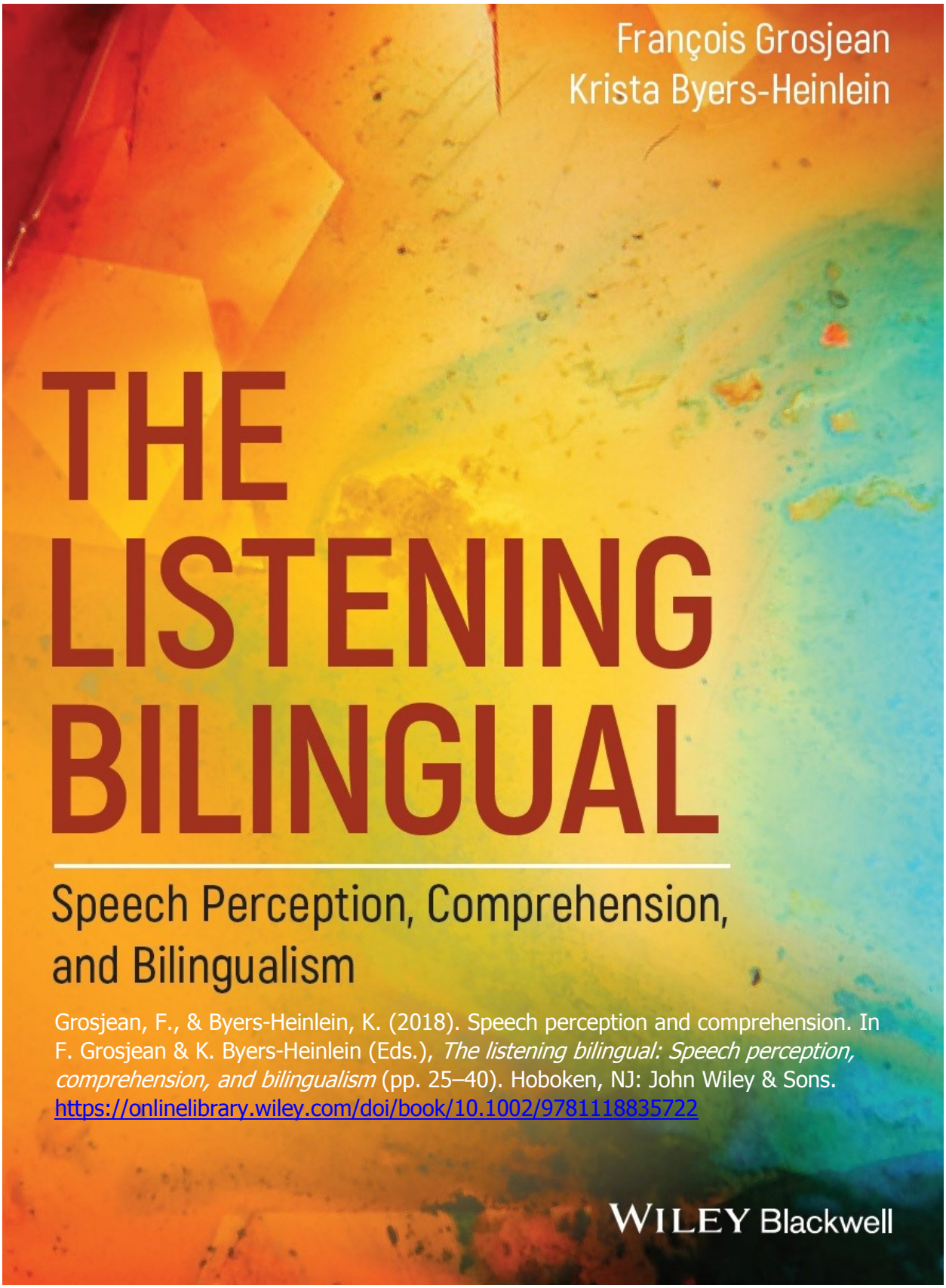
François Grosjean
Krista Byers-Heinlein
THE LISTENING BILINGUAL
Speech Perception, Comprehension, and Bilingualism
Grosjean, F., & Byers-Heinlein, K. (2018). Speech perception and comprehension. In F. Grosjean & K. Byers-Heinlein (Eds.), *The listening bilingual: Speech perception, comprehension, and bilingualism* (pp. 25–40). Hoboken, NJ: John Wiley & Sons.
https://onlinelibrary.wiley.com/doi/book/10.1002/9781118835722
WILEY Blackwell

2

Speech Perception and Comprehension

François Grosjean and Krista Byers-Heinlein

Introduction

Listening to speech, by which we mean processing and comprehending spoken language, is an amazing skill that takes place quickly and efficiently. We process some seven to ten sounds a second, and about 175 words a minute, often in a noisy environment that contains other sounds, and yet we do so effectively and rather effortlessly. The listening process is highly complex, and researchers over the years have tried to understand the processing stages that are involved from perceiving the acoustic wave all the way to building an enriched mental representation of what has been said. The challenge to understand this communication skill is even greater when the listeners, adults and children, have to process two or more spoken languages in their everyday lives.

Before entering into the world of speech perception and comprehension in bilinguals, as we will do systematically in the following chapters, it is important to explain briefly how speech processing takes place generally, and then to evoke the similarities and differences between monolinguals and bilinguals. This chapter therefore has two aims. The first is to give an overview of speech perception and comprehension, that is, to examine the basic components that are involved and the processing mechanisms that are present. We also need to explain briefly how children acquire spoken language since they do so mainly through speech input. The second aim is to present a short introduction of how bilinguals perceive and comprehend two or more languages – either separately or together in the form of mixed speech – and to point out major issues that are involved.

The chapter will be organized in three parts. In the first, we will describe the general architecture of the basic components required for speech processing in a language, in other words, the speech input, linguistic knowledge, processing mechanisms, and other sources of information. We will then detail each processing mechanism: speech perception and prosodic analysis, word recognition, syntactic and semantic processing, and pragmatic processing. In the second part, we will present the processing components involved when two languages are present, first when only one language is being processed and then when two languages are involved, as when the listener is listening to code-switches and borrowings. This will be followed by a discussion of bilingual issues that are common to all processing levels and that are in large part due to the coexistence, and interplay, of two or more languages in the bilingual listener's mind. Finally, the

The Listening Bilingual: Speech Perception, Comprehension, and Bilingualism, First Edition.
François Grosjean and Krista Byers-Heinlein.

third part is dedicated to children. We will present a brief outline of how language acquisition takes place and will cover topics such as language input, speech and phonetic perception, word learning and comprehension, and grammatical and pragmatic development. Then we will discuss some of the factors that impact speech perception and comprehension across processing levels in bilingual children – issues that will be taken up in more depth in later chapters.

From the Speech Wave to the Mental Representation

Listening to speech and processing what is being said – something we do every minute of the day and rather effortlessly – is a very complex process that research has investigated over the years. In this part, we will examine the basic components needed for it to occur, the processing mechanisms that are involved, and various general aspects of processing that make it so efficient.

Figure 2.1 presents the basic components needed for speech processing to take place. First, we have the speech input (or speech wave), which is produced by the speaker. It is presented at the bottom of the figure with an arrow pointing upwards and is often referred to as the "bottom-up" information. In the center part of the figure, we depict "linguistic knowledge", which represents the knowledge a listener has of the language in question (its grammar and lexicon), as well as the "processing mechanisms" that are needed to process the input.

Finally, there are "other sources of information" that listeners use to perceive and understand what is being said. These correspond to the context in which the speech situation is taking place, information about what has been heard so far, knowledge about the world, and so on. These sources are often referred to as "top-down" information (note the downward arrow in the figure) and they play a crucial role in processing, as we will see below. The final outcome of perception and comprehension is referred to

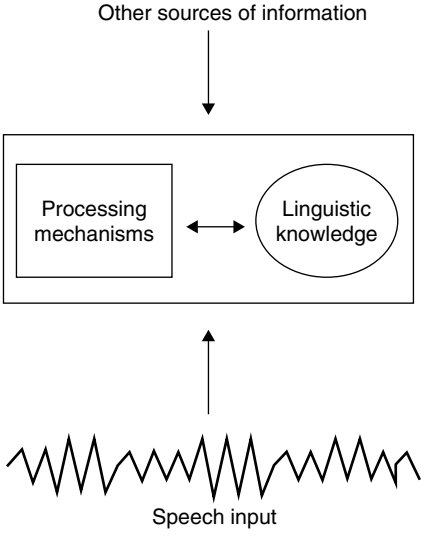

Figure 2.1 The basic components involved in speech perception and comprehension. This figure originally appeared as Figure 2.1 in Grosjean, François. 2013. "Speech perception and comprehension." In Grosjean, François, and Ping Li. 2013. *The Psycholinguistics of Bilingualism*, 30. Hoboken, NJ: Wiley-Blackwell.

as the mental (or interpretative) representation, in other words, the enriched meaning of what has been said.

In what follows, we will concentrate on the processing mechanisms – also referred to as "processing levels" – that are involved in going from the acoustic wave to the mental representation: speech perception and prosodic analysis, word recognition, syntactic and semantic processing, and pragmatic processing. We will deal with each in turn, and as we do so, we will refer to an utterance spoken by a person to another person concerning their common friend Mary: "The librarian gave the book to Mary!", said with a tone of surprise.

Speech perception and prosodic analysis start as soon as the beginning of the utterance is heard. In speech perception, the phonetic units that make up the utterance are identified, first the "th" (/ð/) of "the", then the "e" (/ə/), the "l" of "librarian" (/l/), followed by the "i" (/aɪ/), the "b" (/b/), and so on. Several acoustic cues allow us to identify these sounds and categorize them as elements of the language being processed, in this case English. As for the analysis of the prosody – those aspects of the speech that are not phonetic segments, hence the label often used for them, "suprasegmentals" – we process several acoustic characteristics of the input such as the evolution of the fundamental frequency over the utterance, the duration and the intensity of the phonetic elements, etc. These are combined in various ways to produce percepts such as pitch, length, loudness, rhythm, speech rate, etc. These prosodic variables help us identify phrase and sentence boundaries, specify the type of utterance being uttered (statement, question, etc.), point to words of importance in the utterance, signal the current emotional state of the person speaking, and so on. In the case of our example, "The librarian gave the book to Mary!", we perceive a slightly stronger stress on "gave" and a pitch that corresponds to surprise throughout the utterance.

As soon as the first sounds of the speech wave have been identified, word recognition begins. Based on just a few milliseconds of a word being uttered, that word is activated in our internal vocabulary (our lexicon), along with other words that match the speech signal. These candidates compete with one another and, little by little, as more information arrives, the recognition system narrows down the possibilities and finally selects the word in question. Along with this activation and selection process, various acoustic-phonetic and lexical cues help us identify such things as the beginning and end of the word, and hence they modulate the competition process. Many other factors will speed up or slow down word recognition, such as the frequency and familiarity of a word, the number of competing candidates, and top-down information given by the syntax, semantics, and pragmatics of the utterance. Once words have been recognized, we gain access to the information about them contained in our lexicon – their meaning or meanings (for example, a "book" is a written text), their morphology ("gave" is the past tense of "give"), their grammatical categories ("the" is a determiner, "librarian" is a noun, etc.), as well as the syntactic and semantic structures they can occur in.

With the information obtained during word recognition, syntactic and semantic processing can take place. We compute the syntactic structure of a sentence, a process known as parsing. Thus, "the" and "librarian" are grouped into a noun phrase, as are "the" and "book"; "to" and "Mary" make up a prepositional phrase; "gave", "the book", and "to Mary" are grouped into a verb phrase; and, finally, the first noun phrase ("the librarian") is grouped with this verb phrase to form a sentence. Thematic roles are also assigned, that is, the semantic roles that phrases have in relation to a verb. Thus, the verb "give"

has the structure "AGENT give THEME to RECIPIENT". In our example, "The librarian" is the agent (the instigator of an action), "the book" is the theme (the entity that is moved), and "Mary" is the recipient (the person receiving the theme). The outcome of both syntactic and semantic processing is the literal mental representation of the utterance. In our case, an object (the book) was transferred from one person (the librarian) to another (Mary).

Finally, pragmatic processing will use everything that we know about the context, the speaker, the people referred to, as well as what has been said before, our knowledge of the world, and of the rules of communication, to compute the utterance's enriched representation. In our example, to really understand it fully, we need to know that Mary had tried repeatedly to obtain that specific book from the library but had been told that it could not be borrowed. The fact that the librarian had finally given her the book – hence the tone of surprise detected during the prosodic analysis – means that something happened to make this possible. The final representation that we retain, therefore, is that Mary obtained the book from the librarian and that this was a very surprising event.

In this short description of how we go from the acoustic wave all the way to the enriched mental representation, we have inevitably simplified what takes place at each processing level. The operations are in fact highly complex and each of them, studied by different branches of psycholinguistics, are the object of much research and theorizing.

We will end this part with a few general aspects of processing, some of which researchers agree upon and others for which there is less common ground. The first, which progressively has obtained quite a large consensus, is that during speech perception and comprehension, we construct a mental representation incrementally. The system processes the information as it comes in – from the speech sounds all the way to the syntactic, semantic, and pragmatic elements – and progressively builds a mental representation. Two eye-tracking studies gave evidence for this incrementality quite early on (Tanenhaus et al. 1995; Sedivy et al. 1999) and are described in Chapter 5.

Another aspect, this one more hotly debated, is how interactive the perception and comprehension process really is. Those who defend interactivity, originally proposed in the pioneering work of Marslen-Wilson (1975), state that all levels of processing influence one another from the very start. On the other hand, there are those who propose that certain processes take place independently of others, at least initially. Here are two examples to understand these positions. The first concerns phoneme perception and word recognition, and the question asked over the years is whether lexical processing can influence the speech perception level. A model like TRACE (McClelland and Elman 1986) is highly interactive in that the word level can influence the phoneme level. Other models such as Shortlist (Norris 1994) are basically data-driven, at least when the first list of lexical candidates is proposed. The second example concerns syntactic and semantic processing and when it is that the latter intervenes. In a number of models (e.g., Frazier 1987; Friederici 2002), the first stage builds the syntactic structure and it is only in a second stage that semantic information can be processed. However, in more interactive models such as constraint-based models (e.g., Boland, Tanenhaus, and Garnsey 1990) the system uses numerous sources of information – syntactic, semantic, pragmatic – to do its work. Harley (2014), the author of *The Psychology of Language* and a recognized authority in psycholinguistics, leans towards the fact that language processing is strongly interactive, although there are constraints to this.

Finally, processing is predictive, that is, listeners anticipate what is coming next. Quite early on, Grosjean (1983) listed a number of aspects in the structure of language that facilitates prediction: at the level of discourse, old information usually precedes new information, and definite items often come before nondefinite items. At the level of syntax, some verbs are followed by specific types of phrases. At the level of the word, the sequence of sounds, from beginning to end, will often allow a word to be distinguished from other words long before its end, making the last part predictable. And, at levels below the word, phonological and morphological rules often give cues as to the next phoneme or next morpheme. Prediction is helpful in a number of ways: it reduces the set of possibilities and therefore helps focus the attention of the listener, it helps demarcate domains of processing, and it gives the listener time for other activities such as integrating the information that has been processed, storing it, and preparing a response if needed. Eye-tracking studies have examined predictivity and one such study by Altmann and Kamide (1999) is described at the beginning of Chapter 5.

Processing in Bilingual Adults

Everything that has been said so far concerning the perception and comprehension of speech also pertains to bilinguals. They too analyze the speech signal with a number of processing mechanisms or levels – phonetic and prosodic, lexical, syntactic, semantic, and pragmatic. They too have linguistic knowledge that they use during processing as well as other sources of information, such as what has been said so far, knowledge of the world, the context they are in, etc. And, of course, speech processing in bilinguals, like in monolinguals, is incremental, largely interactive, and predictive.

However, bilinguals perceive and comprehend two or more languages in their everyday life and hence the components depicted in Figure 2.1 will be multiplied by the number of languages concerned. In Figure 2.2, to simplify things, we present the processing components involved when only two languages are present – the listener is just bilingual – but any number of languages could be added. In the diagram, the person is listening to speech input that is monolingual – the language being heard is Language a (La) – and the linguistic knowledge and the processing mechanisms of that language are active, as is seen by the black rim around the left-hand box. The other language, Lb, is present but is deactivated, shown by the light gray rim. Note that the speech input is transmitted to both language systems and that the other sources of information feed into both systems.

If bilinguals find themselves in a bilingual mode, that is, they are speaking to people who share their languages, and code-switching and borrowing is taking place, then the guest language (Lb here) will also be active but less so than the base language (La), which is the main language of the exchange. This is depicted by making the light gray rim darker for Lb, showing greater activation, but it does not reach the level shown for La.

Chapters 3 to 5 will examine what takes place at the speech perception level, the word recognition level, and at the levels above (syntactic, semantic, and pragmatic). They will be followed by a chapter on mixed language processing and one on the clinical assessment of bilinguals. In addition to aspects that are specific to each level,

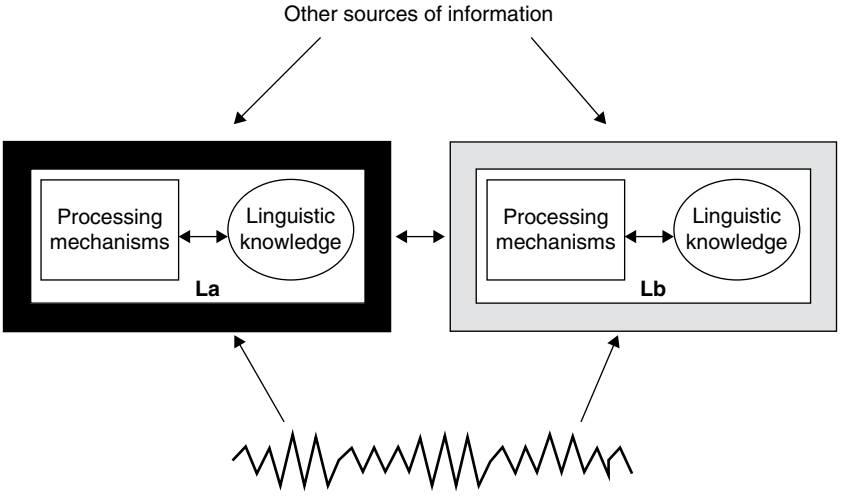

Figure 2.2 The components involved in speech perception and comprehension in bilinguals. This is a slightly modified figure which originally appeared as Figure 2.2 in Grosjean, François. 2013. "Speech perception and comprehension." In Grosjean, François, and Ping Li. 2013. *The Psycholinguistics of Bilingualism*, 33. Hoboken, NJ: Wiley-Blackwell.

bilingual issues that are common to all of them will appear in several chapters, and it is worth giving a short overview of them here. They are in a way transversal to these levels and are largely due to the coexistence of two or more languages in the bilingual listener's mind. As we go through them below, we will point out where it is they come up in the upcoming chapters.

Issues Common to All Processing Levels

The language history of bilinguals as well as the use they make of their languages, and their proficiency in them, will all be crucial in how well processing takes place. As we saw in Chapter 1, a bilingual's language history and language configuration is complex and covers such things as the languages acquired, when and how they were acquired, the pattern of proficiency and use over the years, the language that is dominant at the present time, how well is (are) the other language(s) known and used, and which language(s) is (are) written and read.

The actual languages known by the bilingual will have an impact on processing. Issues such the phonetic categories that are developed in a language (Chapter 3), the segmentation cues that will be used during word processing (Chapter 4), the syntactic representations that will be shared or not (Chapter 5), and so on, all have their origin, in part at least, in how similar or different the bilingual's languages are from one another. A situation that has been studied quite extensively is when the first language has only one sound category and the second language two, which are assimilated to just one category under the influence of the stronger first language. This may then lead to the activation of spurious word candidates during the word recognition process.

It can also be the case that, in addition to categories not existing, some processing mechanisms and strategies are not acquired because the first language does not have them and/or the second language was acquired later. We will see examples of this in Chapters 4 and 5 in domains such as the use of gender, or case information or the use of word segmentation cues such as allophonic variation (different phones are used to pronounce a given phoneme in different contexts). Of course, how well bilinguals know each language, and how much they use it, will also be critical. For example, at the speech perception level (Chapter 3), we will learn that bilinguals who use their first language often are more likely to exhibit first language influence during the perception of their second language, even when the latter was acquired early in life. In Chapter 4 we will also see that greater language proficiency and more language use result in improved use of segmentation cues. As for syntactic processing (Chapter 5), it would appear that when listening to a second language, bilinguals will make use of representations shared with their first language as long as the corresponding structures are similar enough, as stated above for example, they have the same word order, but also that the bilinguals are sufficiently proficient in their second language.

Language proficiency is a variable that permeates all levels of spoken language processing, including speech perception, word recognition, and semantic processing. For example, when bilinguals are equally fluent in their languages, they will be able to access words as rapidly in the one or the other language, but if they are dominant in one language, they will perform better in that language than in the other (see Chapter 4). Bilinguals fluent in a language will also be able to keep the other language at bay, at least to a large extent, and hence not suffer from the perceptual interference of the other language in a monolingual mode. Of course, things change dramatically when the interaction contains mixed speech and the listener has to process both languages within the same utterance (Chapter 6).

The age of acquisition of a language, that is, when it was acquired, will also be an important factor in processing. Thus, we will observe in Chapter 3 that the first language bilinguals acquire will have a persistent influence on the perception of some sounds in the second language, even when bilinguals become dominant in the second language over time. Of course, when the age of acquisition of the second language increases, its sounds are perceived less accurately. In Chapters 4 and 7, we will read how bilinguals who have acquired their language simultaneously often have comparable results to monolinguals when recognizing words and do far better than sequential bilinguals. However, when speech is degraded by noise, multitalker babble, etc., or is speeded up, everything becomes more difficult, even for simultaneous bilinguals (Chapter 7).

In sum, every level of processing in a language will be affected if the bilingual acquired it later in life, does not know it as well as the first language, and does not use it very often. This said, language by essence is highly redundant, communication takes place in context, and hence problems at one level of processing (e.g., the misperception of a sound, the momentary choice of an erroneous word candidate, etc.) will often be compensated by the other levels as well as other sources of information.

In addition to the issues related to language history, language use, and language proficiency, two other issues cross processing levels. They pertain to the presence of the other language (e.g., Language b, during the processing of one language, e.g., Language a). First, there is the activation and role of that language when the language

mode is monolingual and only one language (Language a) is being heard. In Chapter 3, we will see that when completing perceptual tasks, bilinguals may shift their speech category boundaries according to the base language they hear. In other words, they perceive the same physical stimulus as belonging to a different category when the base language changes. This touches on an issue researchers have long asked themselves about: is processing selective, that is, when one language is heard, only that language is active and is processed, or is processing non-selective, that is, the bilingual's other language(s) intervene(s) while processing is taking place? We will see notably in Chapter 4 that the answer depends on such variables as which language is being processed (the first or the second), how well the listener knows that language and the other – deactivated – language, when it was learned, the context, both linguistic and situational, the experimental set-up including the types of stimuli used (e.g., homophones, homographs), and so on.

The other issue that pertains to the presence of the other language during speech perception and comprehension concerns the actual processing of that other language (Language b) when it is brought in by the speaker in the form of code-switches and borrowings. One question that has been examined is whether the perception of code-switches takes extra processing time and, if so, how to account for it. Factors such as where it is the code-switch occurs in the sentence, the density of preceding code-switches, and the language mode the listener is in will all play a role. The time-course of code-switch perception is also an object of interest among researchers (see Chapter 6). As for the actual recognition of guest words (code-switched words and borrowings), many variables that account for their processing have been investigated: the language they are pronounced in, their phonotactic configuration, the preceding sentence context, the presence of near homophones in the base language, various acoustic-phonetic cues, and so on. The perception and comprehension of interferences that occur mainly in a monolingual mode have also been the object of study but to a much lesser degree.

Other issues will also appear transversally in the upcoming chapters. One of them is whether prediction is as efficient in bilinguals as in monolinguals. We will see in Chapter 5 that this is the case when the meaning of words is concerned but that it can be different when gender and case marking are involved. How about the revision of the mental representation in a second language? In Chapter 5, we will observe that initial comprehension processes are very similar in the bilingual's two languages, but that there are more difficulties during syntactic revision in the second language. Another issue is the role of reading (age acquired, preferred language, proficiency, etc.) in speech processing and comprehension.

The issue we will end with concerns how everything we know about the way bilinguals listen and comprehend, across all processing levels, can be of use to applied domains such as the clinical assessment of bilinguals. Some suffer from a breakdown in communication due to hearing and/or processing impairment, which limits them functionally in their everyday life. As we will see in Chapter 7, the clinician will therefore have to decide on the level of processing to test (syllables, words, connected speech, discourse), the language(s) to use, and whether to bring in some form of degradation to the speech (noise, multitalker babble, etc.).

In sum, speech perception and comprehension by bilinguals share many points in common with how monolinguals process speech, but they also have many specificities, proving once again that bilinguals are not two monolinguals in one person.

Processing in Bilingual Children

Like adults, children's comprehension of language requires going from the basic speech wave to a mental representation of what is said. Take, for example, the sentence, "I wish it wasn't raining outside." Comprehending this sentence takes knowledge of English, and involves decoding the speech sounds, locating words, and understanding how the order of the words affects their meaning. Yet, the listener must also have an understanding of abstract concepts such as "wish" and "outside". Older children may process and comprehend such a sentence in a mostly adult-like way. At the other end of the developmental spectrum, newborn infants may simply perceive a series of sounds, syllables, and melodies without extracting any meaning. At most ages, children are somewhere in between: they are experts in some aspects of their native language(s) but are still learning about other aspects. In tandem with the development of language, children are also building the social and cognitive capacities that help support language perception and comprehension. This is an important difference with adults, who are experts when it comes to their native language(s), and who master the many other social and cognitive capacities that come into play when perceiving and comprehending language. It is because of these fundamental developmental differences that we mostly discuss adults and children separately, covering bilingual adults in Chapters 3 to 7 and bilingual children in Chapters 8 to 11.

Children's language development occurs in a fairly predictable sequence, although there is considerable variation as to when each child achieves specific developmental milestones. Variation between children can arise from the particular language(s) being learned, the amount and nature of the language input, and individual factors that children bring to the task of language learning. Some children develop language a little faster and others a little slower, but certain skills are foundational to others and thus tend to develop first. To understand why speech perception and comprehension in bilingual children happens the way it does, it is important to first consider the process and timing of first language acquisition more generally. The next section will trace, in broad strokes, how it takes place, with brief insights into language input to children, speech and phonetic perception, word learning and comprehension, and grammatical and pragmatic development. Then we will discuss some of the factors that impact speech perception and comprehension in bilingual children – issues that will be taken up in more depth in later chapters. Given the focus of this book on language listening, we will not discuss the development of children's language production. However, readers can keep in mind that, in general, the ability to perceive or comprehend an aspect of a language generally develops before the ability to produce it.

The Process of First Language Acquisition

Adults don't speak to children the way they speak to other adults. In most cultures, adults use a special speech register called infant or child-directed speech. Compared to adult-directed speech, it is more melodic and emotional, with greater pitch changes, and shorter sentences (Soderstrom 2007). Even very young children tend to prefer this type of speech (Cooper and Aslin 1990), and older children learn better from it than from adult-directed speech (Graf Estes and Hurley 2013). Children vary significantly in the quantity and quality of input they hear (Hart and Risley 1995), and this plays a

profound role in their language acquisition. Vocabulary grows more quickly in children who hear more words (Weisleder and Fernald 2013) and in those who hear more diverse words (Rowe 2012).

Young children need basic speech perception capabilities to begin to process the input that they encounter. Indeed, newborn infants are far from blank slates: they are born with speech processing capacities and learning mechanisms that will support them in becoming proficient language listeners. Newborn infants prefer speech sounds over non-speech sounds, which helps them focus their attention on the information that is most important for language learning (Vouloumanos and Werker 2007). Infants pay special attention to the rhythmic and melodic characteristics of voices (Langus, Mehler, and Nespor 2016), and these sensitivities help them acquire language (François et al. 2017).

Perceptual sensitivities are particularly important for young infants, because unlike adults, they cannot rely on other contextual information to break into spoken language since they are still learning about the nature of the people and objects around them. However, perceptual information available in the speech stream helps them to begin to decode the speech signal. For example, perceptual abilities can help infants differentiate the parts of speech. Function words such as propositions and articles ("on", "the") tend to be shorter and less emphasized than content words such as nouns and verbs ("running", "bottle"), and it has been shown that newborn infants can indeed discriminate function and content words (Shi, Werker, and Morgan 1999). Newborn infants can also use their perceptual capacities to detect speech units, such as syllables (Bijeljac-Babic, Bertoncini, and Mehler 1993), and tell apart languages with different rhythms (Nazzi, Bertoncini, and Mehler 1998). With experience, many aspects of an infant's speech perception will become tuned to their native language(s). For example, after a few months infants become better at differentiating their native language from a non-native language, but have difficulty differentiating two unfamiliar languages (Nazzi, Jusczyk, and Johnson 2000). Discriminating and differentiating their languages will be particularly important for bilingual infants, as will be seen in Chapter 8.

One important part of early speech perception is phonetic perception, which is the way that listeners perceive individual speech sounds such as /b/ or /a/. At birth, infants can discriminate most speech sound differences that are meaningful in different languages (e.g., they can tell /b/ from /a/ or /p/), and discriminate these contrasts in a way that very much resembles adults (Eimas et al. 1971). Experience with their language(s) will help them refine these categories and determine which sound differences should be attended to and which should be ignored (Kuhl et al. 2007; Werker and Tees 1984). Phonetic perception is one of the earliest-developing capacities for mature speech processing. Monolingual infants perceive phonemes in a native way by around 12 months of age, with vowel perception maturing a few months earlier than consonant perception. We will discuss how phonetic perception unfolds in bilingual infants and children in Chapter 8.

Another crucial aspect is word learning and recognition. Word learning involves, at a minimum, linking a sound sequence to what it refers to in the world. Children begin to understand the meanings of highly familiar words by the age of 6 to 9 months, such as those for body parts, foods, and mommy and daddy (Bergelson and Swingley 2012; Tincoff and Jusczyk 1999). In general, nouns tend to be learned earlier than verbs, which are learned earlier than other types of words such as prepositions and pronouns (Fenson

et al. 2007). Children's word learning begins slowly and then accelerates. Children often seem to go through a vocabulary spurt or explosion around the age of 1.5 years (Nazzi and Bertoncini 2003), when they begin to learn new words more rapidly. Around this age, they also seem to improve in their ability to represent each word's specific sounds, a point we will return to in Chapter 9 when we discuss word learning and recognition in bilingual children.

Another important aspect is grammatical development. As we saw earlier, different types of words play different roles in a sentence (e.g., verbs versus nouns) and are ordered in a specific way to convey a certain meaning (a language's syntax). Words can also have different prefixes and suffixes that depend on a language's grammar, such as the difference between "run" and "running" (a language's morphology). Basic perceptual sensitivity to language structure emerges early. Newborn infants can detect when syllables repeat in a series such as ta-lu-lu, compared to when they do not (ta-lu-fe) (Gervain et al. 2008).

These early-emerging sensitivities to language structure are based on perceptual information, but mature grammatical knowledge depends on understanding the meanings of words. For this reason, the bulk of grammatical development happens after children have begun developing their vocabularies. Thus, by the age of 16 months, English-learning children understand how the order of words affects meaning (Golinkoff et al. 1987), for example, that "Big Bird tickles Cookie Monster" relates to a different action than "Cookie Monster tickles Big Bird". Around the age of 2, Spanish-learners understand how plural words are formed, for example, that "gato" refers to one cat while "gatos" refers to two or more cats (Arias-Trejo et al. 2014). Since each language's grammar is unique and varied, children will learn different grammatical forms in different languages at different ages. We will turn to bilingual children's comprehension of grammatical and other aspects of processing at higher levels in Chapter 10.

The last aspect is pragmatic development. As we discussed with adults, pragmatic processing integrates information contained in the speech signal with everything that the listener knows about the speaker, the context, and the world. Pragmatic skills are some of the last language skills to fully develop, as children's knowledge in these domains can be limited—never mind their ability to fully integrate such knowledge as they listen to real-time speech. This said, the basics of pragmatic processing do begin to emerge in infancy. Starting by 12 months, infants seem to understand that speech is used for communication (Martin, Onishi, and Vouloumanos 2012). They also begin to understand and use gestures such as pointing (Behne et al. 2011) and an adult's eye gaze at an object (Beier and Spelke 2012). Despite these early-emerging sensitivities to pragmatic information, the development of mature pragmatic competence takes quite a while, and children continue to refine their ability to use pragmatic information well into the school years. Chapter 9 will discuss how bilingual children apply their pragmatic skills to learning new words.

Spoken Language Processing in Bilingual Children

All children, be they monolingual or bilingual, go through the same stages of first language acquisition that we have just discussed to acquire the sounds, words, grammar, and pragmatics of their language(s). Monolingual children acquire a single native language and bilingual children acquire two native languages at the same time (simultaneous

bilinguals) or begin to acquire a native language prior to beginning to learn a second language (sequential bilinguals). Some children learn three or more languages either simultaneously or sequentially, but as research on these populations is extremely limited, most of our discussion will focus on bilinguals. As we have noted, four chapters in this book will focus specifically on speech perception and comprehension in bilingual children, their speech perception (Chapter 8), word learning and recognition (Chapter 9), and their processing at higher levels (Chapter 10). Finally, Chapter 11 looks at how language perception and comprehension can be assessed in children with and without language impairment.

Because of the nature of language acquisition, there are some issues particular to bilingual children that we will encounter in several of the chapters. We have already discussed some of these issues in relation to bilingual adults. Language history and the length of exposure to each language can be important for both bilingual children and adults, keeping in mind that children's histories will be much shorter and on average much simpler than those of adults. Language proficiency will also play a role in children as it does with adults, although for children their level of proficiency will be constrained by what is possible given their developmental level. For example, we would not expect children of a certain age to master a grammatical form in their second language if they would not be ready to master it in their first language. Knowing what can be expected at a particular developmental level is why many studies of bilingual children compare them to same-aged monolinguals, although as we will see, this is not always a reasonable comparison. Reading is another factor to keep in mind, although it is important to remember that most children learn to read around ages 5–7, so influences of reading on bilingual children's language will only be relevant for school-aged children.

Some issues that we will encounter are particularly relevant to bilingual children as opposed to adults. One is the nature of input that bilingual children receive. As we have already mentioned, the amount of input monolingual children hear affects how quickly they acquire their native languages. Bilinguals hear language for the same number of hours per day as monolinguals, but their input is split between the languages, meaning that on average they will hear less input in a particular language than a monolingual. Some bilinguals will hear their two languages in relatively equal proportion. Others may hear considerably more input in one language than in the other. The amount of input that bilingual children hear in a particular language plays a very important role in both the development of that language and children's proficiency in it.

A related issue is that bilingual children may receive different types of input in each of their languages. For example, if they hear one language at home and another at school, they may encounter different types of words in each setting (see our discussion of the Complementarity Principle in Chapter 1). This issue will come up many times in both our discussion of children and adults, and will be particularly important to keep in mind for assessment of bilingual children (Chapter 11). On the flipside, for many everyday words, children will encounter and learn them in each of their languages (e.g., English "dog" and French "chien" both refer to the same kind of furry animals). As we will see when we discuss word learning in Chapter 8, bilinguals' need to map two different words to the same object affects their language learning.

A final recurring theme is the issue of competence versus performance. Competence is what children know, while performance reflects what they do in any particular language task. Performance is an especially tricky issue with infants and young children, as

they cannot answer a questionnaire, press a button, or even follow instructions. Young children may be able to follow some instructions, but are limited in terms of their attention, motivation, and cognitive resources. If we use a measure of performance that does not properly tap into children's underlying competence, we will draw very wrong conclusions about what bilingual children are able to perceive and comprehend. A related issue is that speaking is more difficult than comprehending, and bilingual children often understand much more than they can say. Researchers interested in studying bilingual children use unique methods when measuring their language competence that attempt to minimize how much performance issues get in the way. For this reason, we spend considerable time in the chapters focused on children to explain the specific methodologies used in the study of young bilinguals.

Summary

In this chapter, we have given an overview of speech perception and comprehension in bilingual adults and children when they process two or more languages, either separately or together in the form of mixed speech. We have described the general architecture of the basic components involved in spoken language processing and have examined the mechanisms present. We have also discussed the issues that are common to all processing levels in bilinguals. Many are due to the coexistence of two or more languages in the bilingual listener's mind. Finally, after giving a brief outline of how language acquisition takes place in children, we have discussed the factors that play a role in speech perception and comprehension in bilingual children.

References

Altmann, Gerry T. M., and Yuki Kamide. 1999. "Incremental interpretation at verbs: Restricting the domain of subsequent reference." *Cognition*, 73: 247–264. DOI: 10.1016/S0010-0277(99)00059-1.

Arias-Trejo, Natalia, Lisa M. Cantrell, Linda B. Smith, and Elda A. Alva Canto. 2014. "Early comprehension of the Spanish plural." *Journal of Child Language*, 4 (6): 1356–1372. DOI: 10.1017/S0305000913000615.

Behne, Tanya, Ulf Liszkowski, Malinda Carpenter, and Michael Tomasello. 2011. "Twelve-month-olds' comprehension and production of pointing." *British Journal of Developmental Psychology*, 30 (3): 359–375. DOI: 10.1111/j.2044-835X.2011.02043.x.

Beier, Jonathan S., and Elizabeth S. Spelke. 2012. "Infants' developing understanding of social gaze." *Child Development*, 83 (2): 486–496. DOI: 10.1111/j.1467-8624.2011.01702.x.

Bergelson, Elika, and Daniel Swingley. 2012. "At 6–9 months, human infants know the meanings of many common nouns." *Proceedings of the National Academy of Sciences*, 109 (9): 3253–3258. DOI: 10.1073/pnas.1113380109.

Bijeljac-Babic, Ranka, Josiane Bertoncini, and Jacques Mehler. 1993. "How do 4-day-old infants categorize multisyllabic utterances?" *Developmental Psychology*, 29: 711. DOI: 10.1037//0012-1649.29.4.711

Boland, Julie E., Michael K. Tanenhaus, and Susan M. Garnsey. 1990. "Evidence for the immediate use of verb control information in sentence processing." *Journal of Memory and Language*, 29 (4): 413–432.

Cooper, Robin Panneton, and Richard N. Aslin. 1990. "Preference for infant-directed speech in the first month after birth." *Child Development*, 61 (5), 1584–1595. DOI: 10.2307/1130766.

Eimas, Peter D., Einar R. Siqueland, Peter W. Jusczyk, and James Vigorito. 1971. "Speech perception in infants." *Science*, 171 (3968): 303–306. DOI: 10.1126/science.171.3968.303.

Fenson, Larry, Virginia A. Marchman, Donna J. Thal, Philip S. Dale, J. Steven Reznick, and Elizabeth Bates. 2007. *MacArthur-Bates Communicative Development Inventories*, 2nd edn. Baltimore, MD: Brookes.

François, Clément, Maria Teixidó, Sylvain Takerkart, Thaïs Agut, Laura Bosch, and Antoni Rodriguez-Fornells. 2017. "Enhanced neonatal brain responses to sung streams predict vocabulary outcomes by age 18 months." *Scientific Reports*, 7: 12451. DOI: 10.1038/s41598-017-12798-2.

Frazier, Lyn. 1987. "Sentence processing: A tutorial review." In *Attention and Performance 12: The Psychology of Reading*, edited by Max Coltheart, 559–586. Hove, UK: Lawrence Erlbaum Associates.

Friederici, Angela, D. 2002. "Toward a neural basis of auditory sentence processing." *Trends in Cognitive Sciences*, 6 (2): 78–84.

Gervain, Judit, Francesco Macagno, Silvia Cogoi, Marcela Peña, and Jacques Mehler. 2008. "The neonate brain detects speech structure." *Proceedings of the National Academy of Sciences*, 105 (37): 14222. DOI: 10.1073/pnas.0806530105.

Golinkoff, Roberta Michnick, Kathryn Hirsh-Pasek, Kathleen M. Cauley, and Laura Gordon. 1987. "The eyes have it: Lexical and syntactic comprehension in a new paradigm." *Journal of Child Language*, 14 (1): 23–45. DOI: 10.1017/S030500090001271X.

Graf Estes, Katharine, and Karinna Hurley. 2013. "Infant-directed prosody helps infants map sounds to meanings." *Infancy*, 18 (5): 797–824. DOI: 10.1111/infa.12006.

Grosjean, François. 1983. "How long is the sentence? Prediction and prosody in the on-line processing of language." *Linguistics*, 21: 501–529.

Grosjean, François. 2013. "Speech perception and comprehension." In *The Psycholinguistics of Bilingualism*, edited by Grosjean, François and Ping Li, 29–49. Malden, MA: Wiley-Blackwell.

Harley, Trevor A. 2014. *The Psychology of Language: From Data to Theory*. London: Psychology Press.

Hart, Betty, and Todd R. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes.

Kuhl, Patricia K., Barbara T. Conboy, Sharon Coffey-Corina, Denise Padden, Maritza Rivera-Gaxiola, and Tobey Nelson. 2007. "Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-E)." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363 (1493): 979–1000. DOI: 10.1098/rstb.2007.2154.

Langus, Alan, Jacques Mehler, and Marina Nespor. 2016. "Rhythm in language acquisition." *Neuroscience and Biobehavioral Reviews*. Published electronically December 16, 2016. DOI: 10.1016/j.neubiorev.2016.12.012.

Marslen-Wilson, William. 1975. "Sentence perception as an interactive parallel process." *Science*, 189 (4198): 226–228.

Martin, Alia, Kristine H. Onishi, and Athena Vouloumanos. 2012. "Understanding the abstract role of speech in communication at 12 months." *Cognition*, 123 (1): 50–60. DOI: 10.1016/j.cognition.2011.12.003.

McClelland, James L., and Jeffrey L. Elman. 1986. "The TRACE model of speech perception." *Cognitive Psychology*, 18 (1): 1– 86.

Nazzi, Thierry, and Josiane Bertoncini. 2003. "Before and after the vocabulary spurt: Two modes of word acquisition?" *Developmental Science*, 6 (2): 136–142. DOI: 10.1111/1467-7687.00263.

Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler. 1998. "Language discrimination by newborns: Toward an understanding of the role of rhythm." *Journal of Experimental Psychology: Human Perception and Performance*, 24 (3): 756–766. DOI: 10.1037/0096-1523.24.3.756.

Nazzi, Thierry, Peter W. Jusczyk, and Elizabeth K. Johnson. 2000. "Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity." *Journal of Memory and Language*, 43 (1): 1–19. DOI: 10.1006/jmla.2000.2698.

Norris, Dennis. 1994. "Shortlist: A connectionist model of continuous speech recognition." *Cognition*, 52: 189–234.

Rowe, M. L. 2012. "A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development." *Child Development*, 83 (5): 1762–1774. DOI: 10.1111/j.1467-8624.2012.01805.x.

Sedivy, Julie C., Michael K. Tanenhaus, Craig G. Chambers, and Gregory N. Carlson. 1999. "Achieving incremental semantic interpretation through contextual representation." *Cognition*, 71 (2): 109–147. DOI: 10.1016/S0010-0277(99)00025-6.

Shi, Rushen, Janet F. Werker, and James L. Morgan. 1999. "Newborn infants' sensitivity to perceptual cues to lexical and grammatical words." *Cognition*, 72 (2): 11–21. DOI: 10.1016/S0010-0277(99)00047-5.

Soderstrom, Melanie. 2007. "Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants." *Developmental Review*, 27 (4): 501–532. DOI: 10.1016/j.dr.2007.06.002.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathy M. Eberhard, and Julie C. Sedivy. 1995. "Integration of visual and linguistic information in spoken language comprehension". *Science*, 268 (5217): 1632–1634. DOI: 10.1126/science.7777863.

Tincoff, Ruth, and Peter W. Jusczyk. 1999. "Some beginnings of word comprehension in 6-month-olds." *Psychological Science*, 10 (2): 172–175. DOI: 10.2307/40063400.

Vouloumanos, Athena, and Janet F. Werker. 2007. "Listening to language at birth: Evidence for a bias for speech in neonates." *Developmental Science*, 10 (2): 159–164. DOI: 10.1111/j.1467-7687.2007.00549.x.

Weisleder, Adriana, and Anne Fernald. 2013. "Talking to children matters: Early language experience strengthens processing and builds vocabulary." *Psychological Science*, 24 (1): 2143–52. DOI: 10.1177/0956797613488145.

Werker, Janet F., and Richard C. Tees. 1984. "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life." *Infant Behavior and Development*, 7 (1): 49–63. DOI: 10.1016/S0163-6383(84)80022-3.