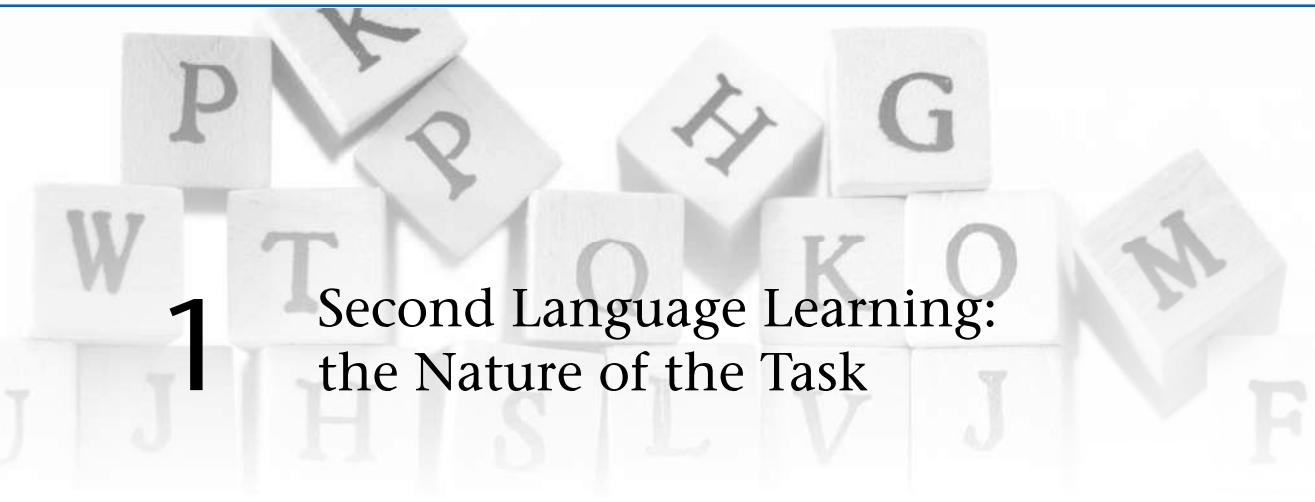

1 Second Language Learning: the Nature of the Task

1.1 Can Anyone Learn a Second (Third, Fourth, …) Language?

Some people are apparently very good at learning languages beyond the one(s) they acquire in infancy. The nineteenth-century British explorer Sir Richard Burton claimed to have spoken more than forty languages and dialects (Farwell, 1963). There are also 'linguistic savants', people who, despite a number of cognitive impairments, have a precocious ability to learn languages. Christopher, a linguistic savant studied by Smith and Tsimpli (1995), has a number of cognitive impairments that make it difficult for him to find his way around or deal with everyday tasks like doing up buttons or shaving. However, 'he can read, write and communicate in any of fifteen to twenty languages' (1995: 1). In contrast to these examples, other people appear to be rather poor at language learning. In 2008, the *Guardian* newspaper reported the attempts of a French teacher at a London school to motivate a Vietnamese boy in his bottom French class, Tommy, for an upcoming French exam. In doing so he was presented with a challenge:

> I told him it wasn't too late. I insisted he could still pass French even at this stage. Tommy scoffed. You learn Vietnamese in three months, he said, and I'll learn French. (*Education Guardian*, Tuesday 26 August 2008, p. 4)

The 52-year-old teacher accepted the challenge, and passed an Institute of Linguists first-level certificate in Vietnamese three months later. Unfortunately, Tommy was not so successful. He failed his French exam with an even lower grade than he was originally predicted to get.

Examples like these sometimes lead to the belief that a special talent is needed to learn languages beyond one's native language. This is not the case. For the majority of people successful second language learning is entirely realistic. Where second, third or further languages need to be acquired for an individual to function normally in society, they mostly can be acquired.

Stenzel (2005) reports the case of the Wanano, a people living along the Vaupés river in South America on the border between Colombia and Brazil. It is the social norm among the Wanano to marry someone outside their language group,

because speakers of the same language are regarded as brothers and sisters. As in most societies, to marry a brother or sister would be incest. It is also the norm for the wife to move to the husband's village and to learn the husband's language. The children of such marriages grow up speaking both languages. Because other wives moving to the same village may speak different languages, other languages may also be learned. In addition, there is a local lingua franca, Tukano, used for communicating with other indigenous communities, and Portuguese and Spanish for communicating with speakers from elsewhere in Brazil and Colombia. A language census conducted in 2002 with thirty-eight Wanano males aged between 12 and 65 shows that individuals speak on average six languages (Stenzel 2005: 20–1).

This is not an isolated example. Similar multilingualism can be found in other parts of the world. Erard (2012: 199) reports the following description of family language use by a woman living in southern India:

> We speak Hindi a lot … All of us speak Telegu, but when more of us are together, we speak Tamil. Then, when we get even more together, we speak Kannada.

It appears, then, that many people can achieve good levels of communicative ability in a second language (henceforth referred to as L2) given the motivation to do so, the opportunity to interact with samples of the target language, and sufficient time for such encounters to turn into acquired mental linguistic knowledge, as the Wanano people, those living in southern India and Tommy's teacher show. The willingness to interact with samples of the target language, and to do so over long periods of time, appear to be key factors. It is interesting to note that Tommy, a Vietnamese speaker, had already been highly successful in acquiring English, where presumably motivation was high and access to samples of English easy for him. Indeed, his success in using English could well have been one cause of his lack of success in French. As his teacher observes, 'it did not help that he preferred to spend the [French] lesson chatting', presumably in English.

Furthermore, Farwell's description of how Burton set about learning a new language suggests high levels of motivation and engagement with the language over a sustained period of time:

> First he bought a simple grammar and vocabulary and underlined the words and rules he felt should be remembered. Putting these books in his pocket, he studied them at every spare moment during the day, never working more than 15 minutes at a time. By this method he was able to learn 300 words a week. When he acquired a basic vocabulary, he chose a simple story book and read it, marking with a pencil any new words he wanted to remember and going over these at least once a day. Then he went on to a more difficult book, at the same time learning the finer points of the grammar … When native teachers were available, he claimed that he always learned the 'swear words' first and laughingly said that after that the rest of the language was easy. (Farwell, 1963: 30)

Modern examples of 'hyperpolyglots' (obsessive acquirers of L2s) appear to confirm that time spent in the company of the L2 is a major factor in its acquisition. Erard (2012: 117) describes one such hyperpolyglot, Alexander Arguelles, who spends around ten hours a day working to maintain the languages he speaks, devoting most time to revising the languages he knows best (English, Arabic, French, German, Latin, Chinese, Spanish, Russian and Korean) but also devoting some time to forty other languages.

This book assumes that, for many people, acquiring L2s is a natural extension of human linguistic ability. Given the right opportunities and sufficient time, good levels of communicative ability in an L2 can be achieved. It aims to show how people do this. You will discover that there are remarkable similarities in the way people who speak different first languages and are acquiring different L2s go about the task, although there are also clear influences from their first language. The accumulation over the past fifty to sixty years of observations from carefully-designed empirical studies has led to the proposal of various hypotheses and theories about L2 learning. Several of these will be outlined and evaluated against the available evidence. Since many people learn L2s in a classroom setting, the relationship between the samples of language they encounter (the 'input') and what they learn will also be discussed, as will the often-asked question: 'What is the best age to learn an L2?' Two further topics that will be addressed are how speakers of L2s use their knowledge in real time, and the role that context plays in L2 learning.

The aim in this chapter is to outline the task facing the L2 learner. This will be presented as a number of 'problems' that the learner has to solve. The discussion starts with how learners 'break into' the L2 in the first place; that is, how they identify its words and assign them the right meanings. A second problem is distinguishing appropriate from inappropriate combinations of words in phrases, clauses and sentences. When words combine to form phrases, those phrases sometimes have meanings that go beyond those of the individual words. Identifying such cases constitutes a third problem. A fourth problem is identifying the way that sentences link to each other in verbal exchanges and written texts. Finally, an L2 learner needs to work out the variety of language to use in specific situations.

1.2 Cracking the Code: the Segmentation Problem

At birth, children with normal hearing ability are confronted with continuous strings of sounds produced by the people around them. To acquire the language(s) that those people are speaking the child has first to identify recurring portions of the sound continuum that are potentially meaningful. S/he then has to associate those chunks of sound with actual meanings. Further, the child has to discover how those sound-meaning chunks – the words of the language – can be combined and recombined to create the multiplicity of new messages s/he will want to understand or will want to communicate to others every day. This is a formidable challenge.

To illustrate, imagine you are an infant living in a community that speaks West Greenlandic (Fortescue, 1984). You will hear people uttering strings of sounds like the following:

1. qimmipinnittuquppai

Even if you can work out from the non-linguistic cues available in the context that it means 'The dog killed the people', you do not know which section of the string might correspond to 'dog', which to 'people' and which to 'kill'. Given that in some languages the verb ('kill') can come at the beginning of a sentence (like Welsh), perhaps some portion of *qimmipin* … means 'kill'? On the other hand, in languages like Japanese and Turkish the verb comes at the end of a sentence, so perhaps some portion of … *tuquppai* means 'kill'. And which bits of sound refer to 'dog' and which bits to 'people'?

Exactly the same 'segmentation problem' faces the L2 learner. Classroom L2 learners are usually helped by having their attention consciously drawn to sound–meaning correspondences and the order in which constituents occur (often with written support). But even if you know consciously that *qimmip* is 'dog', *innit* is 'people' and *tuquppai* is 'kill', such explicit knowledge is difficult to access when trying to 'catch' what native speakers are saying, or produce utterances yourself (as anyone who has learned an L2 on the basis of a 'read-then-listen/speak' method will know). What you need is an ability to automatically segment the fast stream of speech you hear into its meaningful components.

It appears that humans have such an unconscious analytical capacity that is available to infants and older L2 learners alike. This capacity is the ability to analyse rapidly a continuous stream of speech into its component sounds, like the [q], [i], [m] and so on from example (1), and to retain a memory of the frequency with which they co-occur. The unconscious tallying of the frequency with which individual sounds occur together is important because sounds co-occur more frequently within words than between words. To illustrate, the probability of hearing the sequence [kæt] 'cat' as a recurring string in sequences like *the**cat**likesmilk* is greater than the probability of hearing other sequences of adjacent sounds that cross word boundaries: [ðək] 'thec', [ætl] 'atl' or [tlai] 'tli'. If language learners retain a memory for such co-occurrence frequencies, it provides them with an important cue to the identification of words from a continuous sequence of sounds: an important step in solving the segmentation problem.

Research has demonstrated that adults not only retain a memory of the frequency of co-occurrence of sounds from novel strings, but also do so very rapidly. Saffran et al., (1996b) invented a language that consisted of six 3-syllable 'words' like *bidaku, padoti, tupiro* and *golabu*. These words were combined and repeated to form a continuous string as in (2) that was recorded in a female voice.

2. bidakupadotigolabupadotibidakugolabupadoti …

English-speaking participants in their study (undergraduate student volunteers) were told that the language contained 'words', but were not told what they were.

After 21 minutes of listening to the speech stream, participants were presented with a word decision task. They heard pairs of syllable strings where one was a 'word' from the continuous stream they had been listening to (like *bidaku*), and the other either a 'non-word' made up from syllables that never followed each other in the listening phase (like *dadopi*) or a 'part-word' created from two syllables of a 'word' plus a random syllable (such as *bidado*). Participants were asked to decide which of the two strings 'sounded more like a word from the language' (Saffran et al., 1996b: 613) by pressing a key on a computer keyboard. The results show that participants were significantly above chance in identifying both 'words' and 'part words', although they were less certain about the latter (unsurprisingly, since they only partially resembled 'words' in the language they had been listening to).

Saffran et al. interpret these results as showing that adult language learners are sensitive to the probability with which one syllable will follow another in the speech stream. In the stream of sounds they heard, only *da* could follow *bi*, and only *ku* could follow *da* (because *da* is the central syllable in *bidaku*). By contrast, *bi* could be preceded by any of *ti, ro, bu*, and *ku* could be followed by any of *pa, tu, go*. The adult participants unconsciously computed the 'contingencies' between these syllables very rapidly to identify recurring, stable sound sequences in the input they encountered.

In a similar study with 8-month-old infants using the duration of fixation of the infants' gaze on a blinking light to determine whether they were distinguishing words from non-words (Saffran et al., 1996a) the same results were found.[1] So, at a level of perception below conscious awareness, adult L2 learners have the same capacity as infants to compute and retain a memory for recurrent chunks of sound in the stream of speech they are exposed to. The ability to identify such chunks that are potentially meaning-bearing forms in a target second language is an important step in solving the 'segmentation problem'.

Section 1.2 in a Nutshell

Adults have an unconscious mental capacity for computing the frequency with which sounds co-occur in the speech stream that contributes to their identification of word-like sequences. They share this capacity with infants. The ability to identify word-like sequences that are potentially associated with stable meanings is an important step in acquiring a new language.

[1] In the blinking-light procedure an infant sits on a parent/carer's lap and looks at a blinking light on a facing wall while listening to a continuous sequence of invented words, as in (2). At the end of two minutes the light is switched off and another starts blinking on a side wall. This new light is accompanied either by one of the invented words the infant has just been listening to, or a non-word. The average time the infant spends looking at this second blinking light while hearing words is compared with the time spent looking while hearing non-words. Saffran et al. found that the infants studied spent significantly longer looking when they heard non-words than when they heard 'words', suggesting that they were distinguishing the two.

1.3 Matching Form and Meaning: the Categorisation Problem

In real languages, recurrent strings of sounds (or sequences of written symbols) can be associated with three types of meaning, and working out which of these form–meaning relations is appropriate is a second problem to solve.

Some strings are associated with meanings that uniquely identify people, things, ideas, actions, states, locations and qualities in the world (real or imagined). For example: *student, book, freedom, walk, know, in, smooth*, etc., are all recurrent strings in English with uniquely identifiable meanings. Other strings (including those that consist of just one sound segment) have meanings that modify strings with uniquely identifiable meanings. This is the case, for example, with the English articles *the* and *a*, whose function is to indicate whether the uniquely identifiable meanings of nouns (like *student, book*) are to be understood as definite or indefinite: *I saw the student* (a definite description of a particular person – I expect you, the reader, to know who I am talking about); *I saw a student* (an indefinite reference to a person who is a member of the class 'student').

Forms that have a meaning-modifying function need not be words in their own right (although *the* and *a* are). They can attach to strings with uniquely identifiable meanings as affixes. For example, in the strings *students, books, freedoms*, the final *-s* is an affix that turns the reference of *student, book* and *freedom* from singular to plural.

Yet other strings of sound mark a dependency between forms. For example, the *-s* affix of the English string *sings* marks a dependency between *sing* and the constituent that typically precedes it (the **subject** of *sing*). This constituent must be third **person** (i.e. not *I* or *you*) and singular (i.e. not *they*): *He sings* and *She sings* are fine, but ***I sings* and ***They sings* are not fine in standard varieties of English. (Note that this *-s* has the same form as the meaning-modifying plural *-s*, as in *songs*, but the two have different functions.) Dependency-marking strings are often affixes.

The L2 learner must associate recurrent strings of sounds/written symbols that have been identified with one of the categories of meaning: uniquely identifiable, meaning-modifying or dependency-marking. That is, they have to solve a 'categorization problem'.

Where samples of the L2 that are encountered are written, it appears that L2 learners initially deal with the categorisation problem primarily by retaining a memory for form–meaning correspondences where uniquely identifiable meanings are involved. DeKeyser (1995) has shown that in the initial stages of learning a new language, the extent to which adults can identify meaning-modifying and dependency-marking forms simply from exposure to samples of the target language is quite limited.

Like Saffran et al. (1996a, 1996b) DeKeyser used the 'invented language' technique, but unlike the language used by them, his language, called Implexan, had properties that were more like those found in real languages. There were ninety-eight words with uniquely identifiable meanings, for example the nouns

melaks 'apple', *perakt* 'book', and the verbs *wulas* 'peel', *wost* 'read'. In addition, there was a meaning-modifying affix that could attach to nouns and that changed their reference from singular to plural (like English *-s*): *melakson* 'apples', *perakton* 'books'. There were also dependency-marking affixes, one that necessarily attached to verbs when the subject of the verb was female:

3 a. Mari wostin. 'Mary is reading'
b. Rober wost. 'Robert is reading'

Another attached to nouns when they followed the verb (i.e. when they were the verb's direct object):

4 a. Mari wostin peraktus. 'Mary is reading a book'
b. Rober wost peraktus. 'Robert is reading a book'[2]

DeKeyser exposed a group of thirteen adults to this language over twenty training sessions of 25 minutes each. In each session participants saw the same 124 photographs depicting simple actions paired with a written sentence of Implexan (for example, a photo of a woman reading a book paired with the sentence *Mari wostin peraktus*), although the order in which the photographs were presented varied in each session. (Note that this is a written 'language', unlike the spoken 'language' in the Saffran et al. studies. Further, the segmentation problem is greatly reduced because the words of the language are already identified by spaces between letter strings). To stimulate the active processing of meaning in Implexan, participants were presented at regular intervals with additional pictures, some of which were accompanied by sentences whose nouns or verbs did not match the picture. Participants were asked to indicate whether the sentence matched the photo, yes or no. (The affixes were always appropriate during the training phase).

At the end of the training period, participants were asked to produce sentences in Implexan themselves to describe photos. Eight of these could be described using sentences that had been encountered during the training phase, but thirty-six were new photos requiring new combinations of nouns, verbs and affixes. Results showed that participants were very good at using previously encountered sentences to describe previously encountered photos (89.5% accurate). They were also very good at using previously encountered [noun + affix] and [verb + affix] strings to describe new photos (87.3% accuracy). But when they had to use new [noun + affix] or [verb + affix] combinations to describe new photos, their accuracy dropped to 33.3%. DeKeyser observes that 'a score of 33.3% is about what one would expect by chance' (1995: 393). They either failed to add affixes to nouns and verbs where they were required, or overused an affix where it was not required. In the latter case, all the errors involved the overuse of the plural affix *-on* in contexts where the referent of the noun was singular. This is strongly suggestive of a failure by the participants to identify the function of *-on* as a plural marker.

[2] This description of Implexan is simplified, as is the description of DeKeyser's experiment. For full details see DeKeyser (1995).

These results suggest that while the participants had retained a memory for recurrent strings of co-occurring letters (rather than sounds as in the Saffran et al. study) with uniquely identifiable associated meanings, they had not identified the meaning-modifying and dependency-marking affixes within those letter strings. DeKeyser concludes that the participants' unconscious knowledge of the rules underlying the distribution of the affixal forms 'was … virtually non-existent' (1995: 393).

1.4 Identifying Possible and Impossible Word Combinations: the Syntax Problem

Another important feature of 'real' languages that an L2 learner has to acquire is the range of ways in which meaning-bearing strings of sounds (or letters) can be combined, and the kinds of combinations that are not possible. In simple **declarative** (statement) sentences in English, verbs appear after their subjects and before their objects: *Today Mary*$_{[subject]}$ *plays hockey*$_{[object]}$. In West Greenlandic the verb appears at the end of the sentence: *Today Mary hockey plays*. In German, the verb marked for tense and agreement with the subject appears in second position in main clauses (*Today plays Mary hockey*), but at the end of a clause that is subordinate to the main clause (e.g. *Molly said that today Mary hockey plays*). If the subordinate clause comes first in the sentence, its verb appears at the end of that clause, and the tensed/agreeing verb in the main clause immediately follows (because it is in second position after the initial subordinate clause): *Whenever Mary hockey plays, wears she a mouth guard*. The rules that determine the appropriate combinations of words in a language (and exclude inappropriate combinations) are known as its **syntax**.

Rebuschat (2008) (reported in Williams 2009: 331) used a quasi-invented language to examine how quickly adult learners might identify possible and impossible word combinations in a target language. His quasi-invented language consisted of English word-forms with German word order. He presented his English-speaking adult participants (with no previous knowledge of German) with 120 sentences, where the three word orders described above were distributed equally. Their 'training' in the language involved making judgements of the plausibility of the meaning of each sentence (i.e. there was no explicit reference to word order in participants' encounters with the language). After they had made judgements on the 120 sentences, they were given a 'surprise' test in judging whether the word order in a new set of sentences was **grammatical** or ungrammatical (in the invented language they had just encountered). Results show that participants were above chance in accepting patterns of word order combination found in the training sentences, even though the vocabulary was different, but responded randomly to word orders that were ungrammatical. Williams concludes that 'there was rapid incidental learning of abstract word order patterns, but … no learning

of the actual verb placement rules' (2009: 331). In other words, L2 learners in the initial stages can compute and remember linear orders of words in the samples of language they encounter, but their failure to use that information to identify possible and impossible word order patterns suggests that it may be more difficult for them to solve the 'syntax problem' at this very early stage of exposure, the problem of determining the syntactic rules that give rise to particular word orders and not others.

Sections 1.3 and 1.4 in a Nutshell

When L2 learners have identified recurrent strings of sounds (or letters) in the stream of speech (or in connected text), they need to establish what kinds of meaning they have. Those meanings fall into one of three categories: (i) uniquely identifiable; (ii) meaning-modifying; (iii) dependency-marking. In the second and third cases, the forms involved are often affixal. The establishment of form–meaning associations was referred to as the 'categorisation problem'. A study by DeKeyser (1995) of the initial learning of an artificial language (with the characteristics of a 'natural' language) suggests that learners have considerably more difficulty categorising strings that have a meaning-modifying or dependency-marking function (and are also affixes) than categorising strings that have a uniquely identifiable meaning. Results from a study by Rebuschat (2008) suggest that when L2 learners initially encounter novel syntax in an L2 they can apply it to new vocabulary, but they do not immediately infer impossible from possible word orders.

1.5 Working Out the Meaning of Word Combinations: the Semantics Problem

When the syntactic rules of a language combine words into phrases, composite meanings are created. Unremarkably, combining the English words *a, red* and *ball* into *a red ball* creates a phrase that refers to an object that is both a ball and red. Sometimes, however, the meaning of phrases goes beyond the meanings of the individual words of which it is composed. *A heavy smoker* could indeed be a person who is both a smoker and heavy, but more usually it has the meaning 'someone who smokes heavily'. The ambiguity arises not just from the meanings of the words themselves, but from their meanings in conjunction with the structure of the phrase into which they have been combined. Speakers of English know that *heavy* can be understood as modifying the meaning of *smoker* in two ways. The first is additive: it adds the meaning 'heavy' to an entity in the world described as a 'smoker', in the same way that *red* adds the meaning 'red' to an entity in the world described as a 'ball'. The second is qualitative: *heavy* qualifies the activity

engaged in by the entity referred to; that is, 'smoking'. The entity is someone who 'smokes heavily'.

Determining the meanings of word combinations created by the syntactic rules of a language is the job of **semantics**. One of the tasks facing the L2 learner is identifying how composite meanings are derived from the individual meanings of words that have been combined into phrases and sentences. As in the case of the 'categorisation problem', solving the 'semantics problem' may take some time. An illustration of the development of such knowledge is provided in a study by Anderson (2008) of the identification of the additive and qualitative meanings of adjective–noun combinations in French by L2 learners whose L1 is English.

Unlike English, the majority of adjectives in French follow the noun they qualify, for example *un ballon rouge* (literally 'a ball red'), *un voyage fatigant* (lit. 'a journey tiring'). A handful of adjectives typically precede the noun, for example *un vieux château* ('an old castle'), *un bon professeur* ('a good teacher'). There is, however, a set of adjectives that can both precede and follow a noun, but change their meaning with their position. Some examples are given in (5) (from Anderson 2008: 13). In these cases the meaning is qualitative when the adjective precedes and additive when the adjective follows.

5	a.	un **cher** bijou 'a cherished jewel'	un bijou **cher** 'an expensive jewel'
	b.	un **ancien** roi 'a former king'	un roi **ancien** 'an ancient king'
	c.	un **pauvre** village 'a pitiful village'	un village **pauvre** 'a poor village'
	d.	une **grande** gymnaste 'a great gymnast'	une gymnaste **grande** 'a tall gymnast'

Anderson investigated whether English learners of French at different proficiency levels were sensitive to the difference in meanings associated with the different position of eight adjectives in phrases like these. He elicited their intuitions by asking them first to read stories that favoured either the qualitative interpretation or the additive interpretation, and then to rate sentences containing adjective + noun or noun + adjective as 'fine' or 'odd'. For example, one of the stories was about an old lady distributing some of her possessions to her relatives before she dies, but who keeps an imitation diamond that cost little but is of great sentimental value to her. This favours the qualitative 'cherished jewel' interpretation of the word order *cher bijou*. If participants have understood the relationship between the order of adjective and noun and meaning, they should find a sentence containing the phrase *son cher bijou* 'her cherished jewel' fine and a sentence containing the phrase *son bijou cher* 'her expensive jewel' odd. Participants were mainly classroom learners of French in their second, third and fourth year of study at an American university, or university graduates (the advanced group). There was also a **control group** of French native speakers. The results are presented in table 1.1.

The important information to look at in these results is not the absolute ratings of the test items as 'fine' (i.e. whether participants rate the expected cases as 100%