for a single element. In this case the single differentiating element is the final sound. This minimal pair provides evidence that 't' and 'd' represent two separate and contrastive phonemes in the language. The difference is therefore a linguistically significant one. More information about minimal pairs can be found in Chapter 4, Exercise 4A.

There is another difference between the words 'pat' and 'pad' that is usually ignored because it doesn't stand out in differentiating between the words. This difference relates to the characteristics of the vowel sounds. In the word 'pad', the vowel is longer than the vowel in the word 'pat'. This difference does not carry linguistic significance, as speakers/listeners do not recognise it as an important feature. Here we have another example of a context-dependent allophonic difference rather than a phonemic difference. Such allophonic differences help to give the speech its natural sound and native-speaker quality but they do not have a contrastive function. That is, they are not the main cues that differentiate the words in the mind of the speakers/listener.

The selection of allophones is dependent on many factors including the 'rules' of the language that determine how speech sounds are produced when they occur in various phonetic contexts. The language-specific 'rules' or processes necessary to ensure the selection of correct allophones are 'known' to each speaker of the language, albeit subconsciously. The allophones are therefore predictable. Foreign accent is often the result of a speaker using the specific 'rules' or processes of their native language to produce the speech sounds of another language.

1.3 Phonemic and phonetic transcription

In order to convey information about speech production in written form we need a standard set of symbols for this purpose. It is not possible to simply use the letters of the alphabet for transcribing sounds because there are only 26 letters but there are over 40 vowel and consonant sounds in English, and you will be aware that English spelling conventions do not always provide a good indication of pronunciation. Notice, for example, that in AusE the words 'her', 'bird', 'hurt', 'earth', worm' and 'journal' all contain a common vowel sound but the vowel is represented six different ways in the spelling. Alternatively, the words 'dough', 'through', 'rough', 'cough', 'fought' and 'drought' all contain the same sequence of letters 'ough', but each word uses a different vowel sound. Consider words like 'walk', which contains three sounds /wo:k/, 'knee', which contains two sounds /ni:/ and 'music', which contains six sounds /mjʉ:zɪk/. Once we have learned to read and spell, our interpretation of speech patterns is coloured by our knowledge of spelling conventions. Preliterate children tend to interpret the sounds in words phonemically and this can be seen in their early attempts at spelling, for example, 'kwin' for 'queen' and 'crusht' for 'crushed'. The child has interpreted the phonemes correctly but is not yet familiar with the letter to sound correspondence. Transcription of speech sounds requires us to divorce ourselves from the spelling conventions that are so powerful in biasing our interpretation of how words

are pronounced. For example, the word 'ink' does not actually contain an 'n' sound at all. The middle sound is the same as the sound that occurs at the end of the word 'sing'. This sound is represented by the letter sequence 'ng' in 'sing' but 'n' in 'ink'. These are some examples that demonstrate why we cannot use the alphabetic system to represent speech sounds.

Since phonemes are not actual sounds but instead exist as abstract linguistic constructs in the mind, they can be represented by any arbitrary system of symbols. It is convenient, however, to use a standard system of symbols so that we can communicate with each other in writing about speech sounds. The most widely accepted system of symbols for this purpose is the International Phonetic Alphabet (IPA). The IPA is the standard system used for representing the sounds of the world's languages. It is regularly reviewed and maintained by the International Phonetic Association. This alphabet must be used to represent the allophones of a language, in other words it must be used in phonetic transcription, and it is also typically used to represent the phonemes. The IPA does not dictate how phonemes should be analysed or transcribed for particular languages but it does equip us to make informed decisions and choices about the best selection of symbols for phonemic transcription.

When linguists are developing a phonemic description of a language or dialect, they are at liberty to select a subset of symbols from the IPA to represent the important phonemes. They usually represent phonemes using the phonetic symbol associated with the most common or widely distributed allophone of a particular phoneme. When an IPA symbol is used to represent an allophone, it has a very different meaning from when the same symbol is used to represent a phoneme. The symbol for the allophone is a characterisation of the actual sound, whereas the symbol for the phoneme is a label to reflect the group of allophones that comprise the phoneme. For this reason it is necessary to indicate explicitly the type of transcription that has been made so the reader can correctly interpret the transcription. Phonemic transcription must be enclosed in forward slashes / . . . / and phonetic transcription must be enclosed in square brackets [. . .]. This bracketing convention immediately signals to the reader the type of transcription that is intended. The interpretation of the symbols is different in each case.

As an example, take the symbol /k/, which represents an English phoneme. When the /k/ symbol is indicated with forward slashes, the reader interprets this as a phoneme and understands that its actual production can only be ascertained by listening to the sound and examining the characteristics of the allophone produced. Remember that the phoneme is a label for a category of speech sounds. The phoneme /k/ has a very large number of different realisations in speech (allophones). It can be a voiceless aspirated, unaspirated or unreleased velar stop, a palatal stop, a uvular stop or a fricative variant, to name a sample of its realisations. Chapter 2 provides details of

these descriptions. However, if the [k] is indicated with square brackets, the reader can be quite confident that the symbol reflects production of the individual sound. [k] embodies a particular set of articulatory features and uniquely represents the voiceless unaspirated velar oral stop. The IPA chart reproduced at the back of this book shows this symbol in the column for the velar sounds in the top row.

Another important consideration is that the phonetic transcription is language independent and therefore not subject to the 'rules' of any particular language for its interpretation. Phonemic transcription, on the other hand, is language dependent and requires not only a knowledge of the symbol/phoneme relationship but also knowledge of how a phoneme is realised in context in the specific language under investigation. The following examples indicate this language-dependent relationship in two different languages.

In English, the word 'pan' is transcribed phonemically as /pæn/. The initial sound is one of the allophones of the phoneme /p/. Speakers of English 'know' that /p/ in this position (that is, as the single consonant at the beginning of a stressed syllable) is realised as the allophone [p^h], the voiceless aspirated bilabial stop. If the reader of the transcription does not know the 'rules' for allophone selection in English, for instance if they don't speak English as their native language, they may select an allophone (such as the non-aspirated sound) that is not appropriate for the native-like production of the English word. This would lead to a non-native or foreign accented sounding production.

Our second example relates to Fijian language. The town of Labasa is located in the Fijian Islands on the northern island of Vanua Levu. A Fijian phonemic transcription of the town's name Labasa may be similar to /labasa/. Without knowledge of Fijian phonology, a non-native speaker may have difficulty interpreting this phonemic transcription correctly. This is because in Fijian language, consonants like /b/ are prenasalised which means that the /b/ in 'Labasa' is pronounced with a preceding nasal sound like /m/, similar to the combination of 'm' and 'b'. The correct pronunciation of the town's name sounds like 'Lambasa'. A Fijian-speaking reader of the phonemic transcription would be able to interpret the transcription correctly armed with intimate knowledge of Fijian language phonology and the 'rules' for selecting the correct allophones. However, an English reader who does not know Fijian phonology would be unable to interpret this transcription because English does not share the Fijian prenasalisation process. You will notice that the spelling of the Fijian word follows the phonemic convention of just representing the phonemes and not the allophones. If this word was phonetically transcribed, anyone with knowledge of the IPA could reproduce the spoken word correctly because the phonetic transcription would contain the [mb] sequence, the detail required for accurate interpretation.

It is vitally important that you have a clear understanding of the difference between phonemic and phonetic transcription as the material contained in the following chapters is based on these concepts.

In summary, phonemic transcription is useful as a way of indicating the general speech sounds that occur in different words. The phonemic transcription doesn't contain any detail about how each sound is actually produced in different contexts. This is why it is used in dictionaries and as the basis for alphabetic writing systems where exact details of pronunciation may be a hindrance. Phonetic transcription, on the other hand, is a system where specific details of segmental pronunciation are expressed. Here the transcriber and the reader do not need knowledge of the phonological processes of the language in order to interpret the transcription. They do, however, require an intimate knowledge of the IPA and the ability to equate sounds of varying quality with the appropriate phonetic symbols.

1.4 Australian English transcription conventions

English speech sounds vary across dialects, and consequently the transcription of their phonemes can also vary. There is not a single phonemic system that can be considered most applicable for the general transcription of English. For this reason a number of phonemic transcription systems exist for the different dialects of English. There are two main systems used for AusE; these differ from one another mainly in their vowel symbols. The traditional system based on Mitchell (1946) and Mitchell and Delbridge (1965b) (henceforth MD) is employed by the *Macquarie Dictionary* (2009) and is derived from a standard transcription originally devised for British English. This system indicates the congruence between the phonemic systems of SSBE and AusE. The revised system of Harrington, Cox and Evans (1997) (henceforth HCE) uses symbols that are reported to more accurately reflect AusE pronunciation. The revised set of symbols from HCE will be used as the basic system throughout this text but solutions to all phonemic transcription exercises in Chapter 4 will be given in both systems. An evaluation of the two systems for AusE transcription will be provided in Chapter 6.

MD relies on phonemic congruence with SSBE but the cost of this congruence is that it comes at the expense of phonetic accuracy. However, this system does offer parallels with British-based varieties and can therefore provide a general indication of English conventions. HCE is more firmly based on an Australian standard and makes explicit the differences between the British and Australian sound systems. HCE and MD are therefore complementary. It is possible that HCE and MD can coexist as tools for representing AusE, with the choice being dependent on the requirements of the user. The HCE system may be used as an alternative to MD or as a supplement to it. If a reader

requires a system that equates other global British varieties, then MD may be perfectly adequate. However, if a system is required that clearly articulates the phonemic contrasts in the dialect while at the same time providing an indication of unique AusE characteristics, HCE is clearly superior. By more accurately reflecting AusE speech sounds, it can also provide a framework on which to build more detailed phonetic transcription. HCE acknowledges the primacy of the AusE vowel system in the Australian context.

1.5 Australian English

AusE can be considered a regional dialect of English. It is one of the major global English varieties along with other native language Englishes including American English, British English, Canadian English and New Zealand English (Kachru 1985, 1986, McArthur 1987). It is the standard form used in Australia and has its own set of distinctive linguistic characteristics including unique phonology and syntax, and semantic and idiomatic features. Most Australian people use AusE as their native language. It is the dialect of those who are born in Australia or who immigrate during childhood or early adolescence, and it is a powerful symbol of Australian national identity. AusE is increasingly considered an influential new variety of orientation in East Asia (Leitner 2004, Foulkes 2006).

Australia is a diverse multicultural society with approximately 40 per cent of its inhabitants either born overseas or with at least one parent born overseas (ABS 2006). This cultural mix has led to considerable linguistic diversity within the Australian community, with over 200 commonly used languages as well as some vibrant indigenous languages, many of which are endangered (ABS 2006, NILS 2005). Despite the multicultural nature of the society, Australia remains overwhelmingly Anglo-dominant, with 83 per cent of the population speaking only English at home (ABS 2006).

Amongst AusE speakers, there are three major dialect types: Standard Australian English, used by the majority, Australian Aboriginal English, used by many indigenous Australians (Butcher 2008) and Ethnocultural varieties, which are forms used to express non-mainstream or ethnic Australian identity (Clyne et al. 2001). The label 'Australian English' should be considered a superordinate term that embraces all of these various dialectal types. Standard Australian English is the codified variety represented in the national dictionaries such as the *Macquarie Dictionary* (2009) and the *Australian Oxford Dictionary* (Moore 2004), and is the variety used in education, government, the courts, broadcasting and trade. Throughout this text the abbreviation AusE will be used to refer to Standard AusE. The term when used to describe accent does not refer to notions of correctness or prestige. Other AusE