4

The Critical Period Hypothesis

4.1 The view from biology 83
4.2 The two positions in second language acquisition 85
4.3 Global nativelikeness versus different sensitive periods for the separate parts of the grammar 86
4.4 The first and the second language as communicating vessels 90
4.5 Effects of bilingualism: Is the bilingual two monolinguals in one mind? 92
4.6 The importance of the input 95
4.7 An indirect way of appreciating the importance of input 99
4.8 Conclusions 102
4.9 Exercises 104

4.1 The view from biology

If language development is like other neurological development processes, it is likely to have a critical period early in life, after which no amount of experience can bring the development to normal ranges. Hubel and Wiesel (1970) demonstrated such a critical period in the ocular development of kittens. By depriving them from using one of their eyes, they showed that the kittens did not develop areas of the brain receiving input from both eyes, something that is needed for binocular vision. The forced ocular dominance in the early months of life turned out to be irreversible.

With respect to language, the Critical Period Hypothesis was first discussed by Montreal neurologist Wilder Penfield (e.g., Penfield and Roberts 1959). Later, Eric Lenneberg's (1967) seminal work popularized the idea and

especially its implications for language as a biological phenomenon. In essence, Lenneberg argued that there are maturational constraints on first language acquisition due to brain lateralization, which he assumed started at 2 and finished at puberty.[1] If language is not acquired by puberty, some aspects of language such as the lexicon can be learned, but nativelike mastery of grammatical structure cannot be achieved. Some empirical support for this observation comes from the cases of Genie[2] and other children deprived of language input while growing up.

During most of the 20th century, the consensus among neuroscientists was that brain structure is relatively immutable after a critical period during early childhood. This belief has now been challenged by findings revealing that many aspects of the brain remain plastic even into adulthood. Many researchers nowadays argue that language development may have a sensitive period, after which some impairment normally occurs, but problems can be compensated for and reversed (e.g., Hensch 2004). New knowledge about the development and functioning of the brain highlights its plasticity instead. Decades of research have now shown that, in response to experience, substantial changes can occur in the neocortical processing areas[3] that can profoundly alter the pattern of neuron activation. Neuroplasticity is exhibited on a variety of levels, ranging from cellular changes caused by learning, to larger-scale changes, and up to cortical remapping in response to injury.

Of course, a lot of research remains to be done, but neurocognitive studies already indicate that experience can actually change both the brain's physical structure (its anatomy) and functional organization. To cite just one of

[1] This is no longer believed to be the case. Brain lateralization is thought to be complete by early infancy and maybe even at birth (Hahn 1987).

[2] Genie was 13 years old when discovered. She had been kept strapped to a potty chair and was wearing diapers. Her father had judged her retarded at birth and had chosen to isolate her, and so she had remained until her discovery. She appeared to be entirely without language. Although this case of domestic abuse was a tragic event, it afforded scientists a glimpse at "a forbidden experiment:" what happens if a child experiences total lack of language until the age of 12. Although Genie was exposed to ample language input and taught language after that age, she was unable to acquire English completely (Curtis 1977). The degree of her neurological damage (due to the abuse) and the degree to which she acquired language are disputed.

[3] The neocortex is the roof of the cerebral cortex that forms the part of the mammalian brain that has evolved most recently. It makes higher brain functions possible, such as learning.

many examples, a study by Polk and Farah (1998) found that Canadian postal workers were more adept at manipulating the letter and number representational systems, relative to American postal workers. The explanation was that Canadian postal codes include both numbers and letters while US postal codes include numbers only. Neuroscientists are currently engaged in a reconciliation of critical period studies demonstrating the immutability of the brain after development, with the more recent research showing how the brain can, and does, change.

4.2 The two positions in second language acquisition

It is of paramount importance to keep in mind that Lenneberg's formulation of the critical period for language is not a valid description of the second language acquisition process (Lenneberg 1967: 176). L2 speakers already have a native language that was properly engaged at birth, so the brain lateralization argument does not apply to them. While no researcher is disputing the significance of linguistic input for the normal development of child language acquisition, there is still much dispute in this respect within the field of L2 acquisition.

The two positions have recently solidified without getting much closer to each other. On the one hand, Long (2005), DeKeyser and Larson-Hall (2005), and Hyltenstam and Abrahamsson (2003) argue that it is impossible for adult near-native speakers to attain nativelike grammars, when all facets of the grammar such as pronunciation, perception, syntactic representation, and processing are investigated. In these researchers' opinions, the incidence of nativelike performance among early-onset bilinguals is also less common than it was previously assumed.

On the other hand, Birdsong (2005), Donaldson (2011), Flege (2009), Montrul (2009), Muñoz and Singleton (2011), Rothman (2008), Singleton (2005), and Slabakova (2006, 2008) argue that a much more nuanced approach to nativelike attainment is warranted, where quality and quantity of linguistic input as well as language proficiency play bigger roles than previously assumed, and where sensitive periods for some but not for other modules of the grammar, and even for specific grammatical properties, can be uncovered. The critical period is not absolute.

However far apart the two positions may appear to be, they are not irreconcilable. For example, critical period proponents have recently started

calling it a sensitive period (see the new edited volume Granena and Long 2013). It is possible to piece together a composite picture of ultimate attainment in a second language. We will attempt to see what such a picture might look like, after we discuss some more variables involved in the process of second language acquisition. But first, let's turn to some of the evidence for and against a critical period for language.

4.3 Global nativelikeness versus different sensitive periods for the separate parts of the grammar

Three large-scale studies addressing the Critical Period Hypothesis (CPH) will be presented first, those of Johnson and Newport (1989), Abrahamsson and Hyltenstam (2009), and Abrahamsson (2012). These studies offer considerable improvements over previous work addressing the CPH in terms of number of subjects, variety of tests, and statistical treatment of the data. They offer substantial support for the view that complete, in all respects nativelike linguistic performance is impossible for adult L2 learners to achieve.

The classic study in this respect is Johnson and Newport (1989). In this study, the participant's intuitive knowledge of English grammar was measured through a grammaticality judgment task involving listening to and evaluating sentences. Scores on this task obtained from 46 Korean and Chinese long-term residents in the United States were correlated with the participants' age of arrival (AoA) in the country. There was a strong negative correlation between AoA and grammatical knowledge among the early starters (AoA 3–15; $r = -.87$, $p < .01$), and little individual variation was observed. Very early starters (AoA 3–7) invariably performed like native control participants. However, after the AoA of 15, the significant negative correlation with the grammaticality judgment scores disappeared ($r = -.16$, $p > .05$) and was instead replaced by great individual variation. After that age, individual characteristics rather than age determined the grammatical scores. Despite criticism and some indisputable methodological problems, Johnson and Newport's study remains a very influential milestone in this field of research.

Abrahamsson and Hyltenstam (2009) reports on a large-scale, labor-intensive study of the competence of L2 speakers of Swedish, perceived to be near-native by native speakers. The researchers started out with 195 native speakers of Spanish who were very advanced learners of Swedish

and were living in Sweden. A panel of ten native judges evaluated the nativelikeness of those participants' speech samples, interspersed with some native speaker speech samples. The findings demonstrated that only a small minority (5%) of those bilinguals who had started their L2 acquisition after age 12 were perceived as native speakers. However, a majority (62%) of those with an age of acquisition below 12 were deemed to be nativelike by at least nine of the ten judges. A second experiment involved 41 participants that were considered nativelike by a majority (>6) of the native judges; these were divided into a group of child learners and a group of adolescent to adult learners. The experiment comprised ten tests covering a range of phenomena from phonetic perception and production, to grammar and inferencing, to formulaic language. An important advantage of this experimental design was that it established native-speaker ranges for all tests, which allowed the researchers to then check how many of the L2 speakers fell within those ranges. Only three of the child learners and none of the adolescent or adult learners performed as native speakers on the whole array of ten linguistic tests. The authors concluded that complete nativelike acquisition after puberty is impossible.

The findings of Abrahamsson's (2012) study are largely in line with the Johnson and Newport (1989) study. Again, a large number of participants distributed in roughly equal numbers across AoA points ranging from 1 to 30 were tested on a grammaticality judgment task and on a Voice Onset Time (VOT) discrimination task. The latter task measures whether a perceiver distinguishes between a voiceless stop (*p, t, k*) and a voiced stop (*b, d, g*) in Swedish. VOT is the length of time that passes between the release of a stop consonant and the onset of voicing, with the languages involved, Swedish and Spanish, differing in this respect. Abrahamsson found a robust negative correlation between earlier acquirers (AoA < 15) and scores on the two tasks, but a lot of variation among later acquirers.

In sum, these and other studies show that, although sounding nativelike in everyday conversation, when their phonological, perceptual, grammatical, and lexical abilities are examined in detail, people who have acquired their L2 after puberty (12–15) perform significantly below the native-speaker range, and their global L2 performance is distinct from that of native speakers. Researchers who document such performance very often link it to the distinction of implicit versus explicit linguistic knowledge. This link goes back to Lenneberg (1967: 176), who stressed that what disappears around puberty is the ability to attain "automatic acquisition from mere

exposure," the way children learn their native language. After puberty, language acquisition typically involves "a conscious and labored effort." The implication is that, successful or not, language acquisition achieved through explicit learning is irrelevant to the Critical Period Hypothesis. Keep this distinction in mind as we will revisit it below.

On the other side of the debate, many researchers have argued on theoretical grounds, and continue to demonstrate, that nativelike attainment is possible for some learners with respect to some modules of the grammar, as attested by nativelike judgments or performance on individual properties within these modules. There are many studies documenting such successful acquisition in various parts of the grammar, so I have to be fairly selective here. One study representative of this line of research is Donaldson (2011). The researcher used comprehension tasks that examine intuitions and preferences, but he also recorded the production of ten near-native speakers in natural conversation with native speakers. The author was interested in tracking how the near-native learners used *Left Dislocation*, a syntactic construction that marks Topic, as in (1).

(1) Marie$_i$, elle$_i$ vient cet après-midi.
Marie, she is coming this afternoon
'Mary is coming this afternoon.'

In order to use this construction with nativelike frequency and accuracy, L2 speakers of French have to coordinate syntactic knowledge with monitoring who is already mentioned in the discourse. It is precisely this coordination of two types of information (grammar and discourse) that has been argued to present extraordinary difficulty to learners (Sorace, 1993, 2011, Sorace and Filiaci 2006).

Analyzing the near-native speakers' production of left dislocations—the syntactic accuracy of the dislocated subjects and objects, the obligatory resumptive pronoun[4] in the main clause, and how speakers used the construction to promote different types of discourse referents to topic status—Donaldson contends that the near-native participants' mastery of this aspect of discourse organization converges on that of native speakers. Donaldson's near-native speakers demonstrated nativelike performance on a property

[4] A resumptive pronoun is a pronoun that refers to the moved argument, in this case the subject in example (1).

that could not have been transferred from the native language and is rarely explicitly taught in language classrooms.

An earlier study, Montrul and Slabakova (2003) set out to test whether very advanced learners of Spanish with English as their native language had nativelike intuitions on the meaning of the aspectual tenses (Preterit and Imperfect past tenses). These interpretations are deemed to be particularly difficult for learners of Spanish. The test instruments targeted both semantic knowledge that is taught in language classrooms (e.g., the Preterit denotes a complete event in the past; the Imperfect conveys a habitual sequence of events in the past) as well as semantic entailments that are not taught in language classrooms and are difficult, even impossible, to deduce from positive input alone (see Exercise 4.4 for some actual test items).

To take a concrete example, one such subtle meaning distinction is the interpretation of the impersonal subject *se* in the context of Imperfect and Preterit aspectual tenses. In sentences with Imperfect verbs as in (2) *se* refers generically to everyone and is equivalent to the impersonal pronoun 'one.' In sentences with Preterit main verbs as in (3) *se* refers to a specific group of people already mentioned and is equivalent to 'we.'

(2) Se comía-IMPF bien en ese restaurante
'One ate well in this restaurant.'

(3) Se comió-PRET bien en ese restaurante
'We ate well in this restaurant.'

As the reader can ascertain for herself or himself, instruction of such meaning subtleties is quite unlikely to be offered in many language classrooms. In general, such semantic differences are not often discussed explicitly among speakers of a language (unless they happen to be linguists). Therefore the authors argued that this particular semantic contrast represents a Poverty of the Stimulus learning situation.

It was established that 15 out of 17 near-native participants, 5 out of 23 superior participants, and even 2 out of 24 advanced speakers (so 22 out of a total pool of 64 participants) performed within the range of native speakers on a truth value judgment task probing these interpretations. These findings, as well as Donaldson's findings mentioned earlier, address the implicit versus explicit learning issue introduced above. If we assume that L2 learners who attempt L2 acquisition after puberty are only capable of labored,

conscious acquisition that depends on rote learning of rules, Montrul and Slabakova's (2003) and Donaldson's (2011) results remain unexplained. No speaker can reach such advanced state of knowledge through learning explicit rules. Slabakova (2006) and Rothman (2008) make the same argument and give many additional examples of successful acquisition of properties exhibiting poverty of the stimulus. They argue that such results are in direct contradiction to a strict critical period claim.

In summary, this section presented some evidence for the two positions on the critical period for language. Of course, we cannot do justice here to the literature that has been growing over the last 40 years and comprises over a hundred published studies. But I hope to have demonstrated that the issue is no longer "Is there or isn't there a critical or a sensitive period for language acquisition?" The answer is more complicated. One conclusion that we can all agree on is that if age of acquisition is after puberty, there are very few individuals in a group, and frequently none, who display a complete mastery of the second language, including phonetics and phonology. However, looking at specific properties of morphosyntax and meaning acquisition, research frequently uncovers successful adult learners. The last 30 years have revealed considerable complexity in the overall picture of L2 acquisition, and it appears that other factors are at least as important, if not more important, than age of acquisition. We will review more studies in the next sections, where we discuss other factors that can and do affect nativelike acquisition.

4.4 The first and the second language as communicating vessels

The notion of first and second language interference, or interaction, has gathered some credibility in the discussion of critical or sensitive periods. In a nutshell, the idea is that the more the native language grammar is established (entrenched) in the learner's mind/brain, the more difficult comparable acquisition of a second language becomes. A useful visualization of this idea is to imagine the first and the second languages of the bilinguals influencing and interacting with each other as liquid in communicating vessels (Bylund, Hyltenstam, and Abrahamsson 2013).[5] In his Speech Learning

[5] See animation at http://en.wikipedia.org/wiki/File:ANIMvasicomunicanti.gif.

model (Flege 1995), James Flege conceptualizes the native language as one factor that can interfere with nativelike mastery in the second language. The model assumes that the sounds of the L1 and L2 are related to one another and exist in a common "L1±L2" phonological space. Being in contact, the sound systems of the L1 and L2 may mutually influence each other. The key word here is "mutually."

A large-scale study, Yeni-Komshian, Flege, and Liu (2000), assessed the global pronunciation of 240 L1 Korean–L2 English bilinguals divided in ten groups according to age of arrival in the US or Canada: AoA 1–3, AoA 4–5, AoA 6–7, etc. They were recorded as adults, reading and repeating sentences in Korean and English. Ten monolingual judges evaluated the accentedness of the bilinguals' speech in both languages, Korean and English, on a scale of 1 (very strong accent) to 9 (no accent). The bilinguals who had emigrated at a very young age (AoA 1–5 years) had relatively high pronunciation ratings in English, but as a group they were still distinguishable from monolingual speakers. In this respect, the Critical Period Hypothesis was supported in the area of pronunciation. However, another important finding was the significant negative correlation ($r = -.47$)[6] between L1 and L2 pronunciation for bilinguals who had started acquiring English before the age of 12, but not for the those participants who arrived at higher ages. Interestingly, one group, who arrived when they were 10 or 11 years old, had roughly equal ratings in English and in Korean. The authors interpreted their findings as partially supporting the interference hypothesis: the more strongly the native language is represented in the brain, the more strongly it will interfere with the second language. The communicating vessels metaphor works for this and other studies with similar findings: you just have to imagine the communicating vessels tilted to one side, then the other.

However attractive this explanation may seem, another recent study which sets out to test precisely the communicating vessel view of ultimate attainment, Bylund, Abrahamsson, and Hyltenstam (2012), comes to the opposite conclusion. The researchers examined the L1 and L2 knowledge of L1 Spanish–L2 Swedish bilinguals, whose AoA ranged between 1 and 11, but their mean length of residence (LoR) was 23 years. This study focused on grammatical competence and semantic intuitions, which they tested

[6] A negative correlation is a relationship between two variables such that as the value of one variable increases, the other decreases.

through an aural grammaticality judgment task and a cloze test.[7] When compared at the group level, the bilinguals did not perform as well as the native speakers. However, when individual results were scrutinized, it turned out that around half of individual learners performed within the native speaker range. Importantly for us in this section, however, those who did well on Spanish, their L1, also did well on Swedish, their L2. The higher scores were correlated with higher language aptitude (as measured by a standard aptitude test). Thus the negative correlation, or interference, between the first and the second language uncovered by Yeni-Komshian et al. (2000) was not detected in this study. However, this discrepancy can very well be due to the fact that the former study tested pronunciation, while the latter tested grammatical and semantic knowledge.

In sum, research shows that the entrenchment of the first language hampers the establishment of the second language, but it is also possible that this could be the case only in the area of pronunciation. In the other areas of the grammar, however, the native language does not block the achievement of nativelike L2 proficiency.

4.5 Effects of bilingualism: Is the bilingual two monolinguals in one mind?

Put very simply, *bilingualism* is the ability to use two languages. But, as François Grosjean (1989) famously quipped, the bilingual is not two monolinguals in one mind. It has been suggested (Ortega 2009) that, in all fairness, bilinguals should not be compared to monolinguals, and that bilingualism itself makes this comparison flawed/erroneous. This is a line of argumentation certainly worth considering. However, defining bilingualism has turned out to be problematic in more than one respect. Let's unpack this issue a little.

[7] A cloze test is a test of global language proficiency, involving lexical, grammatical, and pragmatic knowledge. Some words are omitted from a complete story and the speaker is invited to fill in the gaps. In some versions of the cloze, every sixth or seventh word is deleted, so the gaps can be grammatical morphemes but also full lexical items. The exact word scoring method is appropriate for such a cloze test, and even native speakers do not score close to 100% on such a test.

Who is a bilingual? Definitions of bilingualism start from minimal proficiency in two languages and range over to nativelike command of two languages. An example of the former would be a British English native speaker with high-school French conversing with a local taxi-driver on a visit to Paris. An operational definition of this minimal level of bilingualism is being able to produce comprehensible and grammatical sentences in the second language. An example of the latter would be a person who grew up bilingual in French and English in Montreal, Canada or a speaker of Spanish as a heritage language who grew up in California. In yet another definition of bilingualism (Grosjean 2001), a person is considered bilingual if she or he uses two or more languages in everyday life; this type of definition capitalizes on frequency and functionality of use.

When we consider the effects of bilingualism on nativelike attainment, we should probably adopt a middle-of-the-road but flexible definition: one that is beyond the incipient view but stops long before full command of two languages. In this sense, a bilingual and an L2 learner are the same, for the purposes of this textbook. The reason for assuming such a practical definition is the robust evidence showing that the two languages of a bilingual are constantly activated, even in situations when only one language is needed to carry out a task. It seems it is virtually impossible to switch off the language not in use. Parallel activation of a bilingual's two languages can be observed in reading, listening, and in planning speech. This parallel activation happens not only at initial stages of acquisition but in advanced proficiency stages as well. As Bialystok puts it, "[T]his situation creates a problem of attentional control that is unique to bilinguals—the need to correctly select a form that meets all the linguistic criteria for form and meaning but is also part of the target language and not the competing system" (2009: 3–4).

Let us review some hard evidence that this parallel activation is happening. Within the mental lexicon, accessing a word in one language leads to the activation of related words in the other language, both in comprehension (Spivey and Marian 1999, Van Heuven, Dijkstra, and Grainger 1998) and in production (Colomé 2001, Costa, Miozzo, and Caramazza 1999). Within syntax, grammatical information from the irrelevant language is constantly available in bilingual individuals processing language. For example, Hartsuiker, Pickering, and Veltkamp (2004) found that Spanish–English bilinguals, describing cards to each other in a dialogue game, used the passive construction in English significantly more often after they had heard it in Spanish than after an intransitive or active sentence. The authors

argued that their findings support a view of syntactic representation being integrated between the two languages of the bilingual.

Secondly, recent brain studies have shown that the neural systems engaged by the bilingual's two languages are largely the same. Since there is a high level of interaction between the two language systems, it has been shown that, unsurprisingly, the native language influences the second language, but more surprisingly, the second language also influences the native language in various ways. A demonstration of this influence is offered by a study on English–Welsh bilinguals, Thierry and Sanoudaki (2012), which showed that even when reading in English, early bilinguals were mentally open to the Welsh order of Noun–Adjective, even though it is ungrammatical in English. In short, a bilingual's two languages may come to function somewhat differently than either language in a monolingual native speaker (Kroll, Bogulski, and McClain 2012).

If a bilingual's mind functions differently from a monolingual's mind, then a comparison between the two is not only unfair but uninformative with respect to nativelike ultimate attainment. By this logic, bilinguals will never be able to achieve monolingual levels, since their mental lexicons and grammatical systems will be influenced by their second (or multiple other) language(s). Their competence will not be deficient with respect to the monolingual gold standard; it will just be different. Comparing monolinguals and bilinguals, then, will be like comparing apples and oranges. Many researchers have indeed adopted the position that age effects on nativelikeness should only be tested among early and late bilinguals, and that the monolingual native speaker's performance should not be used as a yardstick against which the bilingual performance is evaluated (Birdsong 2009, Muñoz, and Singleton 2011).[8]

A related question that deserves our attention is whether bilingualism (the effects of two languages in one mind) completely rules out nativelikeness, irrespective of age of acquisition. The short answer is no, and the evidence in this respect is quite strong. We shall cite some examples concerning the performance of simultaneous bilinguals and of adult bilinguals, both tested on meanings of functional morphology. One caveat to keep in mind, though, is that research to date rarely reports whether the control groups

[8] However, it is a worthy scientific goal to understand the mind of bilinguals and the minds of monolinguals. It is just not fair to use monolingual standards in judging bilinguals.

of native speakers are indeed really monolingual or whether they might also be bilingual in another language (not tested). Thus we are not in a position yet to evaluate bilingual achievements in a two-way comparison: with a monolingual control group and with a bilingual control group.

Kupisch (2012) examined knowledge of specific (e.g., *The cats are on the mat*) and generic nominals (e.g., *Cats are beautiful animals*) in simultaneous bilinguals of German and Italian, individuals who grew up speaking both languages from birth (see next chapter). Some of her participants were dominant in Italian, while some were dominant in German, and there was a positive correlation between dominance and length of stay in one of the countries. Those who grew up in Italy were Italian-dominant, whereas the opposite held for German. The Italian-dominant bilinguals scored at ceiling with accuracy of 97 to 100% at judging the acceptability of bare, definite, and indefinite nominals in Italian. German-dominant bilinguals faced considerable problems in contexts in which Italian differs from German. Such results show that bilingualism on its own does not automatically lead to deviant grammars and that language dominance, obviously conflated with length of exposure and language use, has a crucial effect. This study did not employ a control group but just showed that the dominant bilinguals performed at ceiling in the dominant language.

Studies addressing the same issue, the effect of bilingualism, but looking at adult-onset bilingualism are the Montrul and Slabakova (2003) and the Donaldson (2012) studies discussed in Section 4.3. Since these studies showed that groups of learners as well as individual learners can be demonstrated to perform within the ranges of native speakers, we have to conclude that bilingualism on its own does not preclude nativelikeness in a second language. It is also quite clear that neither age of acquisition nor the effect of bilingualism prevents nativelike performance among nonnative speakers, in some areas of the grammar. We turn to other factors likely to influence bilingual performance in the next section.

4.6 The importance of the input

We already pointed out the importance of this factor when discussing the Kupisch (2012) study above: simultaneous bilinguals with two languages learned at birth, who had not maintained a relatively good balance between their two languages, were not as proficient in their weaker language as they

were in their dominant language. This and many other studies point to the conclusion that age of acquisition is not an all-important or completely decisive factor for ultimate attainment. Maintaining full and diverse linguistic input in the L2 through constant contact and varied usage of that language can be an equally important factor.

As we saw above, researchers who argue for the Critical Period Hypothesis, or at least age effects in second language acquisition, show that the earlier an L2 learner is exposed to the second language, the more likely she is to achieve nativelike competence. Summarizing over the present research, for example, Long (2013: 5) maintains that with respect to pronunciation, nativelike performance is most likely (although not guaranteed) for those with an AoA between 0 and 6, still possible but less likely with an AoA between 6 and 12, and impossible after that. Nativelike morphology and syntax are most likely if the L2 was acquired between the ages of 0 and 6, but highly unlikely after the mid teens. These are strong pronouncements indeed. The flaw of this argument is that it only focuses on one, however important factor, while ignoring other important factors. For instance, as Flege (2009) argues, Long's claims imply that late learners receive input equal in quality and quantity with early learners, but they do not use it in similar ways so as to reach a nativelike grammar. However, there is a dangerous confound in such an interpretation, since in most cases early learners have enjoyed decades of native input while late learners have received significantly less exposure. Furthermore, the differences in the input available to learners can be of different length, but also of different quality. We shall look at both quality and quantity of input below.

Even among simultaneous bilinguals acquiring two languages from birth, quantity of input can influence language development. Elin Thordardottir (2015) reports on vocabulary acquisition in French–English bilingual children in Montreal using the one parent–one language system. Those children who had equal amounts of exposure to the two languages, as measured by parental reports, exhibited monolingual-like receptive vocabulary skills. Those children who did not receive equal exposure to French or English had stronger vocabularies in the language to which they had more exposure. These findings suggest that reduced quantity of communicative interaction with the language provider (the parent) impacts even simultaneous bilinguals.

In order to illustrate the significance of the input, we move from child bilingualism to foreign language learning. The relative effects of input

exposure and starting age on foreign language acquisition[9] were investigated by Muñoz (2014). The author analyzed the oral performance of 160 learners of English in Spain and related it to various input measures: number of years of instruction, number of hours of curricular and extracurricular lessons, number of hours spent abroad in an English-speaking setting, and current contact with the target language. She measured oral performance through lexical diversity, speech rate, syntactic complexity, and overall accuracy on a film retell task. Muñoz shows that input has a stronger association with measures of oral performance than age of acquisition. She argues that cumulative exposure and especially contact with high quality input are better predictors of oral performance in the foreign language than AoA.

The clinching point in this debate is offered by the case of heritage speakers (Montrul 2008), who should be more successful than adult learners since they are child learners and actually are native speakers of the heritage language. When a heritage speaker starts school (or preschool), the majority language exposure and use typically increases considerably, as measured by hours in a typical day, but also in communicative importance, while the home language typically recedes in all these respects. Remember the tilted communicating vessels metaphor? It is certainly true for the hours in a day when a bilingual speaker comprehends and produces her heritage language or the majority (second) language: the more L2, the less L1. Thus the native language, although chronologically first and hence native, gets strong competition from the communicative input of the majority language. What happens as a result? To cite just one example, Montrul (2009) reported overall findings from a large-scale study comparing 70 post-puberty L2 learners and 67 adult heritage speakers in different areas of Spanish morphology and syntax. She demonstrated both comparable error patterns for the two groups and advantages for the early bilinguals in some areas, arguing that an early start may not be crucial in language acquisition, while the input quantity and language use may be just as important. In other words, age of acquisition cannot be the most important factor for nativelike linguistic performance, since heritage language learners have spoken that language from birth, but are not comparable to those native speakers who

[9] Foreign language acquisition refers to learning a language in a country where that language is not the national language, such as English in Spain.

have maintained constant usage (receptive and productive) of that language. In some respects, heritage speakers' knowledge is comparable to the knowledge of adult L2 learners. We will expand on this issue in Chapter 6.

How about quality of the linguistic input? Imagine the situation when a Russian-speaking child is growing up in the USA, in the family of two Russian native speaking parents and even some grandparents. While the child is being raised at home, it is possible to maintain the input quality (and quantity) roughly comparable to those of a monolingual child being raised in Russia. But when the same child starts school, the English spoken there often becomes the dominant language. In addition to the change in hours of language practice, the bilingual child's two languages can become differentiated according to topics or areas of life normally discussed in one or the other language. It is often the case among heritage bilinguals that they can talk freely and fluently about every day, home-related topics in their heritage language but they are more comfortable discussing political events or professional topics in their second language. This functional differentiation affects the mental lexicon as well as the syntactic complexity in the two languages. For instance, it is well known that the English passive construction (*The truck was hit by the car*) is not frequent in everyday communication, while it is more frequent in written language and more elevated registers.[10] Hence, we can expect higher accuracy with the passive in speakers of English as a second but dominant language, who use it for academic and professional purposes.

What kind of linguistic input is most useful for successful acquisition? In other words, what is high quality input? It has to be diverse, wide-ranging and rich in registers (home language, school language, professional language, etc.). It also has to be socially and communicatively important for the individual. Even if we disregard the different length of exposure, which we should not, there is some evidence that the quality of the linguistic input to which younger and older L2 learners are exposed may be different. Jia

[10] A register is a variety of language used for a particular purpose or in a particular social setting. In formal versus informal settings, speakers unconsciously change their way of speaking according to the situation. An example of a formal setting would be talking to a policeman or appearing in court; an example of an informal setting would be talking to one's friends at a party or in the pub.

and Aronson (2003) is a longitudinal study monitoring Mandarin-speaking children and teenagers acquiring English in a naturalistic environment. The study uncovered that the youngest children (ages 5–9) had much richer communicative interaction with native peers while the teenagers (aged 12–16) had less rich social interactions and linguistic environment. The authors attribute the quick change of language dominance, from Mandarin to English, among the younger children to their rich linguistic input.

Finally, a requirement of the linguistic input that is often overlooked is comprehensibility in a communicative situation. It is highly unlikely that one will learn a language by listening to the radio, even if one is exposed to it eight hours a day. Why is that? There are numerous cases in language acquisition when a perfectly acceptable sentence is inappropriate in the context of the previous conversation. A simple illustration would be using a pronoun such as *he* (e.g., *He is coming*) when no referent for that pronoun has been mentioned in the preceding discourse. In that case, the interlocutor is justified in asking, "*Who is coming*?" The utterance *He is coming* is grammatically complete, but contextually inappropriate. Knowing a language includes being able to produce grammatical sentences, but also knowing in what context they are appropriate and relevant. Thus, it is imperative to heed O'Grady, Lee, and Kwak's (2009: 72) warning that "[i]n considering the role of input frequency in language acquisition (first or second), it is vital to bear in mind a key point: what counts is not how many times learners hear a particular form—it is how many times they encounter mappings between a form and its meaning." In sum, language input is useful only when the communicative discourse situation can be mapped onto the linguistic sign and the fit between the two can be evaluated in social interaction.

4.7 An indirect way of appreciating the importance of input

In the previous section, we surveyed some evidence for the claim that the quantity and quality of the linguistic input is a crucial factor in achieving nativelike proficiency in a second language. In this section, we will consider another type of evidence for the importance of input: the effect of linguistic input on monolingual first language acquisition. I will be using two terms widely applied to linguistic performance: variability and optionality. Variability in language acquisition refers to the inconsistent application of the

target language rules, or the deviation from such rules.[11] Optionality refers to the simultaneous use of more than one form with the same meaning, e.g., *I no like chess* and *I don't like chess* at the same stage of development. The two terms are often used interchangeably.

The effect of variable input on child grammar development is demonstrated in a fascinating recent study by Miller and Schmitt (2010). The authors take advantage of existing dialectal differences in Spanish to test a situation that they could not have created experimentally. Chilean Spanish and Mexican Spanish differ in phonetic realization of plural morphology. In Mexican Spanish, plural is overtly realized as [s] on nouns, adjectives, and determiners, while in Chilean Spanish (subject to sociolinguistic variation) this piece of inflectional morphology undergoes a regular process of lenition (weakening) to aspiration or to nothing. In this way, plural morphology is not completely absent in Chilean Spanish, but it is rendered unreliable as linguistic evidence, being pronounced about 50% of the time. Miller and Schmitt (2010) report on the production of adults as well as children (mean age of 5;2–5;3) from different socio-economic groups: middle class and working class, tested by three different tasks. Both younger and older Mexican (working class) children were significantly more accurate on comprehension of plural than their Chilean counterparts. The authors argue that the more variability/ambiguity in the input there is, the longer it will take the learner to converge on the adult grammar.

While Miller and Schmitt (2010) studies the effects of variability in the input on child grammars, Meisel, Elsig, and Bonnesen (2011) addresses the effects of quantity and quality of the input on adult native grammars. French has a variety of interrogative constructions, some used more often in colloquial speech and others used in more formal varieties. For example, subject–verb inversion as in example (4) is almost non-existent in the input to children before they go to school and are exposed to standard French there.

(4) Quand arrive le train?
when arrives the train
'When does the train arrive?'

[11] Variability can also refer to the different rates of acquisition and different outcomes of the process, e.g., not all L2 learners attain the same proficiency after a set amount of exposure, say, two years of high school Japanese or ten years in the L2-speaking country.

The researchers wanted to investigate what happens if children are not exposed to an interrogative construction before they go to school. They reasoned that this construction will then be learned as a second language construction and will be inherently unstable in their grammar. Meisel et al. (2011) employed a grammaticality judgment task and tested adult French native speakers. All of their participants performed consistently on the standard interrogative constructions that are supported by evidence in colloquial French (subject–clitic inversion, complex inversion). In contrast, the question types as in (4) that are only present in standard French exhibit a lot more cross-individual variety as well as more inconsistency within the performance of the same individuals. The authors interpret this behavior of native speakers as essentially indistinguishable from that of second language speakers. Characterizing their native participants' grammar, the authors say: "They do, of course, acquire knowledge about these constructions, but as is evidenced by the broad range of variability in their ratings, it is afflicted by persistent optionality as is typically encountered in L2 acquisition" (Meisel et al. 2011: 380).

This conclusion brings us back to the very interesting question of what constitutes native and nonnative knowledge of language. Obviously, if limited exposure to particular constructions results in optionality in native grammars, then nonnative grammars, also characterized by variability and optionality, are highly nativelike indeed, at least as far as these constructions are concerned. This type of reasoning is supported by another recent body of work, that of Dąbrowska and colleagues (see the epistemological issue of the journal *Linguistic Approaches to Bilingualism* 2:3 for a review and commentaries). One study, Street and Dąbrowska (2010), argues that simple sentences employing the quantifier *every* (e.g., *Every cat is on a mat*) and the reversible passive voice (e.g., *The soldier was hit by the sailor*) present a comprehension challenge to native speakers of English with a low level of education. This is because they do not encounter such sentences often enough to learn them. However, after a brief explanation and practice of the constructions with the adult native speakers, comprehension improved dramatically and for the long term.

A training study, Wells, Christiansen, Race, Acheson, and MacDonald (2009), systematically manipulated participants' exposure to relative clause constructions (e.g., *The senator [that the reporter attacked _____] admitted his error*) over the course of three sessions spanning nearly a month. Two groups of undergraduate students were matched for verbal working

memory: their measures on a reading span test were at the low end. Over three training sessions, the experimental group was exposed to equal amounts of subject and object relatives. The control group received an equivalent amount of complex sentences, but without the inclusion of relative clauses (i.e., they read complex sentential complements and conjoined sentences). After training, the two groups' reading times on relative clauses diverged such that the speed of processing of the experimental group resembled the pattern for high-reading-span individuals, whereas the control group showed the kind of reaction time profile associated with low-reading-span individuals. These findings suggest that the attested differences in processing relative clauses were due to the amount of experience with this structure and not to differences in working memory. Let me reiterate the main argument of this section: variable linguistic performance is attested in monolingual native speakers, too, and it is clearly related to lack of exposure to a certain construction or to input that itself exhibits optionality.

4.8 Conclusions

We set out in this chapter to examine the issue of whether there is a critical period for second language acquisition, after which bilingual individuals cannot achieve nativelike performance. There are several parts of this question that are up for debate. First of all, researchers these days talk of sensitive periods, or age effects in language acquisition. Sensitive periods most likely differ for the different areas of the grammar, with mastering the sounds and intonation of the L2 to nativelike levels being considerably more difficult than mastering the sentence structure or the interpretation of sentences and discourse. It is fairly well established that if L2 acquisition comes after adolescence, complete and global nativelikeness is difficult, by some accounts impossible to achieve. Such nativelikeness will not only make an L2 speaker difficult to distinguish from native speakers, but she will have to perform in the range of native speakers on a variety of challenging and rigorous tests spanning phonology, syntax, and semantics. Linguistic theory offers principled answers to the questions of why some areas of the grammar are more attainable to nativelike levels than others.

We also considered the effects of bilingualism on nativelike attainment, pointing out that the brain of a bilingual and that of a monolingual speaker,

if not morphologically different, are perhaps functionally different as they process language in distinctive ways. Age effects and the effects of bilingualism on nativelike language use can augment each other. Perhaps the most significant challenge to strict critical period claims come from instances of language acquisition where quantity and quality of linguistic input trump age of acquisition, as is the case with heritage speakers. Even among monolingual native speakers, the effects of exposure to a particular grammatical morpheme (the plural) or grammatical construction (passive, relative clause) can lead to variable production and comprehension. If we assume that L2 speakers have fewer opportunities to interact with meaningful language in communicatively significant situations, then their nonnative performance may be due to the input factor. Age of acquisition is certainly of great importance, but the linguistic input may override its importance in the achievement and maintenance of nativelike linguistic competence.

Finally, as in every complex learning process, it is very likely that a variety of factors influence successful acquisition of a second language, including factors such as language aptitude and the learner engagement with the language, also known as motivation. I focused on the importance of bilingualism, exposure and input in this chapter, because these are the factors most closely related to the learner's internal linguistic system, which is at the center of this textbook. For an appreciation of the other factors and their contribution, see the special issue of the journal *Language Acquisition* (2014, issue 4) edited by Carmen Muñoz.

Teaching relevance: Away with the pessimistic message

Why would we worry at all about the Critical Period Hypothesis when we teach a foreign or a second language? How about for psychological reasons? People want to know whether what they do is futile or whether it can make a difference. Of course pronunciation is not taught in most language classrooms with the aim of learners eventually sounding like native speakers, but rather with the aim of comprehensibility. However, teachers should know that their efforts in the classroom are not futile, and that their students can, eventually, become nativelike in many other respects of language competence. We should approach language acquisition with the message of what is possible to acquire, which areas of the grammar students can make real progress in, and where they should concentrate their efforts: the acquisition of morphology, syntax, semantics, and pragmatics.

4.9 Exercises

Exercise 4.1. The following Q and A exchange was found on Yahoo. Read it carefully and discuss. How would YOU answer the question? Make a list of skills or capacities that would help an individual to be classified as a near-native speaker. At all times, consider what the range of native speakers might be.

Question: I am a native speaker of English and I have been describing myself as a "near-native" speaker of Greek ever since I discovered the term. I would just like to check that my definition conforms to the popular consensus:

I started learning my second language (Greek) when I was 8 and this process was continued up until I was 22 (I'm nearly 24 now). I spoke it all the time with my grandmother and have her accent and, like her (who was Cypriot), I am always mistaken for an Athenian because my accent is rather posh for a Cypriot - but still, it's a native Greek one. Following the years of study, practice, immersion (both in and out of Greece and Cyprus) and lessons I can confidently say I'm fluent. The thing is, like anyone, I still make mistakes in phrasing which give me away as an Englishman and not a Greek.

My understanding of a "near-native" speaker is someone who has no trouble understanding anything that is said to him or that he reads and no nuance or play on words is missed and speaks well enough that native speakers feel confident that they can talk at native speed and speak as freely as they want, not restricting themselves to facilitate the understanding of the "foreigner". Also, he can speak so well to a point that until there's a slip up in word-choice or phrasing, the listener believes they are speaking to a fellow-countryman.

I feel I fit into that above category and will be including it on my CV. However, that description might be wrong and I have been mis-using the term to describe myself for ages. Please, tell me your opinions and share your own definitions if mine is incorrect.

Thank you.

One answer: A near-native speaker understands almost 100% of a novel or a paper, listens to a native speaker and understands almost everything, can write with "dignity", recognizes the different "languages" (formal, informal, semi-formal etc.), usually uses good syntax, understands 99% of humor in the foreign language, can follow a phone conversation with a native speaker, is fluent when he speaks etc. From my own experience the most difficult thing to get is a PERFECT PRONUNCIATION. There is usually a "little something" almost "invisible" there that shows you are a foreigner.

I've heard many foreigners in my country who speak very well but whose accent has "something" if they immigrated after when they are 12 years old or the like. The vocabulary level depends a lot on how many books you have read, if you have studied Geography, Biology, Physics, Chemistry, etc. in your near-native language. I think that according to what you say, you could be called a near-native.

Exercise 4.2. Figure 4.1 is from Abrahamsson (2012) (his Figure 2). The study tested 200 participants, Spanish-native L2 speakers of Swedish (LoR in Sweden > 15 years), on an auditory grammaticality judgment task (GJT) and on perception of voice onset time (VOT) with the pair of words *par* and *bar*. Recall that VOT is the length of time that passes between the release of a stop consonant and the onset of voicing, and is heard as how much voicing you put on your consonants.[12] Recall also that Spanish and Swedish, the L1 and L2 of the learners, differ in VOT values. The individual results are plotted below. Each little quadrangle corresponds to the average performance score of an individual, while the line represents central tendency. *R* is the measure of the correlation between the two plotted variables: age of onset of acquisition and performance score. The higher its value, the stronger the correlation.

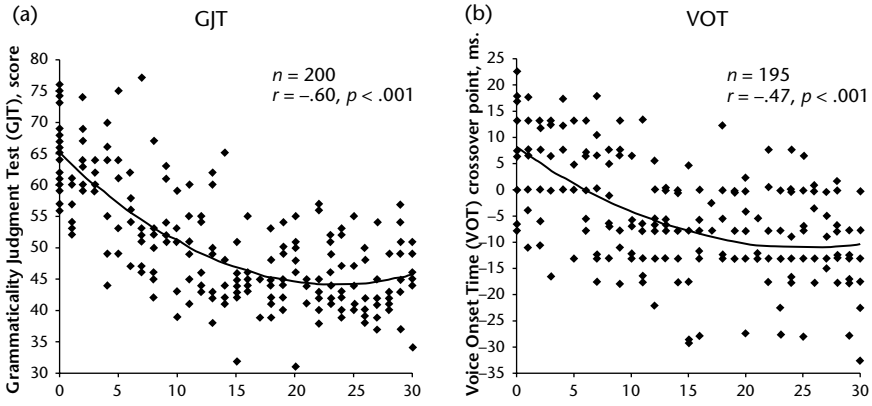

Figure 4.1 Grammaticality Judgment Task scores and VOT crossover points plotted against Age of Onset of acquisition, from N. Abrahamsson (2012)
Reproduced with permission

[12] For additional explanation of VOT with spectrograms, see the following video: https://www.youtube.com/watch?v=b9NuwOLiyss. For a layman's explanation of VOT, check out this one: https://www.youtube.com/watch?v=NKQBWnXMpns.

Looking at Figure 4.1, answer the following questions:

Question 1: What is the range of scores of the native speaker controls on the two tasks?

Question 2: Are there individual learners who score higher than the natives? On which task?

Question 3: Are there individuals who score within the native speaker range at all ages of onset of acquisition?

Question 4: Do you see a critical age in the graphs, after which no individual learner is within native speaker range?

Question 5: Do you think these findings support or refute a strict version of the Critical Period Hypothesis?

Exercise 4.3. Part 1. Abrahamsson (2012) posed another research question: to establish whether the GJT scores and the VOT scores would correlate for early learners, but not for late learners and native speakers. The author explains the theoretical motivation of this hypothesis as follows:

> Behind the prediction lies the assumption that grammatical and phonetic intuitions should develop more or less simultaneously and to a similar degree if the language has been acquired automatically, incidentally, and implicitly as an interdependent or interconnected whole, but not if it was learned consciously, intentionally, and explicitly as independent, separate parts of a whole. This, in turn, would potentially suggest that early and late L2 learners use fundamentally different systems: Although children automatically acquire the morphosyntactic and phonetic–phonological system "from mere exposure" (Lenneberg, 1967, p. 176) through innate, domain-specific mechanisms [. . .] and by using mostly procedural memory resources [. . .], adults have lost most of these abilities and instead must learn the L2 consciously, through formal instruction, and via their domain-general cognitive system, using mostly declarative memory resources. The consequence is that early learners develop implicit linguistic competence (or intuition) very similar to that of native speakers, whereas adults typically end up with mostly explicit (some of which can be equated with metalinguistic) knowledge, which cannot be used as efficiently for spontaneous and effortless language production and perception. (Abrahamsson 2012: 209–10)

Question 1: Do you agree with this reasoning in general? Why or why not? In particular:

a. Why should GJT and VOT scores correlate in early learners but not in natives?
b. Consider the distinction between *implicit, unconscious acquisition*, where different parts of the linguistics system are acquired simultaneously and systematically versus *explicit, intentional learning* where conscious effort and attention is involved. This distinction is made by Paradis (2004, 2009) and other scholars. They relate implicit

acquisition with procedural memory and explicit acquisition with declarative memory, arguing that the latter is predominantly used in SLA after puberty. Does it follow from the implicit–explicit distinction that L2 learners would either learn the phonetics or the syntax of the second language, depending on what they decide to pay attention to?

c. To cite Abrahamsson again:

> . . . the adult learner treats the different levels and sublevels of the L2 as independent puzzles, some of which the learner can choose to focus on in depth, and some of which can—consciously or unconsciously—be disregarded as either uninteresting, unnecessary, or unlearnable, if not left entirely unnoticed." (Abrahamsson 2012: 194).

Do you agree with his logic? Why or why not?

Part 2. Consider the results of the correlations plotted in Abrahamsson's Figure 3, which are reproduced in Figure 4.2.

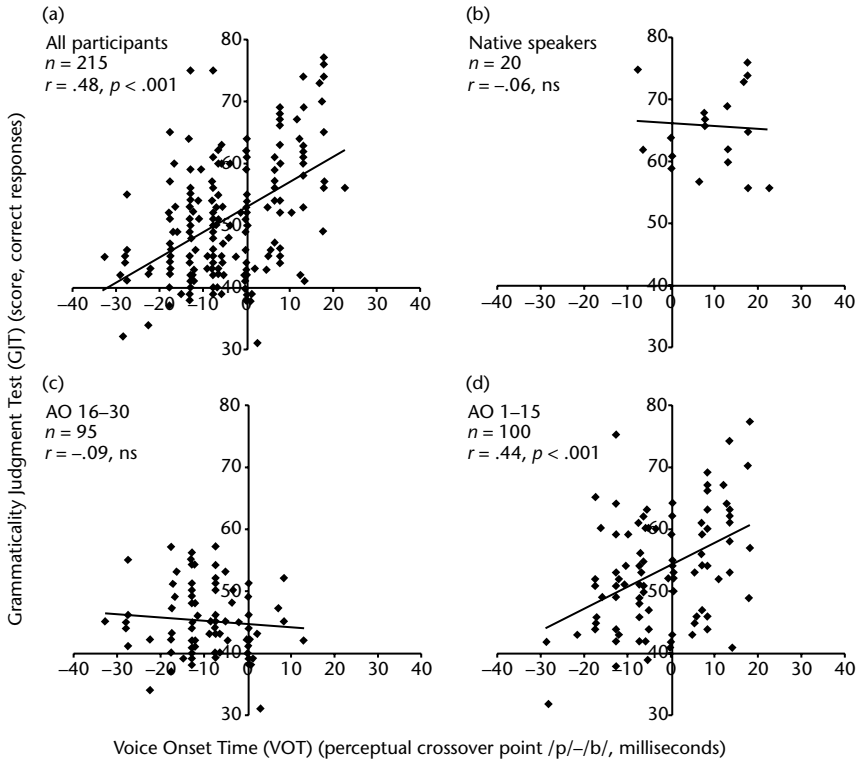

Figure 4.2 Scatterplots with Pearson's correlations between Grammaticality Judgment Task scores and VOT measurements in (a) all participants including the native speakers, (b) just the native speakers, (c) just the late learners, (d) just the early learners, from N. Abrahamsson (2012)

Reproduced with permission

Question 2: Describe what you see in each quadrant. Look at some individual results in quadrants (c) and (d). Point to learners who are good at phonetics and not so good at grammar. Do you see learners who are good at grammar but not nativelike at phonetics? Circle such individual results.
Question 3: Which group shows a significant correlation between the two types of scores (that is, between grammatical and phonetic intuitions)?
Question 4: What does this mean? Are these results in support of Abrahamsson's prediction?

Part 3. Niclas Abrahamsson, the author of this study, writes that the results supporting the prediction of early versus late acquirers using different mechanisms of acquisition is only partially supported by the data, therefore "the empirical results as well as their theoretical interpretation are of a more suggestive and tentative nature."
Question 5: Do you agree with him? Why, or why not?
Question 6: Can you think of a better way to test whether late learners acquire implicit linguistic knowledge?

Exercise 4.4. In the Montrul and Slabakova (2003) study, learners were asked to evaluate the truthfulness of the test sentences in the context of the preceding story. Each story was followed by only one test sentence; the same story with the second test sentence appeared elsewhere in the experiment. Read the stories and test sentences that appear below and answer the following questions:

(i) Generic event story:
Según el periódico, el restaurante de la calle Jefferson era muy bueno y el servicio era excelente. Lamentablemente el restaurante cerró el verano pasado y nunca tuvimos la oportunidad de ir.
'According to the newspaper, the restaurant on Jefferson Street was very good and customers were always happy with the service. Unfortunately, the restaurant closed last summer and we never got to go.'

a. Se comía-IMPF bien en ese restaurante. True
'One ate well at that restaurant.'

b. Se comió-PRET bien en ese restaurante. False
'We ate well at that restaurant.'

(ii) Specific event story:
Según la mayoría de la gente, el restaurante de la calle Jefferson era muy bueno y el servicio era excelente. Fuimos allí a celebrar el

cumpleaños de Carlos y a todos nos gustó mucho. ¡Qué lástima que lo cerraron!

'According to most people's opinion, the restaurant on Jefferson Street was very good and the service was excellent. We went there to celebrate Carlos's birthday and we all liked it a lot. It's a pity that it closed!'

a. Se comía-IMPF bien en ese restaurante. True
'*One ate well at that restaurant.*'

b. Se comió-PRET bien en ese restaurante. True
'*We ate well at that restaurant.*'

Question 1: What is the only difference between the two sentences under each story?

Question 2: What is the tense in the English translations? Is this a syntax–semantics mismatch, in the sense that two Spanish aspectual tenses are translated with the same tense in English?

Question 3: Is knowledge of this syntax–semantics property likely to be implicit or explicit?

Question 4: Recall that we mentioned that meaning can only be learned if the linguistic signal is mapped to an extralinguistic situation exemplifying that meaning. Discuss possible acquisition situations in which this distinction in meanings could have been acquired from observation of extralinguistic situations (such as in these stories) paired with sentences such as these.

Question 5: What kind of feedback should a learner have received, so that she could have learned the mapping of story (i) and that test sentence (b) is not possible? Is it likely that learners hear such corrections from their interlocutors or teachers?

Question 6: Create another quadruple of two stories, one describing a habitual, generic event and the other a similar complete one-time event. Add two test sentences which differ only in the aspectual tense. If you speak Spanish, create the stories and test sentences in Spanish. Other Romance languages such as French and Italian also have similar tenses: check if they work just like Spanish. If you do not speak a Romance language, create the stories in English.