

## 9 The role of the childhood environment for language dominance: A study of adult simultaneous bilingual speakers of German and French

---

*Tanja Kupisch and Joost van de Weijer*

### 9.1 Introduction

The search for the true balanced bilingual depicted in some of the literature on bilingualism is elusive. The notion of balanced bilingualism is an ideal one, which is largely an artifact of a theoretical perspective which takes the monolingual as its point of reference.

(Romaine, 1989: 18)

It has long been noted that most bilinguals, children as well as adults, are not equally proficient in their two languages, having a *stronger* and a *weaker* language. However, it is still unclear what exactly the terms *stronger* and *weaker* mean, and which factors will contribute to these differences. In this chapter, we will be concerned with adult bilinguals. We use the terms *stronger* and *dominant language* synonymously when referring to the language in which a bilingual speaker is more proficient. For the lack of a better alternative – and in keeping with much of the literature – we define proficiency in terms of “resemblance to monolingual norms,” but we will also discuss the problems in doing so.

Considerable previous work on adult bilinguals has been concerned with *heritage language acquisition*. A heritage language is defined as a language that is acquired in the home context and does not correspond to the language of the larger national society, thus being a minority language (e.g., Rothman, 2009; Benmamoun, Montrul, & Polinsky, 2013). Although the heritage language is acquired from birth, heritage speakers (HSs) are often noticeably different from monolingual speakers of the same language. In fact, the acquired proficiency in a heritage language can be so different from monolingual

Many thanks to Dagmar Barton, Virginia Mueller Gathercole, Jason Rothman, and Jeanine Treffers-Daller, as well as two anonymous reviewers for commenting on previous versions of this chapter.

baselines that some researchers have used the term “incomplete acquisition” in this context (Montrul, 2008; Polinsky, 1997), in analogy to second language learners whose development fossilizes before the target grammar has been fully acquired. In the present chapter, we assume that bilingual and monolingual grammars are different, but for various reasons we do not mean to claim that HS grammars are or become “incomplete,” come what may.

As mentioned earlier, HSs are often compared to monolingual speakers, but their input experience can be very different. One reason why the input of HSs may be qualitatively different from that of monolinguals is that some phenomena may have been subject to attrition in the grammars of first generation speakers, who constitute the linguistic model for second generation HSs (for further discussion, see Benmamoun et al., 2013; Kupisch, 2013; Pascual y Cabo & Rothman, 2012). Another reason why this may happen is that linguistic properties may be subject to regional variation or pertain to a specific register. For example, Brazilian Portuguese HSs lack knowledge of inflected infinitives, thus differing from monolingual speakers of Brazilian Portuguese who have had Portuguese as their language of instruction. However, as Pires and Rothman (2009) show, the inflected infinitive is not part of spoken Brazilian Portuguese. Therefore, HSs with no formal instruction in Portuguese have no access to this structure in their (target language) input. By contrast, European Portuguese HSs acquire inflected infinitives because they are part of their input. Thus, insufficiency of input from a standard dialect can affect acquisition. HSs may show deviances from the standard variety because the varieties they were exposed to have been subject to syntactic changes. Monolingual (or bilingual) speakers will learn the respective properties of the standard varieties at school, but if HSs do not receive formal training in their heritage language, such properties will not be part of their linguistic experience.

With these concerns in mind, using the terms *dominant* and *more proficient* synonymously and defining them in terms of resemblance to monolingual norms may seem controversial (see also Montrul, this volume, for arguments in favor of distinguishing dominance from proficiency). Nevertheless, it is a useful starting point when comparing speakers. For instance, it may turn out to be the case that bilinguals, despite having less input, meet monolingual norms in some linguistic properties but not in others.

Previous research comparing bilinguals has shown that majority language bilingual speakers outperform minority language (i.e., heritage) speakers in various grammatical domains (Bianchi, 2013; Kupisch, 2012, 2014; Kupisch, Akpınar, & Stöhr, 2013; Stöhr, Akpınar, Bianchi, & Kupisch, 2012). All these studies compared bilinguals with the same language combination but living in different countries (e.g., German–Italian speakers in Italy vs. Germany). However, they did not compare the bilinguals’ two languages to each other,

thus leaving open how the HSs would have performed on comparable or equivalent phenomena in their stronger language.

Studies in which the acquisition of the heritage or weaker language grammar was deemed incomplete or deviant from some monolingual norm also tend to leave open the sources for this incompleteness or deviance. These may include quantitative and qualitative differences in language input during childhood, or different frequency of usage during adulthood, different ages of onset of the other first language, the prestige of the weaker language in the society, the parents' SES, access to input that is different from the monolingual norm due to L1 attrition (e.g., Pascual y Cabo & Rothman, 2012; Rothman, 2009; Sorace, 2004), or a combination thereof. Another question that has not been systematically addressed with respect to simultaneous bilinguals is whether dominance is still variable during adulthood, for example, when bilinguals change their country of residence.

In this chapter, we investigate language dominance in 20 adult simultaneous bilinguals (German–French). These bilinguals differ in terms of the country of their childhood environment (Germany vs. France), which makes it possible to assess the impact of the childhood environment. Those who grew up in France moved to Germany at different ages during adulthood and continued to use French to different extents, which may reveal potential effects of current language use and length of residence in Germany. We investigate whether performance in a majority language resembles the expected performance of monolingual speakers more closely than performance in a minority language. We make the preliminary assumption that the bilinguals are more proficient in the language that was spoken most frequently in the wider environment during childhood, and we refer to this language as the *stronger* or *dominant* language, while the language that was predominantly spoken at home will be considered the *weaker* language.<sup>1</sup> It needs to be seen whether this assumption can be maintained with respect to different linguistic skills in the participants' two languages, for example, the ability to “sound” native or the ability to fill out a written cloze test. We bear in mind, however, that this assumption is not without problems and that one should be extra-cautious with respect to properties that could be subject to attrition, diatopic, diastratic, or diaphasic (e.g., oral vs. written) variation.

The chapter is structured as follows: Section 9.2 gives some background and motivates the research questions. Section 9.3 presents the methods and Section 9.4 the results. We conclude with a discussion in Section 9.5.

<sup>1</sup> Under this assumption, the weaker language matches the definition of a heritage language, as used by Montrul (2008), Polinsky (1997), Rothman (2009), and others, except that its acquisition need not be “incomplete.”

## 9.2 Previous research on defining and measuring language dominance and open questions

### 9.2.1 Measures of language dominance

There are at least two important issues that emerge from previous accounts of language dominance in the literature. One is that the term *language dominance* is often used without providing any measurement of the subjects' knowledge of either language, thereby leaving it unclear what language dominance means in linguistic terms (Treffers-Daller, 2011: 147). A second issue is that we cannot infer overall language expertise from global dominance tests, since these tests do not take into account how the languages are distributed over domains (Birdsong, this volume; Grosjean, 2010: 35, and this volume).

Language dominance in children is often measured on the basis of mean length of utterance (MLU) (Yip & Matthews, 2007) and the directionality of mixing and code-switching (Berman, 1979; Petersen, 1988). If several different measures are used, they normally identify the same language as being dominant, but some measures yield a stronger difference between both languages than others. For example, in a study by Genesee, Nicoladis, and Paradis (1995), language choice with parents, MMU (number of multimorphemic utterances), and number of different word types brought out sharper contrasts between the languages than MLU. In Kupisch (2007), the amount of different word types (measured incrementally over a period of several months) revealed sharper contrasts than MLU, Upper Bound (longest word per recording), and number of utterances in a given time.

While methods for establishing language dominance in simultaneous bilingual children have been adopted and critically discussed for more than a decade (e.g., De Houwer, 1990; Genesee et al., 1995), dominance in simultaneous bilingual *adults* has been discussed much less. A number of studies on language dominance in adults are concerned with successive bilinguals or monolinguals who have become near-native or native-like later in life (e.g., Dunn & Fox Tree, 2009; Flege, Munro, & MacKay, 1995). In these studies, the term *bilingual* is used regardless of the age at which exposure to a second language began. In this chapter, we will use the term *second language learners* (L2ers) if the second language was acquired after the age of 4;0.

Many studies on bilingual adults are concerned with HSs, and for these it is sometimes simply assumed (i) that the dominant language of the environment turns into the stronger language and (ii) that proficiency in the dominant language mirrors monolingual proficiency. This is perhaps justifiable in the case of heritage speakers at the lower end of the proficiency scale who are not fluent speakers of the language (see, e.g., Benmamoun et al., 2013). However, for fluent bilinguals at the higher end of the proficiency scale, it is harder to determine which of their languages is stronger. More subtle procedures to determine dominance are

therefore needed. Unlike for children, MLU and MMU are not suitable methods here, because one can assume that they are capable of producing long words and complex utterances in both languages. Language mixing is not a reliable predictor either, because fluent bilinguals are generally good at controlling their language choice and adapting it to the situation.

In adults, language dominance is assessed either through biographical reports, for example, length of exposure and self-assessment (e.g., Dunn & Fox Tree, 2009; Flege, MacKay, & Piske, 2002), or through language-based measures, for example, written cloze tests (e.g., Abrahamsson & Hyltenstam, 2009; Bianchi, 2013; Kupisch, 2012; Montrul & Ionin, 2010), lexical richness (e.g., Treffers-Daller, 2011: 148, this volume), vocabulary tests (e.g., Montrul & Ionin, 2010), accent assessment (Flege et al., 2002), mean sentence duration (Flege et al., 2002), or translation tasks (Flege et al., 2002). As a measure for bilinguals at the lower end of the proficiency scale, Polinsky (2008b) has looked at speech rate.

The validity of self-assessment has been criticized, because ratings align more closely with language attitudes than with language proficiency (Hakuta & D'Andrea, 1992). At the same time, there are studies suggesting that self-assessment does mirror relative proficiency (e.g., Dunn & Fox Tree, 2009). Morphosyntactic proficiency measures are rarely used, perhaps owing to the fact that the complexity of these domains is not always comparable across languages (Daller, 2011). The cloze test is commonly seen as a valid tool (Tremblay, 2011), but a justified question is whether it yields representative results for heritage speakers: Even if HSs are literate in their two languages, they tend to have less written input in their minority language, and their written abilities may not reflect their abilities in other aspects of language. Vice versa, adult L2 learners may develop excellent writing skills despite late exposure to a language, and without being fluent speakers of this language.

In this study, we compare results from a written cloze test, an accent rating study, and an elicitation task of grammatical gender to see which of the tests bring out contrasts between the two languages of bilinguals who are very fluent in both of their languages.

### 9.2.2 *Variables determining language dominance*

In simultaneous and successive bilingual children, both age of onset of acquisition (AoO) and input quantity have been found to influence the language acquisition *process* (e.g., Unsworth et al., 2014). With respect to some components of the language, for example, accent, idioms, and proverbs, AoO seems to determine acquisition *outcomes* as well (Abrahamsson & Hyltenstam, 2009). By contrast, acquisition outcomes in syntax appear to be less affected by AoO, at least as long as acquisition takes place during early childhood (age 0–7). For example, German–Turkish adults perform on a par with monolinguals when

judging definiteness violations in existential constructions in both languages, regardless of whether they were exposed to the two languages between 0 and 3 years or between 4 and 7 years (Kupisch, Belikova, Özçelik, Stangen, & White, in press). In the present study, age of acquisition is kept constant, so that potential differences in ultimate attainment cannot be due to AoO.

Only recently, an increasing number of studies have tried to assess input quantity and determine its role in bilingual development. Unsworth and colleagues (Unsworth, 2013b; Unsworth et al., 2014) proposed to distinguish *current exposure* and *cumulative length of exposure* (CLoE) to the two languages. Current exposure corresponds to an overall percentage of exposure to the languages investigated per week during the child's waking hours; CLoE corresponds to exposure *over time* and is measured as a proportion of target language (TL) exposure for each one-year period in a child's life, which can then be added, resulting in a total amount of TL exposure counted in years. Unsworth et al. (2014) show noticeable variation in both input measures for 2L1 and L2 children, both within and across these groups, and the two measures were significant predictor variables for the acquisition of grammatical gender in Dutch and Greek.

Other studies have looked more closely at *exposure to language(s) at home*. Gathercole and Thomas (2009) distinguished English–Welsh bilingual children who at home spoke only English, only Welsh, or both. Amount of exposure to the languages at home had an effect on the children's level of Welsh, the minority language, but not on English, the majority language. Thus, exposure at home seems to mainly have an impact on the language to which the child is generally less exposed.

In contrast to these studies, Pirvulescu, Pérez-Leroux, Roberge, Strik, & Thomas (2014) found no differences in the realization of object clitics in French between balanced bilinguals and English-dominant children with French as their weaker language. In their study, dominance was measured based on parental assessments of the children's language abilities and their report of current language exposure and use (inside and outside the home).

Several researchers have stressed the importance of quality-oriented input factors, or *richness of the environment* for the development of a minority language (e.g., Jia, Aaronson, & Wu, 2002; Jia & Aaronson, 2003; Paradis, 2011), the idea being that a learner should be exposed to a variety of different speakers. This may be achieved by holidays and family visits to countries where the minority language is dominant (De Houwer, 2009) or through activities in the minority languages outside school, such as reading, watching TV, meeting native-speaking friends, etc. (Gutiérrez-Clellen & Kreiter, 2003). Inspired by these ideas, Ågren, Granfeldt, and Thomas (2014) developed a detailed input profile, adding qualitative and quantitative measures (e.g., relative use of the languages at home, vacations, reading habits, access to TV, etc.). They found that the profile correlated with the linguistic development of simultaneous

bilinguals, but not with that of child L2 learners, and that the impact of the input may depend on the characteristics of the linguistic structures under investigation. In other words, some morphosyntactic structures might need more input than others in order to develop in a target-like fashion. The profiles do not provide information about scores on individual variables, though; it therefore remains open which of them are the best predictors of language proficiency.

In summary, most studies report that opportunities to use the language are especially crucial for the minority language (see also La Morgia, this volume). The reconstruction of quantitative and qualitative input conditions is a challenge, in adults even more than in children because the time span to be reconstructed is much longer. For the adult population in this study, we will attempt to assess input conditions during childhood and adolescence as well as adulthood. Given that our biographical data is based on self-assessment and reconstruction, it goes without saying that these data must be treated as subjective rather than objective.

### 9.2.3 *Shifting language dominance in children and adults*

Except for cases in which both parents speak the minority language at home, simultaneous bilingual children are often quite balanced during the first three years of their lives, or more proficient in the national language of the childhood environment (see, e.g., Kupisch, 2007, for relevant case studies), but there are also exceptions where the minority language is dominant (see, e.g., Müller, Kupisch, Schmitz, & Cantone, 2011).

There are many reports of shifting language dominance in the literature. For example, the German–English child, Hildegard (Leopold, 1978) spent the early years of her childhood in the United States and was more proficient in English. Around age six, the family moved to Germany and after only four weeks, Hildegard spoke English with a German accent and had problems constructing simple English utterances. After moving back to the United States, English took over again. Similarly, Shelly, the English–Hebrew child studied in Berman (1979) was more proficient in Hebrew during the first two-and-a-half years of her life that she spent in Israel. English took over when she moved to the United States where she attended an English-speaking kindergarten. On her return to Israel, she seemed to have problems communicating in Hebrew, but after only four months she communicated in Hebrew again. Burling's (1959) son Stephen, who had learnt English during his first 16 months, and then moved to the Garo mountains in North-Eastern India with his parents, developed Garo faster than English because most people around him (except his mother) used Garo. Lukas, the German–Italian bilingual boy studied in Kupisch (2007), was a balanced bilingual until the age of 3;0 years, but developed a strong dominance in German after starting to attend a German kindergarten (Müller et al.,

2011). From today's perspective, one would probably concede that increasing or decreasing frequency of language use may not necessarily imply shifting dominance, as it may also be a question of a child's insufficient processing capacities for inhibition or activation of the grammars. Nevertheless, the studies summarized earlier indicate that the childhood environment (outside home) is a strong determinant of language dominance.

Shifting language dominance has also been reported for L2 learners, i.e., speakers who grew up as monolinguals and immigrated to a foreign country after having completed first language acquisition. A good example is Iverson's (2011) case study of a L1 Spanish speaker from Chile who had been living in Brazil for 30 years, showing the speaker's gradual shift from L1 Spanish towards L2 Brazilian Portuguese dominance. Shifting dominance is also common in early successive bilinguals who speak only a minority language at home before attending a kindergarten or school where only the majority language is spoken, as is typical for the German–Turkish community in Germany (Kupisch et al., in press; Rothweiler, 2007). There are no studies, so far, investigating whether language dominance can shift during adulthood when *simultaneous* bilinguals move to a country where their heritage language is spoken. One might speculate that their language balance is more easily affected as compared to successive bilinguals: successive bilinguals have the opportunity to stabilize their knowledge in one language before they are confronted with the next language.

Summarizing, previous literature leaves open the following questions:

- (1) Which measures of language dominance are appropriate for adult bilinguals who are fluent in both languages?
  - (a) Which measures are stronger indicators of dominance and which measures are weaker indicators?
  - (b) Do all measures point in the same direction? In other words: is it always the same language that is dominant?
- (2) Which variables predict proficiency in adult bilinguals?
- (3) Is dominance in simultaneous bilinguals still variable during adulthood?

### **9.3 Four analyses of proficiency in German–French bilinguals**

#### *9.3.1 Participants*

The data were collected as part of the project E11 at the Research Centre of Multilingualism in Hamburg.<sup>2</sup> The participants in the present study are 20

<sup>2</sup> E11 was funded by the German Science Foundation (DFG) between 2009 and 2011 through a grant to the first author (see Kupisch, Barton, Bianchi, & Stangen, 2012, for further information). It focused on proficient adult simultaneous and successive bilinguals speaking German and French or German and Italian.



German–French bilinguals (2L1ers) who had grown up in bi-national families. All reported that their parents followed the one person, one language principle, i.e., each of them used a different language with their child, either German or French.

The speakers differed in terms of their predominant country of residence during childhood and adolescence. Ten speakers spent the first 19 years<sup>3</sup> of their lives mostly in Germany (2L1G). We will refer to them as heritage speakers of French. Except for three speakers (D7, D14, D19), they traveled to French-speaking countries only for vacations. Eight of them had attended a French school (*Lycée Français de Hambourg*). According to self-reports, all 2L1Gs used more German than French at the time of data collection, and seven used French less than 10% of the time, while they used German 90% of the time. All speakers but one felt more at ease when speaking German than when speaking French. At the time of testing, these speakers were between 20 and 42 years old (mean age 27 years).

The other ten bilinguals spent the first 19 years of their lives mostly in a French-speaking country (2L1F), but moved to Germany as adults. We will refer to them as heritage speakers of German. Before the age of 19 years, most of them went to Germany only for vacations. There were three exceptions (F3, F4, F10), who had spent more time in Germany. Only one bilingual had attended a German school; others had attended a French school and had some formal instruction in German. Even if not all speakers had been schooled in German, they were all literate and used to reading German on a daily basis. At the time of testing, speakers in this group were between 24 and 40 years old (average: 33 years), and had been living in Germany for between six months and 20 years. Although most of these bilinguals continued to use French on a daily basis, six reported that use of German prevailed, and three used French less than 10% of the time, with German being used most of the time. This group showed more mixed language preference patterns than the 2L1G group.

An overview of the participants is provided in Table 9.1. Note that length of residence (LoR) in countries that were neither German- nor French-speaking has been treated as neutral and was subtracted from the total of years.<sup>4</sup> It is worth highlighting two major differences between the two groups. First, most speakers in the 2L1G group, but not in the 2L1F group, had received formal instruction in the heritage language from a relatively early age. Second, most

<sup>3</sup> We used 19 years as a cut-off point because this is when most participants had finished school and began to be more mobile. Using an earlier cut-off point, for example, age 10, would have resulted in a slightly different grouping (D14 and F10 would have been in different groups).

<sup>4</sup> It is debatable whether this is justified. For example, even if a speaker has spent time in Russia, it is likely that she continued to speak one or both of her native languages, but since we have no information to what extent this holds for each individual speaker, we preferred to deduct these periods.

Table 9.1 *Participants*

Group	Participant	Age	LoR in Germany before 19;0	LoR in Germany after 19;0	Relative use of French (%) vs. German	Type of school	Language of mother	Language preference
2L1G	D1	22	19	3	12/1	F	F	G
	D2	21	19	2	11.5/0.5	F	F	G
	D4	42	19	22.5	12/1	F	G	G
	D7	20	14	0	18/6	F	G	G
	D9	20	19	1	24/0.5	G	G	G
	D10	20	19	1	11/3	F	G	G
	D11	37	19	17.75	15/1	G	F	G
	D14	32	10	12	15/1	F	F	G/F
	D15	34	12	13.75	12/0	F	F	G
2L1F	D16	20	19	0	6/6	F	G	G
	F1	24	0	0.5	4/3	G	F	G/F
	F2	36	0	5.5	16/8	F	G	F
	F3	41	3	12.25	12/0.5	F	F	G
	F4	34	1	14	4/4	F	G	G/F
	F5	35	0	14	8/2	F	G	G
	F6	33	0	5	6/7	F	G	F
	F7	39	0	20	10/5	F	G	G/F
	F8	40	0	11	6/6	F	G	G/F
	F9	26	0	3.5	24/1	F	G	F
	F10	24	9	1	2/0	G/F	F	G/F

speakers in the 2L1F group had lived and worked in Germany for a relatively long time, while many of the 2L1G group had never lived nor worked in a French-speaking country.

For the remainder of the analysis, we will use the division into these two groups, i.e., 2L1G and 2L1F, to see whether it can be confirmed that the 2L1ers are more proficient in the language of their childhood environment.

### 9.3.2 *Methods*

The two bilingual groups' performance in their two languages was compared with respect to a cloze test, an accent rating experiment, and an elicitation task targeting grammatical gender.

Note that the same participants were also tested with respect to adjective placement and article use in generic noun phrases. However, these two phenomena have a different degree of complexity in the two languages and are subject to variation, which is why they are not ideal when comparing the two languages.<sup>5</sup> Even in terms of grammatical gender, the two languages are not perfectly comparable, as we will discuss further in the following text.

*Cloze test.* The French cloze test was kindly made available to use by Annie Tremblay (Tremblay, 2011). It was a text of one page, containing 45 gaps, which required the insertion of lexical ( $n = 20$ ) and function words ( $n = 25$ ). The German cloze test was created for the research project E11 (see above) and was comparable to the French one in terms of length, difficulty, and number of gaps.<sup>6</sup> Participants were given 15 minutes to fill it out. The analysis was based on the number of appropriately filled blanks. In many cases, there was more than one acceptable solution.

*Foreign accent rating.* We designed two separate experiments, one for German and one for French, to test the speakers' foreign accent in French and German. Each experiment contained speech samples of 30 speakers, of which 20 were the 2L1ers introduced earlier. There were also five samples from monolingual speakers and five samples from L2 speakers.<sup>7</sup> The criterion for being monolingual was that only one language was used at home during childhood and this was the only language of instruction at school.

<sup>5</sup> Adjectives have only one possible position in German, whereas they have two possible positions in French, depending on contextual factors. Vice versa, article insertion in French is relatively easy to resolve because bare nominals are generally absent in French, while in German the use of articles depends on semantic factors, and displays regional variation.

<sup>6</sup> In both the German and the French cloze test, the topic was similar (environmental issues). Monolinguals taking the test tended to fill the gaps appropriately between 90 and 100% of the time.

<sup>7</sup> The 2L1 and the L2 data are part of the HABLA-corpus (Hamburg Adult Bilingual Language), which can be accessed at: [www.corpora.uni-hamburg.de/sfb538/en\\_e11\\_habla.html](http://www.corpora.uni-hamburg.de/sfb538/en_e11_habla.html).

The speech samples were extracted from naturalistic interviews of between 20 and 30 minutes duration. Two passages from each speaker were selected, one lasting 15 seconds and one lasting 30 seconds ( $\pm 10\%$ ). These passages contained a description of a book or a movie but no information revealing the speakers' cultural background, no morphosyntactic errors, few or no hesitations, and no long pauses.

In each of the two experiments, samples were rated by monolingual speakers of the respective language. The raters for the French experiment ( $n = 23$ ) were recruited in Paris (France), and the raters for the German experiment ( $n = 21$ ) were recruited in Hamburg (Germany). Stimuli were presented over headphones. There was a training period with two examples, one from a L1 speaker, and one from a L2 speaker with a clearly identifiable foreign accent. These two speakers were not part of the experiment. Prior to testing, it was emphasized that a regional accent should be counted as native.

The method was inspired by De Leeuw, Schmid, and Mennen (2010) with some modifications. The raters were asked to judge samples for foreign accent in four steps, but we will only report the results of the first part of this experiment, where the raters heard a speaker for 15 seconds and were to decide whether this speaker sounded foreign or native.<sup>8</sup> The experiment took between 30 and 40 minutes and there were two semi-randomized test versions for each language.

*Gender assignment and gender agreement*<sup>9</sup> German has three genders: masculine (M), feminine (F) and neuter (N), while French has only two genders: masculine and feminine. In both German and French, gender assignment follows semantic, morphological, and phonological rules. While nouns have gender as an inherent property, elements other than nouns receive gender through agreement with the head. Both German and French mark gender on determiners and adjectives, cf. German: *ein<sub>M</sub> grauer<sub>M</sub> Hund<sub>M</sub>* vs. *eine<sub>F</sub> graue<sub>F</sub> Katze<sub>F</sub>*; French: *une<sub>F</sub> vache<sub>F</sub> grise<sub>F</sub>* vs. *un<sub>M</sub> chat<sub>M</sub> gris<sub>M</sub>*.

The study of gender assignment and agreement was based on an elicited production task (EPT) in German and French. In both language versions, the participants were first presented with two objects (e.g., two pianos) that differed in one feature (e.g., color) and were asked to describe each of them using a word that describes "how the object is." The appropriate answer was a combination of two Determiner Phrases (DPs) containing an indefinite article, a

<sup>8</sup> In the second step, raters had to indicate how certain they were between certain, semi-certain, and uncertain. Third, if a rater considered a speaker to sound foreign, he was asked to explain which phonetic or phonological features he had based his decision on. Fourth, raters heard another sample from the same speaker, this time for 30 seconds. After listening to the second sample, they had the chance to revise their judgment.

<sup>9</sup> The French part of the data has been published in Kupisch et al. (2013). One of the speakers studied in that paper was excluded here because he did not take part in the accent study.

noun, and an adjective (indefinite condition). An example of a question and answer is provided in (1).

- (1) (a) Was siehst du?  
 “What do you see?”  
 (b) ein<sub>N</sub> schwarzes<sub>N</sub> Klavier<sub>F</sub> und ein<sub>N</sub> braunes<sub>N</sub> Klavier<sub>N</sub>  
 “a black piano and a brown piano”

If a participant did not use the target noun (e.g., Keyboard instead of Klavier), the experimenter provided the target noun without revealing the gender (e.g., “What you see is called ‘Klavier’”). After seeing the two objects for the first time, the participants were presented with only one of the two objects they had previously seen and were asked to indicate what is missing (definite condition). An example of question and answer is provided in (2).

- (2) (a) Was fehlt jetzt?  
 “What is missing now?”  
 (b) das<sub>N</sub> schwarze<sub>N</sub> Klavier<sub>N</sub>  
 “the black piano”

In both the German and the French test, 18 nouns were elicited. Fillers were added, making participants believe that their memory capacity was being tested. The two language versions differed in that the nouns were balanced across the three genders in German and across two genders in French. Another difference between the two tests was that the French adjectives are gender-marked when combined with indefinite and definite DPs. By contrast, German adjectives are not always marked for gender if the DP is indefinite.

In the analysis, we determined gender assignment on the basis of the article, and agreement on the basis of the article and the adjective. The indefinite condition (1b) was used to analyze gender agreement since German adjectives are not always gender-marked with definite DPs (cf. *der<sub>M</sub> schwarze Hund<sub>M</sub>* “the black dog” vs. *die<sub>F</sub> schwarze Katze<sub>F</sub>* “the black cat”). By contrast, the analysis of gender assignment was based on the definite condition (1b) because the German masculine and neuter indefinite articles are homophonous in the nominative case (cf. *ein Hund<sub>M</sub>* “a dog” vs. *ein Schwein<sub>N</sub>* “a pig”).<sup>10</sup> In French, gender assignment and agreement were determined on the indefinite condition.

*Language background questionnaire* Participants were asked to fill out a questionnaire, including the variables introduced earlier (i.e., AoO for both languages, age, languages spoken at home, predominant place of residence before the age of 19). In addition, the questionnaire also assessed *time spent in the*

<sup>10</sup> Even though the question elicited accusative case, participants sometimes answered with a DP in the nominative case.

*heritage country after age 19* and *current relative use of both languages*, which can be considered quantitative variables, because they reflect the relative frequency with which each language is heard or spoken. The *type of school* (French or German) could be considered as a qualitative measure, because a child who grows up in Germany but attends a French school will receive input, spoken and written, from more different sources than a child who grows up in Germany but attends a German school. The language of the mother (French or German) could also be considered as a qualitative factor, because it is sometimes (perhaps unjustifiably) assumed that children have a closer relation to their mothers than to their fathers. The idea would then be that it may be an advantage for the development of the heritage language if it is the mother who speaks it. In addition, we assessed language preference, which may be considered to be an affective factor (note, however, that we did not explicitly ask for language preference but for the language that participants felt more comfortable using).

*Analysis* In the results section that follows we present descriptive statistics, concentrating on differences between the two groups. Furthermore, for each of the variables described earlier, we performed separate multilevel logistic regression analyses with the background variables as predictors and the outcome on the test (correct/incorrect, native/non-native accent) as outcome variables. This analysis was complicated because the sample was unbalanced (for instance, there were no participants in the 2L1G group who preferred French only), and some predictors correlated considerably with one another (for instance, LoR after 19 strongly correlated with age of the participant). These factors rendered the output of the regression analyses rather unstable, making it difficult to draw clear conclusions. Particularly problematic was the correlation between language preference and LoR before 19. Participants who spent their childhood in Germany tended to prefer German, while participants from the 2L1F group preferred French or had no special preference (see Table 9.1). Our approach was to first look at the predictors individually, after which we constructed a reliable model that contained one or more predictors that were significant in the initial analysis, and also remained significant in combination with other predictors. In most cases, as described in the following text, the principal predictor that explained the test scores could either be LoR before 19 *or* language preference. Since these predictors are so strongly correlated in the data set it was not possible to disentangle their individual effects. In the presentation of the results that follow we prioritize the LoR before 19 predictor.

## 9.4 Results

An overview of the results on all tests is shown in Table 9.2. In the following text we present the results of each test separately. Numbers indicate proportions

Table 9.2 *Proficiency scores, comparing all tests and the two groups*

			2L1G	2L1F
Cloze test	French version	Function words	0.74	0.88
		Lexical words	0.76	0.90
		Total	0.74	0.89
	German version	Function words	0.95	0.86
		Lexical words	0.87	0.77
		Total	0.91	0.82
Foreign accent	Speaking French		0.49	0.90
	Speaking German		0.99	0.36
Gender	French version	Agreement	1.00	1.00
		Assignment	0.98	1.00
	German version	Agreement	1.00	0.99
		Assignment	0.98	0.88

of correctly answered items for the cloze test, the gender assignment test, and the gender agreement test, or proportions rated as “native” for the foreign accent study.

#### 9.4.1 *Cloze test*

The results of the cloze test are given in Table 9.2. The total proportions suggest a difference in performance. The 2L1G group scored better than the 2L1F group on the German version of the test, and the 2L1F group scored better than the 2L1G group on the French version of the test. The difference between lexical and function words appears negligible for both groups.

The output of the regression analysis for the French test suggested that the main significant predictor was LoR before 19 ( $\beta = -.083$ ,  $SE = .028$ ,  $z = -2.995$ ,  $p = .003$ ). For the German cloze test, LoR before 19 was marginally significant ( $\beta = .048$ ,  $SE = .027$ ,  $z = 1.813$ ,  $p = .070$ ), and the native language of the mother was significant ( $\beta = -1.312$ ,  $SE = .4827$ ,  $z = -2.718$ ,  $p = .007$ ). This latter effect indicates that participants with a French mother scored higher on the German cloze test than participants with a German mother.

#### 9.4.2 *Foreign accent study*

As the proportions in Table 9.2 show, the members of the 2L1G group were nearly always perceived as native German but considerably less often as native French. The members of the 2L1F group, on the other hand, were nearly always perceived as native French and considerably less often as native German.

The results of the regression analysis for the participants speaking French showed a clear and significant effect of LoR before 19 ( $\beta = .136$ ,  $SE = .039$ ,

$z = 3.483, p < .001$ ), and a marginally significant effect of hours per day that the participant spoke German ( $\beta = .106, SE = .057, z = 1.862, p = .063$ ). In other words, participants who spoke relatively more German were perceived to be foreign-sounding when speaking French more often than participants who spoke less German per day. For the participants speaking German, only the LoR before 19 was a significant predictor ( $\beta = .421, SE = .087, z = -4.861, p < .001$ ).

#### 9.4.3 *Gender*

The results of the gender test, unlike the other tests, did not reveal striking differences between the groups. As the figures in Table 9.2 show, the overall scores were nearly 100% correct for both groups in both languages and for agreement and assignment alike. The only part of this test that distinguished the groups is that of German gender assignment. The 2L1G group did better on this part than the 2L1F group. We carried out a regression analysis on this part of the data set, but comparisons between the other parts cannot be made since either or both groups scored 100% correct. The effect of LoR before 19 was significant for gender assignment ( $\beta = .159, SE = .045, z = 3.559, p < .001$ ).

#### 9.4.4 *Summary of the results*

In sum, the results of the tests suggest that there are differences in language proficiency as measured by the cloze test and the accent test but with respect to gender assignment only in German. Our analyses point to the country where the participant spent his or her childhood as the main predictor that explains the differences in proficiency, but we also found indications that the native language of the mother may play a role, as well as the number of hours that the participants spoke German. There were no indications that school language, LoR after 19, or participant age correlated with any of the variables measured in the investigation. The only background variable that might be a predictor of language proficiency is language preference, but given the characteristics of the sample, we cannot draw any firm conclusions about its role.

### 9.5 **Discussion**

Unlike most previous work on language dominance in adult bilinguals, we selected a group of simultaneous bilingual speakers with exposure to both languages in their homes from birth. For such speakers, developmental studies found that they separate their two languages from early on, pass through the same stages as monolinguals, and make the same type of developmental errors (Meisel, 2011), even if one of their languages developed faster than the other.



Our study also differs from many previous ones on adult bilinguals in that our speakers were highly fluent in both of their languages. They were at the higher end of the proficiency scale when compared to the HSs that have, so far, figured prominently in the literature (e.g., see Benmamoun et al., 2013, for an overview).<sup>11</sup>

### 9.5.1 Measures of dominance

Our first question was which measures of language dominance are appropriate for adult bilinguals who are fluent in both languages. Table 9.2 shows that with respect to most tests the two groups, 2L1F and 2L1G, achieved different scores, and that the 2L1Fs had higher scores in French, while the 2L1G group had higher scores in German. This can be taken to imply that even highly fluent bilinguals have different levels of proficiency in the two languages, with proficiency (if different) being consistently higher in the language that corresponds to the national language of their environment during childhood and adolescence.

In general terms, our results confirm the view by Ågren et al. (2014), who suggest that in bilingual development different language components are not affected to the same extent by differences in input quantity and quality. We have shown that some components of the language are perfectly mastered despite the fact that one language may be stronger overall. Specifically, the two groups of bilinguals were at ceiling in their performance on gender agreement for both languages, suggesting that their morphosyntactic competence is native-like.

Recall that if gender was not correctly assigned, while agreement between the article and the adjective was correct, we counted this as an instance of correct agreement despite incorrect gender assignment. For example, for someone who produced *\*ein<sub>N</sub> trinkendes<sub>N</sub> Frischling* ‘a drinking young boar’ – *\*das<sub>N</sub> trinkende Frischling* ‘the young boar’ instead of *ein<sub>M</sub> trinkender<sub>M</sub> Frischling<sub>M</sub>–der<sub>M</sub> trinkende Frischling*, agreement was counted as correct, because it was consistent between the determiners and the adjective, although the target gender was masculine.<sup>12</sup> This way, we discovered that agreement tended to be consistent, although some participants had problems assigning the correct gender of some infrequent words in German, even if the suffix provided a gender

<sup>11</sup> There is a reason why less proficient HSs have figured more prominently in the literature. Studies on HSs sought to address the question why early naturalistic acquisition results in the differences exemplified by HSs. Data from more proficient HSs may not be suitable to address this question, even though highly proficient speakers also reveal interesting differences from monolinguals (e.g., Pires & Rothman, 2009).

<sup>12</sup> In cases like these, where masculine and neuter gender are indistinguishable on the indefinite article (except in the accusative case), we took the definite condition into account to determine which gender has been assigned.

cue (*-ling* is typically associated with masculine gender, but speakers might have been tempted to assign neuter gender because the picture showed a baby animal).

Differences in attainment between gender assignment and agreement (Table 9.2) may be due to the fact that agreement is rule-based while gender assignment is (at least partially) a matter of lexical learning. The lexicon is relatively more vulnerable compared to the grammatical and phonological systems, because it is more variable: new items are continuously added or the meaning of existing ones is changed. The advantage of the French over the German HSs in gender assignment may result either from the relatively higher degree of transparency of the French assignment system,<sup>13</sup> or from the fact that the 2L1G group had advantages in their French vocabulary since many of them attended a French school. The latter is less plausible, though, because we did not find any correlation between type of school and proficiency in the regression analyses.

Summarizing, all participants were highly proficient with respect to both gender assignment and agreement, although assignment seems to be relatively more difficult, as it may require familiarity with specific lexical items rather than rule-knowledge. In this respect, our study contrasts with other studies on the *development* of grammatical gender, which have shown that exposure to the languages at home, current exposure, and cumulative exposure had an effect on proficiency in gender-marking (Thomas & Gathercole, 2007; Unsworth, 2013b; Unsworth et al., 2014). The results of the present study could mean that even if their proficiency in a specific domain lags behind in development, bilinguals can catch up during later development. Note, however, that the conditions for gender assignment and agreement in French and German are not perfectly comparable to those in Dutch and Welsh, because a comparison of the rules formulated in the papers by Thomas and Gathercole (2007) and Unsworth (2013b) with those in German and French (see above as well as footnote 13) suggest that Dutch and Welsh have more opaque gender systems, which require relatively more lexical learning.

The results of the cloze test and the accent study showed clearer differences between the two groups of bilinguals, being most pronounced in the accent study (Table 9.2). Again, our participants differ in this respect from the less fluent HSs depicted in the literature, for whom it is generally claimed that “phonological competence seems to be the best preserved aspect of linguistic knowledge” (Benmamoun et al., 2013: 136). Taken together, our results

<sup>13</sup> Both German and French have phonological, morphological, and semantic rules for gender assignment. However, phonological rules are arguably more reliable in French, because there are fewer assignment rules in German and they cover a lower number of lexical items (see Tucker, Lambert, & Rigault, 1977, for French; Köpcke & Zubin, 1983, for German).

suggest that accent is the best indicator of language dominance in adult simultaneous bilinguals who are highly fluent in both languages. Finally, all measures in which we found a contrast in proficiency between the two groups showed higher proficiency in the language of the environment.

### 9.5.2 *Variables determining proficiency in adult bilinguals*

In the regression analysis only LoR during childhood and adolescence were significant predictors of proficiency during adulthood. Neither LoR during adulthood, nor age, nor use of the two languages at the time of testing, nor type of school, nor language of the mother was significant.<sup>14</sup> As we noted in the beginning, one problem in our analysis was that the sample was unbalanced and that some predictors correlated considerably with one another. For instance, LoR before 19 correlated with language preference, which means that we cannot firmly conclude whether language preference, or LoR before 19, or both together were crucial.

Again, our results are interesting in comparison to studies on developing bilinguals. First, Gathercole and Thomas (2009) found that the languages at home may affect bilinguals' two languages during development. Our results suggest that the language spoken at home is not a sufficient factor when predicting different acquisition outcomes. This is because all participants in the present study were exposed to both languages at home; nevertheless, there were differences in the acquisition outcomes between the 2L1F group and the 2L1G group.<sup>15</sup>

Second, Unsworth (2013b) and Unsworth et al. (2014) have shown that current exposure determines proficiency in developing bilinguals. By contrast, current exposure and use seemed to matter less in our sample. The only comparable effect we found was in the accent study. Here, participants who spoke more German per day were more likely to be classified as foreign-sounding in French than participants who spoke less German per day. Vice versa, however, the amount of speaking French per day did not affect the accent in German. In other words, our sample has more instances where frequent use of German goes along with an accent in French than instances where frequent use of French goes along with an accent in German. The explanation may be that there are comparatively more speakers who use German frequently (and these include all 2L1Gs – who tend to have an accent in French). Although speakers who use French frequently (some of the

<sup>14</sup> We found that participants with a French mother scored higher on the German cloze test than participants with a German mother. This was unexpected, because participants with a German mother could have been expected to score higher on the German cloze test, but factors not controlled for (e.g., test-taking ability) could be responsible for this result.

<sup>15</sup> Of course, even if both languages are spoken at home, one may prevail. We have no data on the relative use of languages in our participants' homes.

2L1Fs) tend to have an accent in German, the analysis does not reach significance because there are fewer speakers of this type.

Third, we did not measure cumulative exposure directly. It is likely that most participants had more exposure (cumulatively) to the language spoken in their childhood environment. This may seem to imply that both childhood environment and cumulative exposure could have determined our results. However, some of the 2L1F speakers had more cumulative exposure to German overall (having been raised in France with two languages, then moving to Germany where exposure to German prevailed), and nonetheless they continued to be relatively more proficient in French. This would speak in favor of childhood environment being the crucial determinant.

Finally, the fact that the type of school did not influence language proficiency was unexpected, because schooling in the weaker language can be taken to imply a qualitatively richer input, including more different speakers (in the case of the 2L1G group the teachers were native speakers of French) and more exposure to written language.

### 9.5.3 *Is dominance still variable during adulthood?*

One particularity of our sample was that the 2L1F group had moved to Germany during adulthood and had been living there for periods of varying length. If the language of the environment had an impact on proficiency during adulthood, one would expect that increasing LoR in Germany leads to higher proficiency in German and to lower proficiency in French. Since the regression analysis did not yield any significant correlation between LoR after 19 and proficiency, we conclude that LoR during adulthood will not affect proficiency, at least not in simultaneous bilinguals.<sup>16</sup>

## 9.6 Conclusions

Language dominance can be very subtle in early simultaneous bilinguals who had ideal conditions for acquiring both languages during childhood. However, even highly proficient bilinguals may show different abilities in their two languages. If one language is dominant, it corresponds to the language to which the bilingual had more exposure outside the home. Language dominance may not be visible with respect to all components of the language, and for the population we studied it was most visible in global pronunciation and least visible

<sup>16</sup> One point of criticism might be that we should have tested the 2L1F group individually, as the results may have been influenced by the 2L1G group (highly proficient in German, relatively lower proficiency in French). To be on the safe side, we also looked at whether LoR correlated with proficiency when excluding the 2L1G group from the sample. Again, there was no significant correlation.

in morphosyntax. For highly fluent bilinguals, global accent therefore seems to be a reliable indicator of language dominance. Of course, the question of whether a foreign accent is necessarily an indicator of proficiency depends on how one defines proficiency. Bilingual speakers may not necessarily strive for native pronunciation (unlike many L2ers). Moreover, there may be native accents that sound foreign to some monolingual speakers of a language (e.g., Miami Spanish, Quebec French). In this respect, viewing a foreign accent as an indicator of weaker language proficiency only holds when comparing to monolinguals who grew up in the same area. Finally, language dominance is unlikely to shift during adulthood, even upon changing the place of residence.