CHAPTER 1

The languages of East and Southeast Asia: a first look

1.1 Introductory remarks

This book is all about the similarities and differences among the languages and "language cultures" of East and Southeast Asia. Moving from south to north, the countries we will cover can be grouped as follows: (a) Indonesia, East Timor, Malaysia, Singapore, Brunei, and the Philippines, (b) Myanmar (Burma), Thailand, Cambodia, Laos, Vietnam, (c) China, (d) Korea and Japan.

How many languages are there in this broad area? Many people would probably answer: about a dozen or so, to match the number of countries involved. This would be roughly correct if we were only talking about national languages, i.e. languages which are the official medium of government, education, and media—Indonesian (Bahasa Indonesia) and Malaysian (Bahasa Melayu), Tetum and Portugese (in East Timor), Pilipino (Tagalog), Burmese, Thai, Lao, Khmer, Vietnamese, Chinese (Modern Standard Chinese), Korean, and Japanese. But if we are talking about all

the languages of the area, including the hundreds of minority languages spoken in the hills and jungles, then the total number is vastly greater. There are hundreds of languages in countries such as China, the Philippines, and Indonesia, and dozens in most other countries of the region.

Languages can be grouped in various ways. One way is by their family trees, i.e. by their historical origins. For example, many people would know that the so-called Romance languages (e.g. French, Spanish, Italian) all descend from a common ancestor language (Latin). They form a small "language family", or more accurately a branch of a larger family which, once all the branches of the family tree (e.g. the Germanic branch, the Slavonic branch) are taken into account, embraces most of the languages of Europe. We say that all these languages are genetically related. How many language families, then, are there in East and Southeast Asia? Setting aside various complications which we will look into later, the short answer is that there are six main families (Austronesian, Mon-Khmer, Tai-Kadai, Tibeto-Burman, Sinitic, and Hmong-Mien), plus several major languages (Japanese, Korean) whose ancestry is unclear. From a genetic point of view, the languages of East and Southeast Asia are much more varied than those of Europe.

The family tree is not the only way of grouping languages. Another approach is to group them according to their similarities to one another—in pronunciation, grammar, and vocabulary. This is called a "typological" approach to grouping. If we do this, then the language situation comes out a bit differently. The reason is that many languages of different families have been in contact with one another—living together, so to speak—for many, many years and have mutually influenced each other in various ways. The result is that languages which are not genetically related nevertheless share many features of language structure. In fact, we can say that most of the languages of mainland Southeast Asia fall into a single "linguistic area", on the basis of pervasive similarities which cross-cut the genetic classification. There are also important cultural similarities across regions of East and Southeast Asia—similarities which have their reflections in vocabulary and semantics, in social aspects of language use, and sometimes even in aspects of language structure. For example, many East and Southeast Asian languages have special forms of respectful or deferential language to be used when talking with or about people who are older or socially higher than oneself.

In this chapter, we are going to survey some of the similarities and differences among languages of East and Southeast Asia, especially the kind of features which are different from English. This will also help us establish some descriptive terminology which we will use throughout the book. For the purposes of illustration, we will stick mainly to major national languages. Then in Chapter 2 we will survey the language families, linguistic areas and language situations of the region to get a truer picture of the language diversity of the region. These first two chapters form an

introductory package. After that, the remainder of the book consists of closer examinations of various topics, both sociocultural and structural.

1.2 Lack of inflection

One feature of many—not all—languages of East and Southeast Asia which is very different from English and other European languages is what linguists call "lack of inflection". In simple terms, inflection occurs when the form of a word changes to indicate an added aspect of meaning which is relevant in a particular grammatical context. In English, for example, nouns have a special form (plural) which indicates that we are referring to more than one of the items in question (*dog, dogs*). We say that English nouns inflect for number, i.e. they have various forms indicating the "number" category (singular vs. plural). Similarly, we say that English verbs inflect for tense, meaning that they have different forms (e.g. *kills, see*, present tense; *killed, saw*, past tense) to indicate whether the event is happening now or happened previously. Notice that in English tense inflection is sometimes done by suffixes (endings), and sometimes by modifying the vowel.

If you look at the following example sentences from Thai, you will see that in this language the verbs do not inflect for tense. That is, the verb for 'read' (*àan*) remains the same regardless of whether we are talking about the present, the past, or the future.

(**1**) a. *Khǎw àan nǎngsɨɨ dĭawníi*.
　　 he/she read book   now.

　 b. *Khǎw dây àan nǎngsɨɨ*.
　　 he/she PAST read book.

　 c. *Khǎw cà àan nǎngsɨɨ*.
　　 he/she will read book.

In (1a) the adverb *dĭawníi* 'now' makes it clear that we are talking about the present time. In (1b) and (1c) the fact that the action is in the past or the future is indicated by separate grammatical words—*dây* and *cà*, respectively. However, sentences in Thai can refer to the present or to the past without this being overtly marked in any way. Example (2), for instance, could be translated into English as either 'He/she sat down' or as 'He/she is sitting down', depending on the context.

(**2**) *Khǎw nâng long*.
　　 he/she sit   down.

Again unlike English, Thai nouns do not inflect for number. That is, although the word *nǎngsɨɨ* in examples (1a)–(1c) has been glossed as 'book', it could in fact be referring not to a single book, but to some particular books, or to books in general. Of course, if the Thai speaker wants to specify

the number of books, that is perfectly possible by using a number word or other quantifying word.

You can also notice from examples (1a)–(1c) that Thai doesn't have any words which correspond exactly to *the* and *a*—the so-called "definite article" and "indefinite article" of English. This is also a widespread feature of the languages of East and Southeast Asia, though often these languages use demonstrative words, i.e. words for 'this' and 'that', in "article-like" ways. Another notable feature is that there is no gender distinction in the third person pronouns; that is, the word *khǎw* is used regardless of whether the person is male or female.

If we look at other languages of mainland Southeast Asia, we find that most of them are like Thai in lacking inflection for tense and number. As well, these languages do not show any inflection according to the grammatical role, such as subject or object, that a word or phrase has in a sentence. Consider the following question–answer pair from Malaysian (Bahasa Melayu), which is Thai's southern neighbour but belongs to a completely different language family.

(**3**) a. *Awak marah pada saya ke?*
    you   angry at   I   QUES
    'Are you angry at me?'

   b. *Ye, memang saya marah pada awak.*
    yes certainly I   angry at   you
    'Yes, I certainly am angry at you.'

In English the first person singular pronoun changes its form depending on its position in the sentence, which is linked with its grammatical role. If it comes before the verb (i.e. in subject position), it is *I*, but if it comes after the verb it takes the form *me*. Most English pronouns have two forms (e.g. *he/him*, *she/her*, *we/us*, *they/them*). The English second person pronoun is unusual in retaining the same form (i.e. *you/you*); in older English it was *thou/thee*. All other European languages have several different pronoun forms for different grammatical "cases" (nominative, accusative, dative, etc.). Anyhow, the main thing we can see from examples (3a) and (3b) above is that in Malaysian (Bahasa Melayu), the pronouns stay the same regardless of their position or grammatical role. That is, *saya* is used, whereas in English we would have to switch between *I* and *me*.

The story is the same in Mandarin Chinese (Modern Standard Chinese). The first person pronoun *wǒ* is used in both subject and object roles, without changing its form.

(**4**) a. *Wǒ xǐhuān Susan.*
    I   like   Susan.
    'I like Susan.'

   b. *Susan xǐhuān wǒ.*
    Susan like   me.
    'Susan likes me.'

To summarize: in most languages of mainland Southeast Asia and East Asia, words do not inflect for tense, number, gender, or grammatical role (case). It is important to point out, however, that this generalization does not apply to all languages in this region. Especially in the northern area there are languages such as Japanese and Korean which have plenty of inflection in their verbal systems. Looking at (5a) and (5b) below, we can see that Japanese verbs have suffixes indicating tense. Suffixes *-ru* or *-u* indicate present (or, non-past) tense, while *-ta* or *-da* indicate past tense.

(5) a. *Akiko ga oki-ru.* 'Akiko is getting up.'
      Akiko ga oki-ta. 'Akiko got up.'

    b. *Akiko ga shin-u.* 'Akiko is dying.'
      Akiko ga shin-da. 'Akiko died.'

These examples also show another feature of Japanese which differs from the other languages we have seen so far. The word *ga* which follows the noun phrase *Akiko* is a grammatical word whose function is to mark the subject of the sentence (more on this in Chapter 4). Since *ga* is a separate word rather than a suffix, this does not qualify as inflection, but it is still a way of explicitly marking grammatical function. In Thai, Malay, and Chinese, in contrast, there is no explicit marking of grammatical function.

Table 1.1 lists some inflected forms of Japanese verbs. Comparing the columns, we can see that the form of a suffix is determined by whether the verb stem ends in a consonant or in a vowel. The Japanese verb system is actually a good deal more complicated than this, but this is enough for our purposes. The main point is that although the majority of East and Southeast Asian languages lack inflection, this generalization does not apply to Japanese or Korean, or to some of the minority languages of the region.

A technical term often used to describe languages like Thai, Chinese, and Vietnamese is "isolating". Essentially the term refers to a language in which

Table 1.1. Some Japanese verb forms

	C-final stem	**V-final stem**
	kir- 'to cut'	*ki-* 'to wear (clothes)'
imperative	*kir-e*	*ki-ro*
present	*kir-u*	*ki-ru*
past	*kit-ta*	*ki-ta*
participial	*kit-te*	*ki-te*
provisional	*kir-eba*	*ki-reba*
tentative	*kir-oo*	*ki-yoo*

words have a single, fixed shape; that is, words do not change their shape as a result of affixation or other morphological processes. In European languages, nouns and verbs do change their forms as they inflect for categories such as case, tense, number, and gender, and so these languages are not isolating. In effect, isolating implies non-inflecting.

Since Korean and Japanese do have inflection, they are similar to European languages in this respect, but there is still a significant difference. In European languages it is common for a particular suffix to combine several different inflectional meanings; for example, the English verb suffix -*s* indicates 'third person singular subject, present tense'. In Japanese and Korean, however, it is more usual for each inflectional suffix to represent a single category. If it is appropriate to mark several inflectional categories at the same time, then several different suffixes are used, one after another. Languages whose inflectional morphology bundles several meanings into a single form (like French and German) are known as fusional. Languages which allot one morpheme per inflectional meaning (like Japanese and Korean) are known as agglutinating.

It's important to point out that a language can be isolating and still have complex word forms. Being an isolating language is not a matter of the internal complexity of words, but rather of how words behave when they occur in different grammatical contexts. Some of the classic isolating languages of Asia, such as Mandarin Chinese and Vietnamese, have a high proportion of complex words formed by compounding or by reduplication.

1.3 Word order (constituent order)

Linguists often use the term "word order" in a rather loose way to refer, not to the order of words in a strict sense, but to the order of the major constituents of a simple clause, i.e. the subject (S), the object (O), and the verb (V). The most common order of these constituents in a simple, ordinary clause is often called the "basic word order" of a language. Various permutations are found in the world's languages. For example, the verb (V) can come in the middle (medial position), as in English; it can come at the beginning of the sentence (initial position); or it can come at the end of the sentence (final position). Regardless of the position of the verb, however, most of the world's languages place the subject before the object. This means that the most common constituent orders are SVO (as in English), SOV, and VSO. All three of these are found in languages of East and Southeast Asia, though VSO order is much rarer than the others. In Chapter 4, we will expand this discussion by distinguishing between the notions of subject and topic, and even questioning whether "subject" is a valid grammatical category in some languages of the region, but we can ignore these complications for now.

1.3.1 Verb-medial, verb-final, and verb-initial languages

Most of the mainland Southeast Asian and East Asian languages have SVO order, including Thai, Malay, Vietnamese, Khmer, and all the Sinitic (Chinese) languages (Clark 1989). A few more examples of SVO order follow. In (7) and (8) I use square brackets to indicate multi-word noun phrases.

Vietnamese

(6) *Tám yêu Hiến.*
    Tam love Hien
    'Tam loves Hien.'

Cantonese

(7) *Ngóh jìdou [go daapon].*
    I   know   CL answer
    'I know the answer.'

Hmong

(8) *[Tus dev] tom [tus npua].*
    CL dog   bite CL pig
    'The dog bit the pig.'

Most Tibeto-Burman languages have the verb at the end of the clause in SOV order (Okell 1994), and so do the North Asian languages, Korean and Japanese. Notice that Korean marks the grammatical role of noun phrases by means of special little grammatical words (particles), as does Japanese.

Burmese

(9) *Ko thè hlain-gá əko-shi-go   meʔgəzìn-de hmanhman*
    name-SUBJ   brother-place-to magazine-PL regularly
    pó-pè-ba-dɛ.
    send-give-POL-TENSE
    'Ko Htay Hlaing sent magazines regularly to his brother.'

Korean

(10) *[John i]   [Mary ege] [chaek eul] ju-ess-da.*
     John SUBJ Mary DAT   book OBJ   give-EXP-PAST
     'John gave Mary the book.'

As for the rarer verb-initial order, this is found in Pilipino (also known as Tagalog), the national language of the Philippines (Schachter 1987). Don't worry about the unfamiliar glosses. The main thing is that the verbs *magaalis* 'will take out' and *nakita* 'saw' come first in the sentences.

(11) *Magaalis   [ang bata] [ng laruan] [sa kahon].*
AT:will.take.out TRIG child PAT toy   DIR box
'The child will take the toy from the box.'

(12) *Nakita   [ng Juan] [si   Maria] kahpon.*
PT:PERF:see ACT Juan   TRIG Maria yesterday
'Juan saw Maria yesterday.'

After this brief look at the main constituent order variations, I should point out that we have only been talking about the most common or neutral order in each language. As a matter of fact, almost all languages of East and Southeast Asia allow some variation in the constituent order of a simple sentence. It is often possible to swap around the order of the subject and object: that is, to put the object first in order to focus some extra attention on it (in some languages, this kind of switch-around is accompanied by some other grammatical changes as well). Generally speaking, the languages of East and Southeast Asia tend to have a more flexible and "expressive" word order than English.

1.3.2 Prepositions or postpositions?

There is often a correlation between the basic constituent order of a clause and the order in which other kinds of elements occur (which is one of the reasons linguists are interested in constituent order). The most relevant factor is generally the position of the verb. Verb-final languages tend to have "postpositions" instead of prepositions. Postpositions are simply words which are functionally equivalent to prepositions but which come after the noun phrases they relate to.

Since Japanese and Korean are verb-final languages, we would expect them to show this property—and they do. For example, in the Japanese sentences in (13a)–(13c) below, the postposition *e* indicates motional 'to'; postposition *ni* indicates location 'at' a place; and postposition *de* indicates 'by (means of)'.

(13) a. *Taroo ga   gakkoo e ikimasu.*
Taroo SUBJ school to goes
'Taroo goes to school.'

b. *Masumi ga   gakkoo ni imasu.*
Masumi SUBJ school at is
'Masumi is at school.'

c. *Kazuko ga   basu de kimasu.*
Kazuko SUBJ bus by comes
'Kazuko comes by bus.'

As we've already seen, both Japanese and Korean make use of particles (such as *ga*) to indicate grammatical function, e.g. subject or object. These particles too come after the phrases they mark.

1.3.3 Word order in questions

In English, and most other European languages, questions have a different word order from affirmative sentences. For example, in questions containing an interrogative pronoun such as 'who', 'what', and 'where' (information questions or "wh-questions"), the interrogative pronoun generally comes at the beginning of the sentence. The special word order is one of the grammatical indications that a question is being asked. Many languages, however, do not use any special word order in questions. In Thai, for example, the order of constituents in a question is the same as in a statement.

(14) *Khǎw àan àray dǐawníi?*
he/she read what now
'What is he reading now?'

Another kind of question is known as a "polar question" ("yes/no question"). In English, this kind of question is formed by means of "subject–auxiliary inversion", i.e. instead of the normal English word order in which the subject comes first, an auxiliary verb (such as *will* or *do*) comes first. In Thai, polar questions are formed simply by adding a question-particle to the end of the sentence. There are two main question-particles: *máy* and *rɯ̌ɯ*. Particle *máy* creates a more or less neutral question, whereas *rɯ̌ɯ* indicates that the speaker expects that the answer will be positive.

(15) a. *Khun cà   pay hǎa phɯ̂an máy?*
    you   will go see friend QUES
    'Will you go and see a friend?'

b. *Khun cà   pay   hǎa   phɯ̂an rɯ̌ɯ?*
    you   will   go   see   friend   QUES
    'So you're going to see a friend, are you?'

Similar examples could be given from Chinese, Japanese, or Korean. On the other hand, there are also languages (e.g. Malay) which do have a special word order for questions, or at least, for wh-questions. This is another reminder that whatever generalizations we make about the languages of East and Southeast Asia, there are almost certainly going to be exceptions somewhere.

1.4 Sounds and writing

From a pronunciation point of view, the most famous aspect of the languages of East and Southeast Asia is, no doubt, the existence of "tone" in languages such as Chinese, Vietnamese, and Thai. This means, roughly speaking, that words in these languages each have a distinctive intonation pattern and have to be pronounced with the correct intonation in order to be recognizable. We will look at tone in the next section. Meanwhile, however, what about the consonants and vowels?

As far as consonants are concerned, few of the major national languages present very severe pronunciation problems for an English speaker—at least, not compared with languages from various other parts of the world, e.g. central America, southern Africa, the Caucasus, which have a great number of difficult or "exotic" sounds to contend with. The most straightforward are languages like Malay and Tagalog, in the south of Southeast Asia, and Japanese toward the top of North Asia. These languages have fairly small numbers of consonants, which are distinguished from one another in ways which are fairly familiar to speakers of English. For example, they have distinctive "voicing" of stops; e.g. contrasts between /p/ and /b/, between /t/ and /d/, and between /k/ and /g/.

In the languages of mainland Southeast Asia and China, voicing is usually not distinctive. Instead, one commonly finds contrasts between "aspirated" and "unaspirated" stops; e.g. contrasts between /p/ and /p^h/, between /t/ and /t^h/, between /k/ and /k^h/. Some of these languages also have consonants which are pronounced at different places of articulation to those of English. There can also be unfamiliar manners of articulation, as with stops in Korean or Khmer. Generally speaking, however, the overall size of the consonant inventories in the languages of mainland Southeast Asia and China is not very large, if we stick to national languages.

Still on the subject of consonants, most languages of East and Southeast Asia have more restricted possibilities for consonant clusters than does English, which has many two-consonant clusters, e.g. in words like *snake*, *tree*, and *clean*, and even some three-consonant clusters, e.g. in words like *strike*. It is quite typical for these languages to allow no more than a single consonant at the beginning of a word. Usually, just about any consonant is acceptable word-initially, but at the end of a word, the possibilities are usually more restricted. In most languages only a handful of consonants can occur word-finally, in some languages none at all.

When it comes to vowels, the situation is rather different. It is true that Japanese and some of the languages of insular Southeast Asia have fairly small and simple vowel systems, but most other languages of the region have fairly complex vowel systems. Even when the number of vowel phonemes is smallish (say, five or six as in Chinese), they can often occur in diphthongs

(two-vowel combinations) and even triphthongs (three-vowel com-binations)—thus multiplying the overall number of vowel sounds in the language. In fact, Southeast Asia is home to languages (such as Khmer) which can lay claim to having the greatest number of vowel sounds in the world.

We can make two generalizations about the sound systems of this broad region. First, the languages in the centre of the area (mainland Southeast Asia and China) tend to be more phonologically complex than those on the periphery. Second, phonological complexity tends to be concentrated in the realm of vowel sounds. However, as before, we have to qualify such gen-eralizations with the acknowledgement that there are exceptions, sometimes striking ones, as one would expect of an assemblage of languages as numerous and as diverse as those in East and Southeast Asia.

Most people know that important national languages such as Chinese Japanese, Korean, and Thai are not usually written using roman letters, i.e. the familiar ABC letters used for English and other languages of Western Europe. In their own countries these languages are usually written in entirely different scripts, like those illustrated below. For outsiders such scripts are completely unintelligible. We can't even tell where one word ends and another begins, let alone figure out any of the pronunciation.

คำภาษาไทย
'words in the Thai language'   Thai

中 國 文 字 最 為 美
'Chinese characters are the most beautiful'   Chinese

お習字　で　習った　字
'the characters that we learned in calligraphy'   Japanese

우리　나라의　말은
'our country's language'   Korean

For anyone learning one of these languages, learning how to read and write in the national scripts is indispensable. In some cases, as with Chinese and Japanese, it is a task which requires a great deal of time and study. Fortunately, for the purposes of this book we do not need to learn any new writing systems. We will use roman letters to represent the words of all the languages we will discuss.

Although writing systems are the topic of a chapter of their own (Chapter 6), it is worth quickly mentioning a couple of points here to dispel some common misconceptions. You may have heard it said that Chinese characters stand not for sounds (in the way that roman letters do, however imperfectly) but for ideas or meanings. You may even have heard that Chinese characters are like pictures—albeit cryptic ones—of the meanings they stand for. Though there is a grain of truth in these ideas, they are

essentially mistaken. First, the pictorial basis for Chinese characters is very slight. Second, 90 per cent of commonly used characters do not represent meanings alone: they have both phonetic (i.e. sound-based) and semantic (i.e. meaning-based) components.

A second misconception about the non-roman writing systems of East and Southeast Asia is that they are all like Chinese in being based on "characters". This belief is partly true in relation to Japanese, but it is completely false in relation to Burmese, Thai, and Korean. The writing systems used for these languages are essentially alphabetic systems, just like the roman letter system we use for English—the main difference is simply that the letters in Burmese, Thai, and Korean writing are different. Of course, we should not forget that many languages of the region, such as Indonesian, Pilipino (Tagalog), and Vietnamese, are normally written using ordinary roman letters.

1.5 Lexical tone

One of the most famous aspects of many languages of East and Southeast Asia is generally known as "tone". Most people probably know that in Chinese, for example, the same sequence of sounds can have several different meanings depending on the tone in which it is pronounced. Using a more precise terminology we call this phenomenon "lexical tone", to make it clear that tone is part of the lexical identity of words in Chinese. That is, it is an integral part of the pronunciation of the word, and it serves to distinguish one word from another. For speakers of a tonal language, the tones are not extra features added to the segmental shape of a word. Speakers often have no auditory image of the word at all independent of tone, and a mispronunciation of tone can make the word completely unrecognizable.

But what is this thing called tone? Basically, to pronounce words with tones means to use the pitch and quality of your voice in various ways, not just to convey emphasis or to convey attitudes or feelings as we do in English, but on a word-by-word basis. Some tones are "level", meaning that they do not change very much in pitch. In Mandarin Chinese, for example, one of the four tones is called the "high-level" tone. It is found in the word ma^{55} 'mother'. The superscript numerals refer to a set of five pitch levels in your voice, with 5 as the highest and 1 as the lowest level. The first digit indicates the starting pitch level and the second digit indicates the final pitch level. So a tone notated as 55 means using a consistently high pitch. Of course, the actual pitch level varies according to every individual's voice. A 55 tone spoken by someone with a very low voice (e.g. an old man) would be much lower, in absolute terms, than a 55 tone spoken by someone whose voice is generally higher (e.g. a young girl). It is not the absolute pitch that matters but how it sits relative to each individual's voice.

Aside from "level" tones (technically called register tones), there are also "moving" tones (technically called contour tones), which change pitch as the word is being pronounced. Mandarin Chinese has three moving tones. The word ma^{35} means 'hemp', for example, and ma^{51} means 'to scold'. So a 35 tone starts somewhere in the middle of a person's pitch range, and moves up towards the top of the range. This tone is usually called "high rising" (though "rising to high" would be a more accurate description). The 51 tone starts at the top of the range and falls steeply. It is often labelled as "high falling" (though "falling from high" would be a better description). With both the 35 tone and the 51 tone, the pitch change moves in only one direction (rising or falling, respectively). Another type of moving tone involves a fall followed by a rise, or a rise followed by a fall. The final Mandarin tone is of this type. For example, the word ma^{214} means 'horse'. It starts a little low, dips even lower, then moves up (hence the label "falling-rising").

The four tones of Mandarin are sometimes notated as shown in the second column of Table 1.2. These are simplified graphs of the pitch over time. The vertical line at the right of each one just serves as a reference for the possible pitch range.

Tones can be a problem for writing systems. It isn't very economical to use two superscript numbers. So for practical purposes various other systems are usually used. The first is just to assign each tone a single identifying number: for example, we could call the 55 tone "tone 1", the 35 tone "tone 2", the 214 tone "tone 3", and the 51 tone "tone 4". Thus we could write the four words in Table 1.2 as ma^1, ma^2, ma^3, and ma^4. In this kind of system the numbers do not indicate the pitch level: they are just identifying numbers.

Another approach is not to use numbers at all. After all, they look messy and take up extra space. Instead we can use diacritics (i.e. extra marks) above the vowels of each syllable. This is what is done in the official Pinyin system for writing Mandarin Chinese using roman letters, as shown in the third column of Table 1.2. As you can see, the diacritics correspond to the pitch graphs. I can't resist illustrating here with one of the many "tongue twisters" which have been devised based on tone contrasts: *Māma qí mǎ. Mǎ màn. Māma mà mǎ.* 'Mother rides horse. Horse slow. Mother scolds horse' (Crystal 1987: 172). Strictly speaking, tongue twister is not really an accurate description, however, because it is the larynx or voice-box which is involved

Table 1.2. Four [*ma*] words in Mandarin Chinese, Beijing dialect

[*ma*]^{55}		*mā*	'mother'
[*ma*]^{35}		*má*	'hemp'
[*ma*]^{51}		*mà*	'to scold'
[*ma*]^{214}		*mǎ*	'horse'

in tone production, not the tongue (see chapter 5). Other sets of four-way tone contrasts are: *fān* 'turn over', *fán* 'annoying', *fǎn* 'opposite', *fàn* 'meal'; and *tāng* 'soup', *táng* 'sugar', *tǎng* 'lie', *tàng* 'hot'.

Tone diacritics work fine for Mandarin Chinese, but they don't work so well for languages with more than four tones, such as Cantonese or Thai, because the number of diacritics gets confusing and difficult to read. Another option is to coopt some letters of the alphabet to indicate tones, either by themselves or in combination with diacritics. For example, the Romanized Popular Alphabet for Hmong (which has eight tones) uses final consonant symbols to indicate tones: *-b* for 'high level', *-j* for 'high falling', *-v* for 'mid rising', *-g* for 'low falling breathy', and so on. This is made possible by the fact that all words in Hmong end with a vowel or with /ŋ/. A limited use of letters to indicate tone is found in Matthews and Yip's (1994) system for writing Cantonese. They use a syllable-final letter *h* to signal that the tone is one of the low tones, with diacritics to add further specification. For example, *ah* stands for the vowel /a/ with a 'low level' tone, *áh* stands for the same vowel with 'low rising' tone, and *àh* for 'low falling' tone. It is feasible to use the letter *h* in this way because the phoneme /h/ is never found at the end of a syllable in Cantonese.

Languages with lexical tone can differ substantially from one another in the nature of the tonal system. Not only can there be different numbers of tones, but the quality of tones differs from language to language. For example, Cantonese has six contrastively different tones, including three "level" tones as shown below (Matthews and Yip 1994: 21). The mid level and low level tones have no counterparts in Mandarin Chinese. Vietnamese also has six tones (Thompson 1987), but they do not match those of Cantonese.

Hong Kong Cantonese

high level	55 *yau* 'worry'	high rising	35/25 *yau* 'paint'
mid level	33 *yau* 'thin'	low rising	23/13 *yau* 'to have'
low level	22 *yau* 'again'	low falling	21/11 *yau* 'oil'

Some languages have been reported to have even more tones. As just mentioned, Hmong has eight tones and so does Chaozhou (a Sinitic language). In many languages, tones can combine and influence one another in speech. The rules governing this (called "tone sandhi" rules) also differ from language to language. Although I have been concentrating on pitch contour, it has to be stressed that in many languages other aspects of "voice quality" work together with pitch in constituting the distinctiveness of tones. For example, Burmese is said to have four tones, but one of the high-pitched tones is distinguished by "creaky" pronunciation (the syllable is pronounced with tension or constriction in the throat) and another by a final glottal stop, which shortens the duration of the vowel. Several of the

Vietnamese tones are also distinguished by strong glottalization. We will return to tones in more detail in Chapter 5.

Finally, remember that although many East and Southeast Asian languages have lexical tone, many others are not tonal. In the south, the national languages Malaysian (Bahasa Melayu), Indonesian (Bahasa Indonesia), and Pilipino have no traces of lexical tone whatsoever. To the far north, Japanese and Korean are not tonal. Most of the minority languages of mainland and insular Southeast Asia are not tonal either.

1.6 Classifier constructions

Another celebrated feature of the languages of East and Southeast Asia is the existence of classifier constructions. We say that a language has classifier constructions when it has grammatical devices which, in certain contexts, oblige speakers somehow to categorize the person or thing they are speaking about (the referent) in terms of certain semantic dimensions. For living things, these dimensions may include whether or not the referent is human, and if not, what kind of life form (e.g. animal, fish, plant) or functional category (e.g. edible, dangerous) it belongs to. For inanimate things, semantic dimensions may include physical properties (e.g. shape, size, material) and functions (e.g. vehicle, tool).

Many different types of classifier construction are found in the world's languages (Aikhenvald 2000), but in East and Southeast Asian languages the most common type is the numeral classifier construction. Numeral classifiers are found within certain kinds of noun phrase, appearing next to numerals and other quantifying expressions, and sometimes also with demonstratives. The classifier may be a word or, less commonly, an affix. To get a quick idea of how classifiers work, look over the following extract from a Japanese shopping list. In Japanese, classifiers are suffixes which attach to the numerals.

nasu [eggplant]   *nana* [seven] *-ko* [CL: smallish, solid thing]   'seven eggplants'
kyuuri [cucumber] *hachi* [eight] *-hon* [CL: long cylindrical thing] 'eight cucumbers'
hamu [ham]   *juu* [ten] *-mai*   [CL: flat thing]   'ten slices of ham'

The examples below show classifier constructions in Malaysian (Bahasa Melayu), Thai, and Cantonese, languages which are more typical of East and Southeast Asia in having classifiers which are separate words. By the way, as you can see from the Cantonese example in (18), the term "numeral classifier" is a slight misnomer, because classifiers are found with quantifying expressions of all kinds, including words like 'many' and 'several', as well as numerals in the strict sense.

Malaysian (Bahasa Melayu)

(16) *empat ekor   kucing*
     four   CL:ANIMAL cat
     'four cats'

Thai

(17) *burìi   sɔ̌ɔng múan*
     cigarette two   CL:STICK-LIKE
     'two cigarettes'

Cantonese

(18) *géi   gāan   ngūk*
     several CL:BUILDING house
     'several buildings'

Classifier constructions have a structure similar to English expressions like *two sheets of paper*, *two drops of water*, *two articles of furniture*, and *two members of the family*. English words like *sheet*, *drop*, *article*, and *member* are not true classifiers, however, because they do not categorize the things being spoken about. They are better termed either "measure words" or "unit counters". We call a word a "measure word" when it specifies the form in which an amount of a mass substance is found; for example, in phrases like *two sheets of paper* or *three drops of water*. The term "measure word" is also applied to units of measurement and "containerfuls", for example, in phrases like *two litres of milk* and *three cups of sugar*. The term "unit counter" is used for words which single out a number of individuals from a collective; for example, in phrases like *two articles of furniture* and *three members of the family*. The languages of East and Southeast Asia also have various measure words and unit counters, in addition to classifiers. Unfortunately, some descriptions of these languages describe measure words and unit counters as subcategories of classifiers. This is confusing, because it gives the impression that these languages have scores or hundreds of classifiers, which is not true.

Numeral classifiers originate historically from ordinary nouns with concrete meanings. In some languages, classifiers are still identical in form with ordinary nouns. In Malaysian (Bahasa Melayu) the classifier *ekor*, for example, is identical with the ordinary noun meaning 'tail'. On the other hand, in many other languages the classifiers are completely unrelated to ordinary nouns.

It might head off potential misunderstandings to point out the main difference between a numeral classifier system and a system of gender (noun class), such as we find in German, French, and many other European languages. A gender system is essentially a classification of nouns, i.e. of words. Though there is always a semantic basis for parts of a gender system (e.g. for words referring to humans), for the most part the system is not semantically based. Gender is also intimately connected with the grammatical process of

"agreement"; that is, the phenomenon whereby other words in the clause adopt different forms to "agree with" the gender of a noun. The number of genders is usually fairly small and quite fixed, e.g. in European languages: masculine, feminine, and (sometimes) neuter.

Numeral classifiers do not classify nouns but the referents of nouns—the actual things in the world which the speaker "picks out" to say something about on a particular occasion. The basis for the system is always predominantly, if not exclusively, semantic. Because of this, it is usual in some circumstances to have a choice of classifiers, depending on how one is viewing the referent in question. To illustrate this from Cantonese (Matthews and Yip 1994): *chēung* 'gun' could be classified as *jī* (by its cylindrical shape) or as *bá* (by its function); and *syùhn* 'sailing vessel' could be classified as *ga* (large vehicle) or as *jek* (small object), yielding the meanings 'ship' vs. 'boat', respectively. Unlike genders, numeral classifiers are not normally involved in grammatical agreement processes. They are often found in sizable numbers, and sometimes form a "semi-open" class, lacking clear boundaries.

Classifier systems may differ according to: (a) the kinds of semantic parameters involved, (b) whether the system is exhaustive or partial, i.e. do all nouns receive classifiers, or only some, (c) whether the use of classifiers is obligatory or not. We will return to classifiers in greater detail in Chapter 4.

1.7 Serial verb constructions

A serial verb construction involves two or more verbs, all of them sharing a single grammatical subject, packed into a single clause without any intervening conjunctions. They are widespead in the languages of East and Southeast Asia, especially in those of the isolating type. As we will see later (in Chapter 4) there are several different kinds of serial verb constructions and they can serve various grammatical purposes. Just to get the general idea, however, take a look at these examples from three genetically unrelated languages—Mandarin Chinese, Khmer, and Hmong. They are typical of serial constructions in that they involve at least one verb which is a verb of motion or posture.

Mandarin Chinese

(19) *Tā gùi-xialai   qíu Zhāng-sān.*
     3SG kneel.down beg Zhang-san
     'He knelt down begging (or: knelt down and begged) Zhang-san.'

Khmer

(20) *Viə deek luək.*
     3SG lie   sleep
     'He lay sleeping (or: lay down and slept).'

Hmong

(**21**) *Nws khiav rov   qab mus tsev*.
    3SG run   return back go   home
    'He ran back home.'

In each case, you can see that the sentence contains more than one verb—'kneel down' and 'beg' in (19), 'lie' and 'sleep' in (20), 'run', 'return' and 'go' in (21)—inside a single clause. Notice that there are no conjunctions (i.e. words like English 'and') separating the verbs.

You may be able to tell from the translations that in each case there is a very close semantic connection between the verbs. In a sense, each of these sentences depicts a single event, even though this event can be seen as consisting or two or more closely related "sub-events". For example, sentence (19) can be understood as depicting either a single event of "kneeling down begging" (both actions happening at the same time) or a single event of "kneeling down, then begging" (one action following the other but both seen as part of a single two-part event).

In cases where the events are understood as happening simultaneously, English has a somewhat similar grammatical option involving the so-called "gerund" form of the verb, i.e. the *ing*-form; as in *He knelt down begging* or *He lay sleeping*. The difference between this kind of English sentence and a genuine serial construction are: (a) the *ing*-form is a marked verbal form, whereas in the Chinese, Khmer and Hmong examples above all the verbs appear in their normal, unmarked forms; (b) the *ing*-form can only be used where simultaneous actions are involved, whereas serial constructions can indicate sequences of closely related actions; (c) in many cases, it is possible to combine three or more verbs in a serial construction.

Be careful not to assume that any old sequence of verbs is a serial verb construction. In most East and Southeast Asian languages, it is possible to get sequences of verbs which do not fit the definition of serialization either because they are not all in a single clause or because they do not all share the same subject. We will only look at some examples of the first kind, from Lao (Enfield 1994: 26; 1998 p.c.).

(**22**) *Khòòj⁵ jaak⁵ lin⁵   phaj⁴*.
    I   want play cards
    'I want (to) play cards.'

(**23**) *Laaw² sùù⁴ khaw⁵-niaw³ kin³*.
    3SG   buy rice-sticky   eat
    'She bought some sticky rice (to) eat.'

In (22) the phrase *lin⁵ phaj⁴* 'play cards' is a complement of the verb *jaak⁵* 'want', but because there is no explicit marking of the fact, it looks like a serial construction. In (23) the word *kin³* 'eat' is a purposive clause even

though it is not marked as such (in fact, the purposive marker *phùa¹* 'in order to, so that' can be inserted in front of *kin³* 'eat' with no change of meaning).

1.8 Multiple pronouns and other systems of address

One of the most interesting features of many of the languages of East and Southeast Asia is that aspects of the language structure seem to be specially adapted to allow people to express "social messages", especially messages to do with differences in social standing, respect, deference, and the like. From the perspective of speakers of these languages, the pronouns of modern-day English are particularly insensitive to social distinctions. There is only one form each for 'I' and for 'you', and English speakers feel free to use these words even with social superiors, such as respected older relatives, one's boss, or teachers. At least in other European languages there are two different ways of saying 'you', e.g. French *tu* and *vous*, German *du* and *Sie*, one of which expresses greater social distance than the other. Many East and Southeast Asian languages go far beyond European languages in having a whole range of alternative pronominal and quasi-pronominal forms (such as kin terms).

Table 1.3 shows some address forms in Thai, a language with a particularly wide range of choices. The choice depends on the speaker's attitude towards the addressee, which is normally conditioned by the nature of the social relationship and by factors such as age and sex. You will see that there are some gaps in Table 1.3. This reflects the fact that in some situations one would not feel comfortable using pronouns at all (preferring to use names, or to avoid direct reference altogether). Also, there are additional pronouns *kuu* 'I' and *mɨng* 'you', which are plainer and "lower" than any shown in Table 1.3. For third persons, the choices are fewer. *Khǎw* is the general polite form, and *thân* expresses special respect for social superiors. There is a third form, *man*, which is used for social inferiors, for non-humans, and when expressing anger. Additional, special forms are used with royalty and clergy.

A more typical-sized system is found in Malay, which offers its speakers essentially two choices each for first and second person pronouns. As in Thai, the choice ultimately depends on one's attitude to the person one is speaking to, but this is strongly conditioned by factors such as social standing and relative age. One would not normally use the plain forms *aku* 'I' and *kau* 'you' unless one is on rather casual or intimate terms with the addressee. The plain forms are used in "unguarded" situations, for example, in thinking aloud to oneself, among intimates and familiars in informal situations, and (sometimes) from senior family members to junior family members. A good description of pronominal variants and usages can be found in Koh (1990: 104–46). She describes *aku* 'I' and *kau* 'you' as

Table 1.3. Some Thai forms of address (after Hudak 1987: 41; adapted with input from Nick Enfield)

Situation	First person 'I'	Second person 'you'
Polite conversation with strangers and acquaintances	*phǒm* (used by males)[a]	*khun*
Speaking to a superior, showing deference	*phǒm* (used by males)[a]	*thân*
'Plain' forms: for informal conversation with close friends/family	*chǎn*	*thəə*
Between intimates of same sex	*kan*	*kɛɛ*
Adult to child	*chǎn*, or kin-term	*nǔu* (lit. 'mouse') or kin-term
Child to adult, e.g. teacher	*nǔu* (lit. 'mouse')	kin-term
Child to older sibling	*nǔu* (lit. 'mouse')	(*phǐi* 'elder sibling')

[a] According to standard descriptions of Thai, *dìchǎn* can be used by females, but it is very formal and extremely rare in face-to-face conversation.

Table 1.4. First and second person pronouns in Malay and Japanese

		First person 'I'	Second person 'you'
Malay	Plain	*aku*	*kau* (or: *engkau*)
	polite or formal	*saya*	*awak, kamu, anda*
Japanese	plain ('rough')	*ore*	*omae*
	casual	*boku*	*kimi*
	normal ('polite')	*watashi*	*anata*
	formal	*watakushi*	—

"non-polite" rather than as "impolite", saying: "these forms do not indicate impoliteness when used correctly. They merely do not express extra politeness" (1990: 111). Because Malay society is extremely conscious of verbal politeness, however, the more formal pronouns (such as *saya* 'I', *awak/kamu* 'you') are required in a wide range of social situations.

Japanese also has linguistic devices for expressing social messages (see the section on "honorifics" below). When it comes to pronouns, there are several options depending on the "social tone" which the speaker wants to convey. Using the very "plainest" pronouns *ore* 'I' and *omae* 'you' often conveys a "rough" tone. The alternatives *boku* 'I' and *kimi* 'you' convey

a "casual" tone. In traditional society, both these sets of pronouns tended to be confined to men and boys. It was seen as socially inappropriate for woman and girls to use them. The normal pronoun choices in polite interaction are *watashi* 'I' and *anata* 'you'. In addition, there is a specially formal variant for 'I', namely *watakushi*.
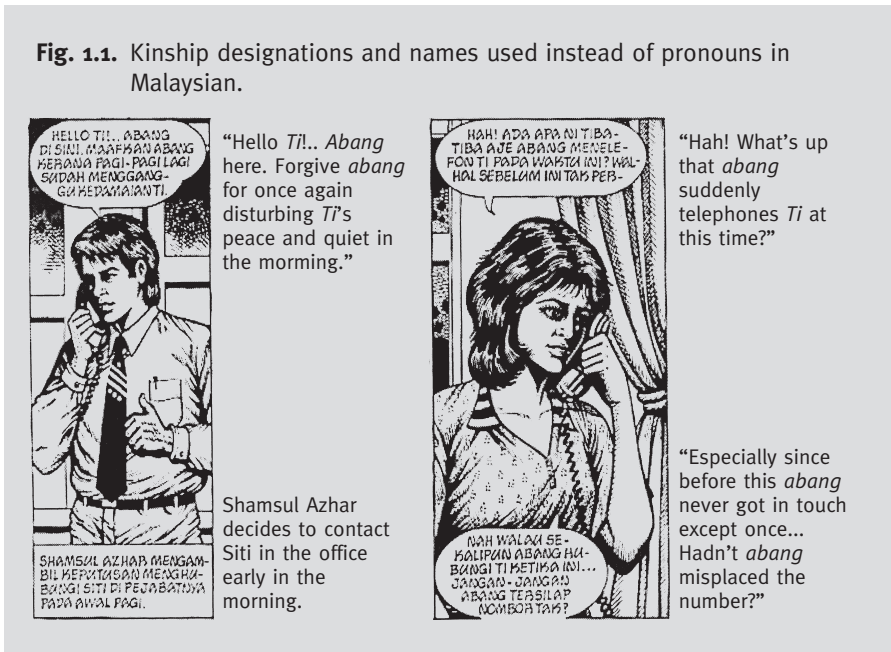
Notice that the three examples we have looked at here come from Southeast Asia and from North Asia. It is an interesting fact that the languages which are geographically in-between these regions, i.e. the Sinitic (Chinese) languages, do not have elaborate pronoun forms (though they use various other means of expressing respect and deference in speech).

Despite the range of pronoun choices, there are many situations in which it may feel better not to use pronouns at all. Alternative strategies include using some other way of referring, such as personal names, kinship designations, or words for a person's occupation or role, or avoiding explicit references altogether. In Malaysia, it is normal in everyday interaction to refer to oneself and one's interlocutors not by any of the pronouns, but by using a person's name or "kinship" designation, such as *abang* 'big brother', *kakak* (or *kak* for short) 'big sister', and *adik* 'younger brother or sister'. The inverted commas around the term "kinship" are necessary because terms like *abang*, *kakak*, and *adik* are used not only between real family members but in a range of other situations where there is a small or moderate age difference between the people involved. Similarly, in conversation with someone much older, it is normal to address the older person as *pakcik* 'uncle' or *makcik* 'auntie', and for that person to refer to him or herself in the same way.

The cartoon in Fig. 1.1 shows names and kinship designations being used in place of pronouns. Shamsul Azhar is telephoning his close friend Siti to apologize about a misunderstanding. Throughout the exchange both participants refer to him as *abang*, literally, 'big brother', and to her as *Ti* (a shortened form of *Siti*). No pronouns are used at all.

Our Malaysian example concerned a friendly social situation, between people who knew each other well and were more or less equal in social status. In asymmetrical social situations, i.e. when the people involved are of different social level or status, it is even more common to find that pronouns are avoided. For example, in Japanese one would not use the polite *anata* 'you' when speaking to a higher-status person. As Inoue (1979: 284) puts it: "Should one dare to use *anata* in speaking to a person of higher status, it would convey a special message—either of a sense of camaraderie or of indignation or rebellion."

In Korean, a respected person would be addressed in many situations using a job title or role term, or a kinship word, often in combination with the honorific suffix *-nim* (Lee and Ramsey 2000: 225–38). For example, *seonsaeng-nim* 'respected teacher', *sojang-nim* 'respected director', *gisa-nim* 'respected engineer, technician', *son-nim* 'respected guest', *ajwumeni* 'Auntie'. Calling someone by their bare personal name in Korea is a highly

Fig. 1.1. Kinship designations and names used instead of pronouns in Malaysian.

HELLO TI!.. ABANG DI SINI. MAAFKAN ABANG KERANA PAGI-PAGI LAGI SUDAH MENGGANG-GU KEPADAMAIANTI

"Hello *Ti!.. Abang* here. Forgive *abang* for once again disturbing *Ti*'s peace and quiet in the morning."

SHAMSUL AZHAR MENGAM-BIL KEPUTUSAN MENGHU-BUNGI SITI DI PEJABATNYA PADA AWAL PAGI.

Shamsul Azhar decides to contact Siti in the office early in the morning.

HAH! ADA APA NI TIBA-TIBA AJE ABANG MENELE-FON TI PADA WAKTU INI? WAL-HAL SEBELUM INI TAK PER-

"Hah! What's up that *abang* suddenly telephones *Ti* at this time?"

NAH WALAU SE-KALIPUN ABANG HU-BUNGI TI KETIKA INI... JANGAN-JANGAN ABANG TERSILAP NOMBOR TAK?

"Especially since before this *abang* never got in touch except once... Hadn't *abang* misplaced the number?"

sensitive matter. To address anyone by name who is more than five or six years older than oneself would be unusual, no matter how close one feels towards them, and even the closest friends start to feel uncomfortable using each other's names once they reach 30 or 40 years of age. Within the family the only people one can address by name are one's children or younger siblings.

A widespread strategy for avoiding pronouns is simply to avoid explicit references to oneself or one's interlocutor altogether. In English this is rather difficult, but it is easier if the grammatical structure of a language allows "incomplete" sentences, i.e. sentences in which the subject or object is not stated explicitly. This phenomenon (known as ellipsis) is extremely common in Japanese, Korean, Sinitic languages, and in mainland Southeast Asian languages. Accordingly, in these languages the solution to the problem of how to refer to someone is often not to refer them at all.

1.9   Honorific forms

Some languages of East and Southeast Asia—the best-known in this respect being Japanese, Korean, and Javanese—have morphological and lexical means of adding "respect" messages for either (a) the person being spoken to, i.e. addressee, or (b) the person or people being spoken about, i.e. the referent. We can illustrate quickly from Japanese: the verbal suffix *-masu*

Box 1.1. Thai nicknames (*chîi lên* 'play name')

Given names in Thai often have two, three, or more syllables and generally express elegant or "high-sounding" meanings; for example, Chalermwan 'celebrated beauty' or Kittichai 'famous victory'. Except in formal contexts and with people one doesn't know well, however, one often goes by a monosyllabic "nickname". It is usually chosen by parents or other family members soon after the baby is born—before the formal name is chosen, in fact—on the basis of some appealing or distinctive characteristic of the baby. Here are some common ones grouped by meaning type. The names are given in approximate English spellings, rather than in the phonemic spelling used elsewhere in this book. Actually there is a lot of variation, so you may see variants on all or most of these spellings, e.g. Nueng for Neung, Noy for Noi, Choy for Joy, Gop for Kop, Khao for Khaw, Moo for Mu.

Birth order and/or size	Animal	Fruit and flowers	Appearance and character	Colour
Neung, Ek (one)	Maew (cat)	Som (orange)	Ke (smart-looking)	Daeng (red)
Song (two)	Pla (fish)	Ple (apple)	Uan (fat)	Leuang (yellow)
Nong (young)	Nok (bird)	Kluay (banana)	Neng (bald)	Faa (blue)
Tow (big)	Mu (pig)	Ngo (rambutan)	Geng (clever)	Kheow (green)
Yai (big)	Kai (chicken)		Gaen (naughty)	Dum (black)
Lek (little)	Chang (elephant)		Jaew (bright-eyed)	Khaw (white)
Noi (little)	Kop (frog)		Tui (chubby)	

Nicknames can also be a shortened form of one's true name, e.g. Bat from Sombat, Pat for Patcharee, Da from Ladda, or a blend of syllables from the parents' names, or come from another language such as English, e.g. Baby, Pepsi, May, A, B, Rose. The well-known Thai tennis player Paradorn Srichaphan comes from a dedicated tennis family who gave him the nickname Ball. During the World Cup soccer, a number of Thai babies also got the name Ball, after the soccer ball.

expresses respect for the addressee, whereas the nominal prefix *o-* expresses respect for the person being spoken about.

Some scholars draw a terminological distinction on this basis. They point out that linguistic devices which code respect for the addressee can be seen as belonging under the broad rubric of "politeness". Linguistic devices for expressing respect for a referent, on the other hand, seem to be rather different in nature, and deserving of a special term. For this reason, some scholars reserve the term "honorific" (or "honorification") for referent-related phenomena. This seems like a good idea—or it would be, except for the fact that other scholars use the terms "honorific" (or "honorification") indifferently for both addressee-related and referent-related phenomena.

In Chapter 7 we will look closely at the Japanese honorific system, but here let us look briefly at Korean. Various Korean nouns have not only their

Table 1.5. Some Korean honorific nouns and verbs

	Honorific noun	Plain noun
'father'	*abeonim*	*abeoji*
'grandmother'	*halmeonim*	*halmeoni*
'food, meal'	*jinji*	*bap*
'name'	*seongham*	*ireum*
'body'	*okche*	*mom*
'word, speech'	*malsseum*	*mal*
	Honorific verb	**Plain verb**
'to eat'	*japsusi-*	*meok-*
'to sleep'	*jumusi-*	*ja-*
'to be, to exist'	*gyesi-*	*iss-*

ordinary forms but also a special honorific form. Honorific nouns express special respect for the person referred to by such a noun (or, in the case of objects, for the person associated with the object). Honorific verbs are formed by means of the honorific suffix *-si/-eusi*. In addition, there are a few plain verbs which bear no obvious relationship to their honorific counterparts. Examples of pairs of plain nouns and honorific nouns are given in Table 1.5 (Lee 1989: 57–8).

Honorific nouns and verbs tend to occur together. Plain nouns which lack honorific counterparts can also occur in sentences with honorific verbs. A couple of examples (Lee 1989: 59, 87; cf. 104):

Korean

(**24**) *Abeonim   i   jumusi-n-da.*
    father(HON) SUBJ sleep:HON-PRES-DECL
    'My father is sleeping.'

(**25**) *Gunin   i   bange   gyesi-eo-yo.*
    soldier SUBJ in.room is:HON-DECL-PCT
    'A soldier is in the room.'

1.10 Other common features

Other linguistic features are widespread among East and Southeast Asian languages, at least in some regions. We will come to these later in the book. They include: (a) sentence-final particles, i.e. invariable little words which express nuances of the speaker's attitudes and feelings, usually appearing at the end of a sentence (see Chapter 4); (b) widespead ellipsis, also known

as zero anaphora, i.e. the phenomenon whereby noun phrases are simply omitted if the identity of the referent can be understood from context; (c) topic prominence, i.e. when the organization of a sentence is heavily influenced by factors such as where the speaker's focus of interest and attention is, what the speaker and listener are taking for granted, and other so-called "discourse" factors (see Chapter 4).

Key technical terms

agglutinating	lexical tone
constituent order	measure word
contour tone	minority language
diacritic	national language
honorific	numeral classifier
inflection	postposition
isolating	register tone
language family	