Introduction
The mother tongue – the first and, for many, only language of the early childhood – is special, differing in characteristic ways from any language we learn later in life: For one thing, the acquisition of some fundamental properties of the first language (L1) – such as intonation (Mampe, Friederici, Christophe, & Wermke, 2009; Mehler et al., 1988) – already starts before birth. At a slightly later age, a crucial part of the learning of other characteristics – such as the L1 phoneme inventory – takes place. While we do not retain any episodic memories of this early period in our lives (Tus-tin & Hayne, 2010), the linguistic knowledge that we have acquired remains accessible and lays the foundation for all additional representations and procedures that we acquire during childhood and adolescence. Also, it is only in the L1 that these early stages of learning are supported by language input that, for a large share, is tailored to the acquisition needs of children (Kuhl, 2004). As a result, the L1 is more deeply ingrained in our minds than other languages we start learning at an older age – to the extent that we may experience it as part of our identity. Citing Wilhelm von Humboldt, German writer Thomas Mann once called his mother tongue “wahre Heimat” (Mann, 2009, p. 708), a true home. The deep-rootedness of the L1 is reflected in the fact that it is the language that is most readily used in highly emotional situations, for inner speech and in mental calculations (Dewaele, 2004, 2007; Martinovic & Altarriba, 2012). Some studies have found emotional reactions to native-language stimuli to be stronger than reactions to foreign-language stimuli (Ivaz, Costa, & Duñabeitia, 2015; but cf. Ponari et al., 2015).

In spite of all that, the L1 does not remain in an uncontested position throughout the lives of all native speakers. With each additional language we learn and use, less time is left to speak our L1 and listen to it. Other circumstances, such as emigration, can have an even stronger influence on patterns of language use: In environments where the L1 is not or hardly spoken, exposure to and use of the mother tongue can be strongly reduced or even cease entirely. In such cases, a second language (L2), typically the language of the living environment, takes centre stage.

What are the consequences for the L1 when it is used less frequently in favour of an L2? We know that semantic information – which, like knowledge of facts or life events, is stored in declarative memory – decays when it has not been remembered for a long time and eventually will not be accessible any longer (Cansino, 2009). However, language as a skill is also represented in procedural memory (Ullman, 2001a). In view of this dual entrenchment of language in memory, the question arises if the
L1 can be forgotten or lost at all. Researchers of the past four decades have studied the linguistic behaviour of L1 speakers in L2-dominant settings to answer this question. Their results show that when speakers are immersed in a dominant L2, their L1 perception and production can indeed change. This process is referred to as language attrition and the speakers (potentially) affected by it as attriters.

The nature and scope of these changes have not been thoroughly investigated for all linguistic levels, though, and the factors that influence the strength of language attrition in individual speakers are still not fully understood. Also, the important role that the L2 plays in L1 attrition calls for a closer integration of attrition research into the broader context of bilingualism research. Constantly dealing with more than one language may result in language interference and competition. In the L2 of bilingual speakers, it can be difficult to investigate such effects, especially when the data come from individuals with different proficiency levels that they have reached through different learning strategies. With respect to their L1, however, attriters form a population that is more homogeneous: These speakers have acquired their native language in the same typical way, that is, just like speakers who have remained monolingual. This makes for a reliable baseline of L1 proficiency prior to emigration and L2 immersion. This baseline is the starting point of the investigations into changes to L1 proficiency that will be undertaken in this dissertation.

L1 attrition has been the topic of dedicated studies in which individual linguistic levels were studied in detail, for instance phonetics (Chang, 2010; de Leeuw, 2008), morphology (Keijzer, 2007; Leisiö, 2001), syntax (Gürel, 2002) or the lexicon (Gross, 2000; Hulsen, 2000). However, the results from one domain are not easily generalised to other domains, which means that we do not know if, for instance, the vocabulary or morphosyntax of even the most strongly accented speaker in a phonetic study was affected by attrition as well.

This dissertation takes a different approach by investigating different facets of native-likeness in one group of German native speakers who have migrated to North America as adults. It will not only contribute to answering the question if language attrition does indeed affect different linguistic skills differently in the same speakers, but also study in how far L1 attriters behave similarly to monolingual native speakers, whose early linguistic experiences they share, and late L2 speakers, with whom they have been sharing the experience of bilingualism since their emigration.
The four areas of investigation in this thesis – selected with a view to covering as wide a range of linguistic levels as possible – are segmental phonetics, speech fluency, the use of formulaic language and morphosyntactic processing.

Due to its special role in language acquisition, phonetics is a particularly interesting aspect of proficiency: For late L2 speakers, native-like pronunciation is notoriously difficult to achieve (Levis, 2015; Piske, MacKay, & Flege, 2001). In view of this and the early onset of phonetic learning in the L1, it has to be asked if sound production is particularly deeply entrenched and more resistant to L1 attrition than other skills.

Our investigation of speech (dis)fluency focuses on the topic of parallel language activation. Spontaneous speech is the convergence of the output of various modules – such as lemma selection, syntactic structure building and phonetic production – all of which may operate differently in bilingual speakers. A study of the hitches and delays in this process, which also occur in monolingual speakers, provides a window into a natural way of using language that may be particularly challenging for attriting speakers.

In our study of formulaic language, we investigate one of the factors that are assumed to play an important role in speech fluency. It has been suggested that the use of prefabricated lexical chunks, which can be retrieved en bloc, saves speakers time that they can then put into utterance planning (Pawley & Syder, 1983). While L2 speakers have been shown to benefit from increased use of such chunks (Kormos & Dénes, 2004), our analysis will be the first to study this nexus with respect to L1 attriters.

We return to the topic of parallel language activation in our study on morphosyntactic violations: Non-native-like processing of agreement has been found in bilinguals, for example late L2 speakers (e.g., Sabourin & Stowe, 2008), but in these speakers, it may be caused by a lack of proficiency or by interference from another language. We have investigated the role of the latter factor in one of the first event-related potential (ERP) experiments with L1 attriters. Focusing on gender concord, our study will show if grammatical features that differ between the L1 and the L2 are subject to attrition.

The results of these four studies will be discussed in the light of three theoretical approaches to L1 attrition: generative, usage-based and dynamic. Particular regard will be given to factors that try to explain individual variability. Wide variability of ultimate attainment is one of salient aspects of late L2 acquisition with
speakers often showing considerable divergence from one another in terms of proficiency. Monolingual natives, by contrast, are typically distributed within a much narrower (and higher) proficiency range. Our studies on the L1 of bilingual speakers will help answer two questions: Is the increased variability seen in L2 speakers concomitant with any type of bilingualism? And, provided that this is the case, is the variability mediated by the behaviour of these speakers? Intuitively, a relationship between the strength of L1 attrition and factors like length of residence in the L2 environment (henceforth: L2 residence) or frequency of L1 use (L1 use) seems to suggest itself. However, prior investigations on L1 attrition have often failed and only sometimes succeeded in establishing such a relationship. If increased variability is present in attriters and a robust relationship with behavioural factors is absent, the influence of other factors would have to be considered.

In this introductory chapter, which is made up of four parts, a definition of late L1 attrition will be given first. I will then review the relevant experimental evidence for effects of attrition and bilingualism in the areas of phonetics, morphology and syntax, lexicon, and speech fluency. Each of these four topics will be explored in more detail in empirical studies that constitute the main part of the thesis. In the third part of the introduction, I will present three theoretical approaches that try to account for the data summarised in the literature review. Research questions and hypotheses, based on these theoretical frameworks, will be formulated in the last part of the introduction.

1.1 What is late first-language attrition?
The term *attrition*, defined by the *Oxford Dictionary of English* as “the process of reducing something’s strength or effectiveness through sustained attack or pressure” (Stevenson, 2010), is derived from Latin. It goes back to *attritum*, the perfect passive participle of *atterere*, meaning ‘to rub against; wear away; weaken’.

In linguistic usage, *attrition* has initially been used to refer to a wide variety of concepts. This included phenomena such as the intergenerational decline of the use of a minority language or dialect within a speaker community or the deterioration of linguistic skills in the healthy elderly (Seliger & Vago, 1991a). An early classification tried to disentangle the various uses of the term by distinguishing four types of language loss, each defined by what is lost and in which environment (de Bot & Welten, 1985; van Els, 1986). Not all of these types of loss are still classified as attrition.
in present-day research, but the second type is: “loss of L1 in an L2 environment”. This type is what this thesis is concerned with, more precisely with L1 attrition in German native speakers from Germany who have emigrated to Anglophone North America (i.e., the US or Canada) as adults. This latter aspect, concerning the age of emigration of the participants, is crucial because a language can also attrite before it has been acquired fully. This type of change, called *incomplete acquisition*, is experienced by children and adolescents who are exposed to the complex dynamics of learning and forgetting several languages at a time (Montrul, 2008).

While the details of the workings of language attrition are disputed, the conditions under which it comes about are not: It can happen in healthy, unimpaired individuals who, after fully acquiring their L1, migrate to an environment where their L1 is used only rarely. By consequence, they will also use this language and be exposed to it more rarely than before migration (Schmid, 2011). With 244 million international migrants in the world in 2015 (United Nations, 2016), the societal relevance of the topic is evident. By their own choice or not, most migrants live in environments in which their first language – or the variety of their first language they speak – is not widely used. The decline in use and exposure experienced by these speakers can result in changes to – and in most cases a decline of – their linguistic abilities. It is important to emphasise that speakers who migrate to a non-L1 environment must not be seen as attriters by default. Rather, attrition is a process that *can* ensue on some linguistic levels after prolonged lack of exposure to a language. More appropriately, one could call these speakers *potential attriters* (Cherciov, 2011; Ecke, 2004; Keijzer, 2007), but for the sake of brevity and in keeping with common usage, the term *attriters* will also be used in this thesis.

The process of change and potential decline has been described as a “deactualization” of the language (Hyltenstam & Viberg, 1993). Using this word or *attrition* – rather than for example *language loss* – is terminologically more precise: On the one hand, loss can be understood as referring to the complete breakdown of a linguistic system. The data in the literature review and in the main part of this thesis show that this is not a typical development in post-puberty attriters. On the other hand, language attrition is a phenomenon that can only be observed at the surface, that is, in the current behaviour of speakers, making it impossible to conjecture about the future development of their linguistic performance. It has even been stated that actual loss in adult migrants probably does not happen at all. Some
see temporary inaccessibility, which includes the possibility of regaining access to what has become inaccessible, as a more likely scenario (Paradis, 2004, 2007). The move away from the term *loss* is also an attempt to avoid the stigma associated with it: L1 attrition is not a purely linguistic phenomenon, but has social-psychological consequences as well and raises questions of identity and belonging (Ecke, 2004; Prescher, 2007). Also, the term *attrition* is better suited to the predominant view that attrition is “not an exotic and isolated phenomenon experienced [...] under extreme circumstances but part and parcel of the overall process of language development” (Schmid, 2013, p. 96).

Given that attrition always happens in a second-language environment, it is unsurprising that the L2 has been found to shape the way the first language of a given individual changes. An entire volume has been devoted to investigations of the specific influence of the second on the first language (Cook, 2003). Classifications of the ways in which the two languages interact in the mind of the bilingual speaker have been devised, such as the *Crosslinguistic Influence* framework (Pavlenko, 2000, 2004). Still, one has to recognise that not every aspect of L1 attrition can be traced back to the L2. It was already in the early phase of attrition research that a distinction between externally induced attrition, that is, changes caused by the L2, and internally induced attrition, that is, changes “motivated by universal principles [of language acquisition; c. b.] or [...] related to some fact in the particular grammar of L1” (Seliger & Vago, 1991b, p. 10) was suggested. It was assumed that internally induced attrition would manifest itself most profoundly in morphosyntax and phonology. The distinction between internally and externally induced attrition is still occasionally applied, albeit mainly in studies that are based on nativist theories (Gürel, 2008). However, any kind of linguistic change that can be traced back to general principles of cognitive organisation, rather than the specific rules of an L2, could be seen as a case of internally induced attrition.

Regardless of whether change in an L1 can be directly related to an L2, the general methodological question arises how change can be defined at all, particularly in cross-sectional research. In most such studies, no data are available about the linguistic behaviour of the participants prior to emigration. This was sometimes compensated for by focusing on phenomena that were assumed to be absent from the language use of typical native speakers, that is, errors (Håkansson, 1995; Hulsen, 2000; Köpke, 1999; Schmid, 2002). It has been pointed out that this type of analysis
is likely to paint a simplistic picture of language attrition (Schmid, 2004), leading, on the one hand, to a strong focus on abilities that are lost, without putting them in perspective.

On the other hand, restricting oneself to overt errors can result in missing aspects of language attrition that are not visible to the naked eye. Comparing the data of potential attriters to those of a control group of non-emigrated native speakers in experimental settings or spontaneous speech was seen as a remedy for the excessive focus on errors (Schmid, 2004). In such analyses, it is not necessary any more to start with the premise that the group of emigrants will display ‘deteriorated’ L1 performance, but it becomes possible to focus on subtler manifestations of language change. This approach, however, has disadvantages of its own, due to the fact that the linguistic environment they live in is not the only difference between attriters and non-emigrated native speakers of the same language. Many attriters are also competent speakers of their L2, that is, bilingual, whereas control group speakers are functionally monolingual. There is ample experimental evidence of the profound changes that the presence of a second language brings about in the mind of a speaker (see Costa & Sebastián-Gallés, 2014, and Dijkstra, 2005, for overviews; the topic of crosslinguistic influence will be discussed in more detail in Chapter 4). This means that comparing monolingual non-attriters to bilingual potential attriters confounds two phenomena, namely (potential) language attrition and bilingualism, thereby making it difficult to attribute differences between the groups to one of the factors. Still, hardly any feasible alternatives exist, except for longitudinal studies, which are associated with their own set of practical problems.

In this thesis, the potential attriters in North America will be compared to a control group of predominantly monolingual L1 speakers of German in Germany and, wherever possible, a group of advanced L2 speakers of German with English as their L1.

1.2 Experimental evidence on L1 attrition
This section will provide a look into four domains in which L1 attrition has been found. Literature reviews will be presented as well as short introductions to the methodology of the research papers in the main part. The length of the literature reviews varies: The one on phonetics and phonology is brief, summarising the longer discussion in the paper itself; the one on speech disfluencies is concise for the same reason and because this topic has rarely been researched in attriters.
Morphological and syntactic attrition, by contrast, will be discussed in depth to complement the short article in the main part. The review on lexical and semantic attrition is detailed as well, with a focus on phenomena other than formulaic language, which the study in the main part deals with.

1.2.1 Phonetics and phonology
1.2.1.1 Literature review
The study of L1 attrition in segmental sound production has mainly relied on two paradigms: native-likeness ratings on the one hand and acoustical measurements on the other. The former approach consists in having native speakers judge the pronunciation of other speakers. The advantage of these ratings is that they provide a global impression of changes in someone’s speech and make clear in how far such changes are perceptible or communicatively relevant. Its disadvantage is that ratings are unspecific with respect to the features that cause a non-native-like pronunciation. Acoustical measurements are used to fill this gap. Here, recordings of the tokens of individual sounds are analysed, using software to measure, for instance, voice-onset time (i.e., the time between the burst of a stop consonant and the first vibration of the vocal folds in the subsequent vowel) or vowel formants (i.e., characteristic spectral peaks in the sound spectrum), to name just two phenomena that have been investigated frequently. Most of the literature on phonetic (and, to a lesser extent, phonological) attrition looks for externally induced attrition, that is, it is explicitly or tacitly assumed that crosslinguistic influence between the L1 and the L2 will lead to sound changes in the L1.

Voice-onset time (VOT) has been studied acoustically because there are salient, easily detectable differences between two groups of languages: English and German, on the one hand, have long lags (>35 ms) between stops and vowels, whereas many Romance languages, on the other, have short lags. Several studies have indeed confirmed that the L1 speakers of a language of either group will be influenced by the longer or shorter L2 VOT value. Short VOTs getting longer under the influence of a long-lag language have been found for L1 speakers of French (Flege, 1987), Brazilian Portuguese (Major, 1992), Spanish (Williams, 1980) and Dutch (Mayr, Price, & Mennen, 2012), mostly under the L2 influence of English. Shorter VOTs, compared to monolingual natives, have been found for L1 English speakers with French as their
L2 (Flege, 1987). The literature suggests that immersion, that is, being in an environment where most linguistic interactions happen in the L2, is a mediating factor for the strength of L2 influence (see Chang, 2012; but cf. Flege & Eefting, 1987). Of consonantal phenomena other than VOT, the production of the liquids /l/ and /r/ has been investigated, albeit only in L1 speakers of German in Anglophone settings. In both sounds, L1 speakers showed changes towards the L2 (de Leeuw, 2008; Ulbrich & Ordin, 2014).

The acoustical analysis of vowels has received less attention and has been less focused on a small set of contrasts. Studies on the vowels of English-Korean bilinguals, for instance, have produced mixed results: Changes to the fundamental frequency (f0) of English vowels have been attested in beginning and advanced learners of Korean (Chang, 2012, 2013), but the vowels of L1 speakers of Korean remained unchanged in English immersion (Baker & Trofimovich, 2005). The authors of the latter study speculated that the relatively low level of L2 proficiency in their participants might have played a role – an important reminder that L2 influence on the L1 often seems to be mediated by L2 proficiency (but not always, as Chang’s results of beginning learners of Korean demonstrate). Changes in the L1 have been attested in bilingual speakers of L1 Quichua and L2 Spanish (Guion, 2003). Interestingly, the vowels of these speakers did not shift towards L2 values, but in the other direction, presumably in an attempt to establish a more robust differentiation between the L1 and the L2.

The extent to which such measurable changes to the pronunciation are noticeable by other L1 speakers of the same language has been gauged in a small number of studies, all of them dealing with L1 speakers of German. Based on native-likeness ratings of (partially) the same recordings of spontaneous speech, two studies found conflicting results: In one case, the bilinguals in English or Dutch immersion were rated as sounding less native-like than a control group of monolingual speakers of German (de Leeuw, Schmid, & Mennen, 2010), whereas the other found bilinguals to be indistinguishable from monolinguals (Hopp & Schmid, 2013). The crucial difference between these two studies is that in the latter, the speech samples of the attriters were mixed with samples of monolinguals and L2 speakers of German (L1 English). It may be the case that the presence of the latter group of more heavily accented speakers made raters less sensitive to the minute differences between the speech of controls and attriters that were noticed by raters in the first study,
which did not include samples of L2 speakers. It has also been demonstrated that not only speaker, but also rater effects can influence the results of a native-likeness rating (Schmid & Hopp, 2014).

1.2.1.2 The present study
Even this brief literature review makes it clear that many questions with respect to L1 attrition in phonetics and phonology remain to be answered. We will contribute by focusing on vowels, which have been much less widely studied than consonants, and by combining the abovementioned two strands of research – acoustical analyses and native-likeness ratings – in one study. Based on the Speech Learning Model (Flege, 1995, 2002), which predicts mutual influence between the sounds of a speaker’s L1 and L2, the link between audible and measurable changes will be investigated.

We present the results of a native-likeness rating in which we have assessed if spontaneous speech samples of potential L1 attriters of German with English as their L2 sound less native-like than samples of monolingual speakers (n = 33 for both speaker groups). This rating showed that the bilingual speakers could be divided into two groups: one that did not sound native-like to native raters and one that did. We subsequently acoustically analysed tokens of four sounds of German – three vowels and one consonant – to test if the two groups of bilinguals differed from monolinguals (n = 10 for all groups) in the production of these sounds. Additionally, the results of the ratings and the acoustical analysis were entered in a statistical model.

The results of all analyses will be presented in Chapter 3.

1.2.2 Morphology and syntax
1.2.2.1 Literature review
Morphological and syntactic aspects of L1 attrition have mainly been researched with respect to two parts of speech, namely (pro)nouns and verbs.

Inflection – that is, the obligatory marking of grammatical gender, case, tense or aspect on these or accompanying elements – can cause persistent difficulties for L2 speakers, even after prolonged exposure (Hopp, 2010; White, 2003). Studies in attriters, by contrast, have found that these speakers seem to struggle less with nominal and verbal inflection, while still not remaining fully native-like in all cases.

For the nominal domain, this was shown in an investigation of two groups of L1 speakers of German: one in English Canada, the other in France (Köpke, 1999).
Just like English, French does not require case marking, but grammatical gender is marked on determiners, adjectives and verbs. In a picture description task, bilinguals with L2 English produced more marking errors (2.02 per 1000 words) than speakers with L2 French (1.64) and more than twice as many as monolingual controls (0.86). A statistical analysis of the raw counts, however, showed that these group differences were not significant. Descriptively, case errors were the most frequent type of mistake in all three groups, followed by gender errors. In a sentence construction task, by contrast, both groups of bilinguals did produce significantly more errors than the monolinguals. These conflicting results may stem from differences between the tasks: In the picture description task, participants could avoid structures they were not sure how to use, whereas the sentence construction task narrowly constrained their responses.

Köpke’s results are similar to those of an analysis of the L1 speech of German natives who migrated from Nazi Germany to English-speaking countries (Schmid, 2002). In a corpus of biographical interviews, interferences with respect to grammatical gender, case and tense marking made up 15% of all interferences. For methodological reasons, no monolingual control group could be tested. When we compare the data of these bilinguals to those in Köpke’s study, the normalised number of interferences with respect to gender and case marking is slightly higher. This could be due to differences in years of residence in the L2 environment (x̄ = 29.4 for Köpke’s Anglophone group and 17.6 for her Francophone group, compared to > 45 in Schmid’s Anglophone participants).

In an investigation of L1 speakers of Dutch in English Canada, spontaneous speech and experimentally elicited production data were collected (Keijzer, 2007, 2010); this combination is reminiscent of the methodology in Köpke’s study of speakers of German. The formation of plurals, diminutives and agentive nouns as well as adjective inflection and determiner selection were studied. In free speech, use of the wrong determiner was significantly more frequent in the bilingual speakers than in a monolingual control group. In all other domains, the bilinguals only produced a negligible number of non-canonical forms. In the controlled elicitation task, by contrast, significant differences between monolinguals and attriters were evident in all domains except one type of plural formation and determiner use for the neuter gender. The conclusions that can be drawn from these (and Köpke’s) results seem relatively straightforward: When it comes to nominal inflection,
bilinguals do differ from monolinguals, but more often than not, attrition can only be measured when specific forms are elicited using a controlled paradigm. When attriters speak without experimental constraints, they succeed in circumnavigating forms they might be unable to produce correctly.

The processing of gender-marked determiners seems to be even less strongly affected than the production of inflected forms. Using a priming paradigm, it was assessed if grammatical gender marking on adjectives influenced the speed and accuracy of word/non-word decisions on subsequent nouns (Scherag, Demuth, Rösler, Neville, & Röder, 2004). Potential attriters of German with English as their L2 performed indistinguishably from a monolingual control group (whereas L2 speakers of German with English as their L1 were slower and less accurate).

Studies of the attrition of verbal inflection include an analysis of the spontaneous speech of L1 Spanish speakers in the United States with L2 English (Silva-Corvalán, 1991). In first-generation immigrants, that is, late bilinguals, no attrition in the use of the tense/mood/aspect system of Spanish was evident.

As mentioned above, Köpke (1999) did not find a significant difference between German monolinguals and bilinguals with English as their L2 in the number of grammatical errors in the picture description task. Descriptively, the number of tense or other verb inflection errors was extremely low. In the sentence construction task, there was a significant difference, with verbal inflection errors being about as frequent as errors in noun inflection in the attriters. An even higher number of errors was found by Schmid (2002), whose participants with L1 German/L2 English produced nearly three times as many interferences as the speakers in Köpke’s study. It was shown that irregular verb inflection, despite being relatively common in high-frequency verbs in German, was more vulnerable to interferences than regular inflection. This may partly be a task effect with detailed autobiographical interviews, used in Schmid’s analysis, calling for more complex language than short picture descriptions, as in Köpke’s study.

In the study of L1 Dutch/L2 English speakers by Keijzer (2007), the analysis of verbal inflection showed a pattern that was very similar to what we have seen with respect to nominal inflection: In the free spoken data, very few interferences were found (with the exception of past tense inflection), whereas the controlled elicitation task showed that attriters made significantly more errors in all domains.
Interferences in nominal and verbal inflection were also found in L1 speakers of Russian (L2 English; Pavlenko, 2003), French (L2 Hebrew; Ben-Rafael, 2004) and Greek (L2 English; Pelc, 2001).

Syntactic attrition has been demonstrated for a wide range of phenomena including middle constructions in French (L2 English; Balcom, 2003), word order in Spanish wh-questions (L2 English; Perpiñán, 2011) and complementiser use in German infinitive clauses (L2 Dutch; Ribbert & Kuiken, 2010). No effects of attrition, by contrast, were found in a study on wh-scrambling in Turkish (L2 English; Gürel, 2015) and differential object marking in Spanish (L2 English; Chamorro, Sturt, & Sorace, 2015).

Embedded clauses are one topic that several studies have been concerned with, obtaining conflicting results in some cases: For the processing of relative clauses, an experiment showed an influence of the L2, English, on the acceptability judgments and event-related potentials (ERPs) of L1 speakers of Italian (Kasparian, Vespignani, & Steinhauer, 2014). Bilinguals, unlike monolinguals, rejected grammatical Italian relative clauses when the word order was ungrammatical in English; also, attriters “showed morphosyntactic violation effects (LAN or P600), whereas Italian controls showed reliance on semantic cues (N400) and garden-path effects (P3a)” (p. 159). The difference between attriters and controls was largest for the bilinguals who gave the lowest ratings to Italian structures that are ungrammatical in English, that is, applied the rules of the L2 grammar to the L1.

The production of relative clauses has been investigated in an experiment with L1 speakers of Turkish (L2: Australian English), who had to rearrange scrambled constituents to form structurally correct sentences (Yağmur, 1997, 2004). The bilingual speakers produced significantly fewer acceptable sentences than monolinguals; more highly educated participants performed better in both groups. In a more recent analysis of Turkish embeddings in spontaneous speech, the difference between monolinguals and bilinguals was less striking (Yılmaz, 2011). Here, bilinguals, who spoke Dutch as their L2, used only one out of four types of embeddings (prepositional clauses) less frequently than monolinguals did and showed a minimal share of incorrect structures. Again, this difference is reminiscent of the results by Köpke and Keijzer, cited above, in whose studies the differences between monolinguals and bilinguals were largest in controlled experiments and often absent in analyses of spontaneous speech. In the case of embeddings in Turkish, it has to be added that the
task in Yağmur’s study was a particularly artificial assignment, the results of which might not be representative of natural language use in either participant group.

Structural differences between relative clauses in Turkish, on the one hand, and Germanic languages like English and Dutch, on the other, also lead to differences in the binding properties of pronouns as a potential gateway for attrition. In an English sentence like _John believes that he is intelligent_, the personal pronoun in the embedded clause can refer either to _John_, the subject of the matrix clause (bound reference), or to some individual outside the sentence (disjoint reference). In Turkish, by contrast, overt personal pronouns cannot be bound by the matrix subject, but only allow a disjoint reading. Reflexive and null pronouns, by contrast, allow both bound and disjoint readings. In a series of experiments, it was found that this rule of Turkish L1 grammar is weakened in native speakers in English L2 immersion. Bilingual speakers were more likely to accept not only disjoint, but also bound reference of overt pronouns, which is considered to be ungrammatical by monolinguals, and they had a preference for bound reference in reflexive and null pronouns, which is grammatical, but less strongly preferred by monolinguals (Gürel, 2004a, 2004b). Similar results were found in a group of L1 speakers of Turkish with Dutch as their L2 (Gürel & Yılmaz, 2011).

Investigations of phenomena at the syntax-discourse interface have focused on the distribution, position and interpretation of (pro)nouns. One strand of research has dealt with crosslinguistic differences in the interpretation of ambiguous relative pronouns. In a sentence like _Someone shot the servant of the actress who was on the balcony_, the relative pronoun _who_ (and, by extension, the entire relative clause) can refer either to the servant or to the actress. Native speakers of so-called low-attachment languages preferentially read the relative clause as a reference to the second NP (i.e., the actress); this group of languages includes English, Arabic, Romanian, Swedish and Norwegian. In high-attachment languages, by contrast, the pronoun is preferentially attached to the first NP (i.e., the servant); this is the case for Spanish, Greek, German, Dutch and French. In three studies with L1 speakers of Spanish in L2 English immersion, it was investigated if the L1 preference was influenced by the conflicting L2 preference. This was confirmed by questionnaire data, showing that the L1 preference for high attachment was reversed in bilinguals, and in a self-paced reading experiment, where sentences encouraging high attachment were read more slowly by bilinguals (Dussias, 2003). An eye-tracking experiment, using
similar materials, corroborated these findings: Here, L1 attriters, unlike monolingual controls, took longer to read the disambiguating element in a relative clause that was attached to the first NP (Dussias, 2004). The role of immersion in the reversal of attachment preferences in bilinguals was highlighted in another eye-tracking experiment that used the same materials. This time, two groups of L1 speakers of Spanish who were highly proficient in English participated: speakers who were immersed in an English setting, but also speakers who lived in Spain and were trained there to be translators for English. Interestingly, a preference reversal, measured through longer reading times for sentences encouraging high attachment, was only found in the immersed bilinguals. The equally proficient bilinguals in an L1 setting processed the sentences much like monolingual speakers (Dussias & Sagarra, 2007).

The interpretation of personal pronouns is another source of potential interference from English that has been investigated with respect to L1 attrition. Languages like Italian, Spanish, Greek and Turkish show an alternation, which does not exist in English, between overt pronouns and null pronouns in the subject position. In Italian, both Lui è partito (He has left) and È partito (* Has left) is grammatical, whereas in English, only the sentence with an overt pronoun is. In null-subject languages (also called pro-drop languages), the contrast between null and overt pronouns is exploited for interpretative differences: The null option, as the default, is found in wide focus contexts, whereas overt pronouns are used to introduce new referents, to signal topic shifts or to place a narrow focus on the subject. Given the unavailability of null subjects in English, it is hypothesised that attriters in an Anglophone setting are more likely to produce and accept sentences with a non-canonical use of pronouns (e.g., overt pronouns in unmarked contexts). This was confirmed for L1 speakers of Italian with English as their L2 in an offline sentence interpretation task of subordinate clauses (Tsimpli, Sorace, Heycock, & Filiaci, 2004). In both anaphoric and cataphoric contexts, bilinguals were more inclined to attach overt pronouns to the subject of the matrix clause, compared to monolinguals who preferred reading overt pronouns as object reference or sentence-external reference. The interpretation of null subjects in bilinguals, by contrast, showed reinforcement, rather than a reversal, of the preferences that are found in monolinguals: Whereas the monolinguals occasionally construed null subjects as object reference, the bilinguals hardly ever did so. Similar results were obtained for L1 speakers of Spanish with L2 English (Chamorro, Sorace, & Sturt, 2015): In an acceptability judgment task,
no differences between potential attriters and monolinguals were found, but an
eye-tracking experiment revealed a reduced sensitivity for non-canonical reference
in attriters. Whereas monolinguals had longer reading times for overt pronouns
with subject reference, attriters did not show a pronoun × antecedent interaction.

An influence of English as an L2 has also been demonstrated for the interpreta-
tion of anaphoric pronouns in L1 German (Wilson, 2009): In eye-tracking experi-
ments, German monolinguals were found to prefer postverbal non-topics as ante-
cedents of demonstratives (der, as opposed to the personal pronoun er; exp. 5). This
tendency was not found among bilingual L1 German/L2 English speakers. The longer
they had lived in the L2 environment, the weaker their preference for a postverbal
referent became with some speakers eventually shifting to a preference for pre-
verbal referents (exp. 9). Wilson suggests a similarity between attrition effects on
the interpretation of demonstratives in German and of overt pronouns in Italian
(Tsimpili et al., 2004).

Another difference between English on the one hand and Italian or Greek on the
other concerns postverbal subjects: In Italian, for example, Gianni è partito (John
has left) and È partito Gianni (* Has left John) is both grammatical, whereas only the
former word order is available in English. Postverbal subjects are particularly com-
mon in Greek, a vs(o) language. An influence of the L2, in which the subject can only
appear in front of the verb, would predict a wholesale increase of the share of prever-
bal subjects in the L1 of attriters. This is indeed what was found in a sentence produc-
tion task with L1 speakers of Greek with L2 English (Tsimpili et al., 2004). Also, these
bilinguals were less sensitive to interpretative differences arising from the contrast
between pre- and postverbal subjects. In a sentence interpretation task, bilinguals
were more inclined to read preverbal subjects, which are typically topics (i.e., old
information) in Greek, as new information. Monolinguals showed no such tendency.

Summing up, we can see a divide between morphology on the one side and syn-
tax as well as interface phenomena on the other. Nominal and verbal inflection, as
the morphological topics that have been studied most widely, are not impervious to
change in the L1, but the amount of attrition found in bilingual speakers can hardly
be called dramatic. Also, there are several studies that did not find any significant
differences between monolingual and bilingual speakers, particularly in analyses
of spontaneous speech. In the syntactic domain, the number of studies in which
monolingual and bilingual groups perform indistinguishably is smaller. Without a
doubt, this is partly due to methodological decisions, favouring controlled experiments over analyses of spontaneous speech in many cases. The differences between monolinguals and bilinguals are larger still in interface phenomena. This could be due to the relative infrequency of such structures in L1 input (as compared to, for instance, gender-marked determiners or inflected verbs), rendering them more vulnerable to influence from the dominant L2.

1.2.2.2 The present study
A topic that has not been thoroughly studied in attrition research so far is the processing of morphological marking. Also, with the exception of Kasparian et al. (2014), no experiments using electrophysiological measurements with potential attriters have been reported yet. For the study in the main part, we have analysed event-related potentials (ERPs) of monolingual and bilingual L1 speakers of German that were collected during the processing of morphological violations.

Our analysis included two types of structures: (1) combinations of gender-marked determiners and nouns; (2) combinations of auxiliaries/modals and non-finite verb forms. Half of the sentences were incorrect in both types of structures: In the first condition, the marking of grammatical gender on the determiner did not match the gender of the noun (e.g., pairing a determiner marked for neuter with a masculine noun); in the second condition, finite and non-finite verb forms were incorrectly paired (e.g., combining an auxiliary with an infinitive, rather than with a past participle).

The grammatical gender system of German will be outlined in Chapter 6. The results of all analyses will be presented in a subsection of this chapter.

1.2.3 Lexicon and semantics
1.2.3.1 Literature review
Research on L1 attrition in the domain of the lexicon and semantics has centred upon the production of words, both in experimental settings and in spontaneous speech.

Anecdotally and by early attrition researchers, it has been stated that the mental lexicon is affected first or is particularly vulnerable to loss (e.g., Andersen, 1982). Such claims have been disputed on practical and theoretical grounds (Schmid & Köpke, 2009). First, longitudinal data, which could show that the lexicon attrites before other linguistic areas, are not available; second, a weighted comparison between different linguistic areas seems difficult in the absence of shared variables that are
measured, an objection that was already discussed in early bilingualism research (Weinreich, 1953, p. 58).

Experimental studies in this domain have produced mixed results, for instance when it comes to the elicited production of individual words. A study with Dutch emigrants to Australia (L2: English), using a picture-naming task, found that these speakers were less accurate in naming drawings, but not slower in doing so (Hulsen, 2000; Hulsen, de Bot, & Weltens, 2002). There was a significant negative correlation between the reaction times and the use of the L1 in everyday life. Earlier research on a similar group of speakers (L1 Dutch; L2: Australian English) had also found lexical attrition (Ammerlaan, 1996). In this case, it was shown that the performance of the bilinguals was influenced by lexical similarities between the L1 and the L2: Cognates (i.e., words with similar meaning and form in L1 and L2) were produced more accurately than the average of all test items, whereas false friends (i.e., words with similar form, but divergent meaning) led to more errors. No lexical attrition, by contrast, was found in a picture-naming experiment with L1 speakers of Turkish. The bilingual group (with Dutch as their L2) was as accurate and fast in naming drawings as a monolingual control group on both high- and low-frequency items (Yilmaz & Schmid, 2012).

Verbal fluency (VF) tests have also revealed lexical attrition in some speaker groups. In these tasks, participants are asked to produce as many words as possible from a specific lexical category (semantic VF) or with a specific first sound/letter (phonemic VF) in a predefined time interval (usually 60 s). Successful performance requires a combination of speed and accuracy in lexical access, which may be decreased in bilingual speakers due to language competition. A semantic VF task with L1 speakers of German and Dutch in English immersion and L1 speakers of German in Dutch immersion showed that bilinguals in all three groups named significantly less items than monolinguals (Schmid & Keijzer, 2009), but an analysis of the German subset of these data revealed that the score on the VF task was not correlated with L1 use (Schmid, 2007). Lower scores in a semantic VF task were also found in L1 speakers of Romanian in English immersion (Cherciov, 2011). However, no attrition was found in a group of L1 speakers of German in Ireland with English as their L2. In this study, a number of semantic and letter VF tasks were used, but bilinguals did not differ significantly from a monolingual control group on any of these (Opitz, 2010, 2011). In a group of L1 speakers of British or Irish English in
Germany (L2: German), two semantic VF tasks did not reveal any attrition either (Dostert, 2009). The differences between the former and the latter studies may be explained by for how long the participants have been residing in the L2 environment. The average length of residence was 26 years in Dostert’s (2009) English natives and 19.5 years in Opitz’ (2010, 2011) German natives, but more than 35 years in all three participants groups of Schmid & Keijzer (2009).

A particularly interesting study on single-word production was conducted with a group of L1 speakers of Russian in American English immersion (Pavlenko & Malt, 2011), who were shown pictures of drinking containers and asked to name them (e.g., cup, mug or glass). Their L1 and L2 naming preferences were compared with those of monolingual speakers of Russian and English. As groups, late bilinguals (the study also included two other groups of bilinguals) and monolinguals differed on the naming of 20% of the items. On a quarter of these divergent items, bilinguals used broad terms that are also applied to a range of other objects, whereas monolinguals preferred more specific words. In some cases, the tendency in the bilinguals to use cover terms seems to have been reinforced by naming patterns in the L2: Two objects, for instance, were called рюмка (rjumka), a narrow term for a specific type of shot glass, by Russian monolinguals, whereas bilinguals dominantly used стакан (stakan). The latter word is a broad term that translates to glass, which is also the word that the English monolinguals preferentially used for these objects. In other objects, bilinguals showed the reverse tendency by avoiding stakan, a term that, according to the preferences of the monolingual speakers, has a wider semantic range than glass in their L2. The experiment demonstrates how different processes in late bilinguals – ‘loss’ of low-frequency words as well as expansion and narrowing of semantic boundaries, partially due to L2 influence – contribute to a modest but measurable shift away from monolingual language use.

Lexical attrition has not only been found in experimental studies, but also in analyses of spontaneous speech. One study compared the naming of objects in a story retelling by L1 speakers of American English in Hebrew immersion (Olshtain & Barzilay, 1991). A number of low-frequency objects were identified for which all monolinguals used the same one or two words, whereas bilinguals came up with a variety of more or less precise terms. No statistical group comparison was made, probably due to differences in sample size, but descriptively it seems that bilinguals experience difficulty with accessing low-frequency items. When lexical access in
the L1 fails altogether, speakers have been found to use L2 lexical items in their L1 speech – a phenomenon known as code-switching or, when done word-internally, code-blending. It is, however, difficult to evaluate the amount of code-switching because monolingual speakers, who by definition produce no code-switches, cannot be used as a reference group. For this reason, the code-switching behaviour of different groups of bilinguals has been compared: One study included L1 speakers of German with L2 French and L2 English, of which the latter produced more than three times as many switches in picture-description and sentence construction tasks (Köpke, 1999, 2002). The difference could be attributed to the attitude of the speakers towards switches, which might be shaped by how well predominantly monolingual speakers of German speak their respective L2: French, on the one hand, is not widely spoken in Germany, so switching to French during an L1 conversation is unlikely to result in successful communication. On the other hand, chances are that L2 speakers of English – a language many Germans have at least basic proficiency in – will still be understood when interspersing their L1 speech with L2 words. Code-switching and related phenomena have also been found in L1 speakers of English (L2 Spanish; Porte, 2003) and French (L2 Hebrew; Ben-Rafael, 2004), but no quantitative analysis was done for either group. It is clear that longitudinal studies on the incidence of code-switches would be a valuable addition to the literature.

Studies at the single-word level have proven useful for exploring changes in clearly delimited semantic fields. For capturing the entire breadth of lexical usage, automatised analyses of spontaneous speech have been conducted. One simple way of assessing the lexical diversity of a speech sample is the type/token ratio (TTR), calculated by dividing the number of unique words by the number of concrete instances of these words (e.g., when the word girl is used five times, this would be one type and five tokens). A well-known disadvantage of the TTR is that it is negatively correlated with sample size, due to the recurrence of highly frequent (function) words. A number of formulae have been suggested to remedy this shortcoming, with the Uber score and the D score being considered the most accurate (Jarvis, 2002; McCarthy & Jarvis, 2010). For elicitation, many studies have asked participants to retell the storyline of an extract from Modern Times, a film by and with Charlie Chaplin (for details on the film and the typical procedure, see Chapter 2).

Many analyses have found no differences in lexical diversity between monolingual and bilingual L1 speakers (unless specified otherwise, the data were collected
using the Charlie Chaplin film and analysed by calculating the D score). This was true for language combinations including L1 Russian/L2 English (Dewaele & Pavlenko, 2003; Uber score), L1 English/L2 German (Dostert, 2009), L1 Romanian/L2 English (Cherciov, 2011), L1 German/L2 English (Opitz, 2011) and L1 Hungarian/L2 Danish (Varga, 2012). By contrast, significantly reduced lexical diversity in bilinguals has been found for bilingual speakers of L1 Dutch/L2 English (Keijzer, 2007), L1 German/L2 English or Dutch (Schmid, 2007; Schmid & Dusseldorp, 2010) and L1 Turkish/L2 Dutch (Yılmaz & Schmid, 2012). The reason for these divergent results has been sought in differences of sample size: For several of the analyses that showed no group differences, only a small number of speakers had been tested (n < 25). Therefore, the absence of a difference may be a type II error, that is, a failure to reject a false null hypothesis. Schmid & Jarvis (2014), who put forward this explanation, conducted a more sophisticated analysis of the spontaneous speech of L1 speakers of German with L2 English or Dutch (n = 53 for both bilingual and a monolingual group) from whom two types of speech samples – a biographical interview and a film retelling – were elicited. For both samples, type/token-based variables hardly differed between groups; differences between the distributions of items were only evident in the interviews and had small effect sizes. However, the bilinguals were found to overuse high-frequency and underuse low-frequency vocabulary, compared to monolingual speakers. Here, effect sizes were also small and a clear pattern of differences only emerged when the distribution of the frequency classes was based on the corpus analysed in this study, rather than on a larger external corpus. In addition to that, it was shown that only the German/Dutch, but not the German/English bilinguals used a higher share of words that are L1/L2 cognates. To identify predictors that are associated with attrition, a linear regression analysis was conducted. This did not produce significant results for both speech samples, demonstrating once again the difficulty of identifying external factors that make some speakers more susceptible to attrition than others.

Only one recent study has dealt with semantic processing in potentially attriting bilinguals (Kasparian et al., 2014). In an ERP experiment, L1 speakers of Italian in an Anglophone immersion setting listened to sentences with subtle and salient lexical mismatches: In one condition, fitting words were replaced by similar, but incongruous words (e.g., mento by menta, meaning ‘chin’ and ‘mint’, respectively); in the other, the replacement bore no phonological nor semantic similarity with
the fitting word. Participants had to give an acceptability rating after each sentence. Both types of violation elicited large N400 effects in monolingual speakers. In attriters, the presence of a native-like effect was modulated by general L1 proficiency: Highly proficient speakers did not differ from monolinguals, whereas speakers in the lower half of the proficiency range showed a P600 effect for the first type of violation (menta vs. mento) as a sign of “more elaborated processing further downstream” (p. 159), according to the authors. The processing differences revealed in this experiment may be an example of internally induced L1 attrition, given that the L2 had no specific role in rendering the bilinguals less sensitive to the subtle contrasts that were tested in the experiment. Rather, less exposure to the L1 over a long time seems to have led these speakers to confuse similar words.

In sum, the results of experiments and the analysis of free speech data are at odds with the impression that the lexicon is disproportionally affected by L1 attrition. In both picture-naming and verbal fluency tasks, a sizeable share of studies did not find any significant differences between monolingual and bilingual speakers. Lexical diversity in spontaneous speech has also been found to remain unaffected in many groups of bilinguals, although the non-significant results of some studies have been called into question for methodological reasons. Two in-depth studies – one at the single-word level (Pavlenko & Malt, 2011), one on global lexical diversity (Schmid & Jarvis, 2014) – have demonstrated that changes in the lexicon come in a number of subtle guises, including gradual boundary shifts and over- or underuse of certain frequency classes. The advantage of the approaches of these two studies lies in the fact that their analyses concern aspects of lexical usage that are below the threshold of perception and, consequently, not under the conscious control of a speaker.

1.2.3.2 The present study
A topic that has not yet been investigated in attrition research is the use of fixed multi-word expressions in spontaneous speech. Many studies cited above deal with the production of words in a natural sentential context, but none of these analyses have focused on groups of words that form a meaningful unit. Given the paucity of data on the use of formulaic language in bilinguals, it is difficult to form hypotheses about whether and how attrition in this type of language use would differ from attrition in individual words. On the one hand, the access to formulaic sequences in the mental lexicon may be impeded by the fact that each sequence
has a lower lexical frequency than its elements (which also appear in isolation or in other sequences). On the other hand, some types of formulaic sequences may be easier to remember for their use of metaphorical or figurative language. Also, formulaic sequences seem to play a beneficial role in the speech fluency of bilingual speakers, acting as “islands of reliability” (Dechert, 1983), that is, as prefabricated units the use of which buys the speaker some time for further speech planning.

Our analysis is based on an existing classification of formulaic sequences (Granger & Paquot, 2008) that was developed for English. This system differentiates between a large number of categories, of which only a small share appeared in the retelling corpus that we have used. The data come from potential L1 attriters and L2 speakers of German (with English as their L2 or L1, respectively) as well as a monolingual control group. We have analysed the frequency of different types of formulaic sequences and correlated individual participant data with measures of general proficiency, lexical diversity and speech fluency.

The results of all analyses will be presented in Chapter 5.

1.2.4 Disfluency
1.2.4.1 Literature review
The previous three subsections of this chapter have mainly dealt with discrete elements that are susceptible to L1 attrition: sounds, morphemes, words. This subsection and the corresponding paper in the main part focus on how these elements are strung together in spontaneous speech. It is not a new insight that speech production “in any speaker is a highly fragmented and discontinuous activity” (Goldman-Eisler, 1968, p. 31), characterised by frequent hesitations and the use of disfluency markers. The distribution of such markers has been found to differ among groups with, for instance, L2 speakers using more hesitation markers and longer pauses than L1 speakers of the same language (Hieke, 1981; Riazantseva, 2001; Wiese, 1984). The reason for these differences lies in the fact that disfluency markers reflect underlying processes of speech planning. The smooth execution of these processes crucially depends on two factors: The elements that have to be combined must be rapidly accessible (Timmermans, Schriefers, Sprenger, & Dijkstra, 2012), but the procedure of combining them also has to be rehearsed frequently to get deeply entrenched (R. Schmidt, 1992). In L2 speakers, both factors can be problematic, the former due to a lack of, for example, lexical knowledge, the latter due to a reduced
amount of language experience, compared to native speakers. For this reason, it seems interesting to investigate the behaviour of potential attriters, that is, a group of bilingual speakers who have fully acquired the language in question, but who, since their emigration, have used the procedural and declarative knowledge involved in sentence formulation much less frequently and who, just like other bilinguals, experience crosslinguistic influence from their L2.

The number of studies on speech disfluencies in L1 attriters is still low. Save one case study on incomplete acquirers of Japanese (Yukawa, 1998), only two full-fledged analyses of late attriters’ speech have been published: One of them focuses on L1 attriters of German and Dutch in Anglophone Canada and attriters of German in the Netherlands (Schmid & Beers Fägersten, 2010). Four types of disfluencies were investigated: empty pauses, filled pauses (elements like *uh*), repetitions and self-corrections (i.e., alterations of original material before an interruption; also called *retracings*). All attriting groups used more empty pauses, repetitions and self-corrections than corresponding control groups of monolingual speakers. The use of filled pauses was only increased in German emigrants to the Netherlands. An additional analysis concerning the position of disfluency markers in an utterance revealed two interesting details: First, it was shown that the differences between monolinguals and attriters were largest clause-internally, that is, in the middle of the process of formulating a sentence, rather than clause-finally. This suggests that bilinguals face greater challenges during the microplanning of utterances, but not during the macroplanning of storylines that happens between utterances. Second, a particularly high incidence of disfluency markers in the speech of attriters was detected before nouns and verbs, that is, before content rather than function words. This is reminiscent of earlier findings that access to L1 lexical items in bilinguals is prone to attrition (Dostert, 2009; Köpke, 2002; Schmid, 2007).

The second study, in which L1 speakers of Turkish with L2 Dutch were tested, seems to support the relationship between lexical access and disfluency (Yilmaz & Schmid, 2012). These bilinguals performed as well as monolinguals on a picture naming task, but had a lower lexical diversity and a higher use of disfluency markers in spontaneous speech. This suggests that lexical access, which was unimpaired in the designated task, may become too slow to feed quickly enough into automatic sentence generation procedures under the demands of spontaneous production.
1.2.4.2 The present study
Neither of the abovementioned two studies on disfluency has included L2 speakers in addition to L1 attriters. In the study that will be presented in the main part, we have analysed spontaneous speech data from both groups of speakers. The reasoning behind that lies in the factors that we assume to have an impact on speech production: One of these factors, namely crosslinguistic influence, is shared by attriters and L2 speakers. The other factor, namely a potential lack of proficiency in the testing language, is limited to the L2 speakers. Due to the combination of these two factors, the latter group should be more disfluent than L1 attriters. The degree of disfluency in the attriters can then be used as an estimate of the disfluency that arises from crosslinguistic influence.

Our analysis included data from the two bilingual groups (attriters: L1 German/L2 English; L2 speakers vice versa) and a group of monolingual L1 speakers of German (n = 20 for all groups). We have analysed the incidence of three disfluency markers: pauses (empty and filled), repetitions and self-corrections. Self-corrections were subdivided by contextual motivation to differentiate, for example, corrections that sought to rectify an error as opposed to corrections that served stylistic purposes. Also, the count of all types of disfluency markers was correlated with the results of a cloze test that provided an estimate of the global language proficiency of the participants.

The results of all analyses will be presented in Chapter 4.

1.3 Theoretical frameworks
The reviews of previous research have shown that attrition does not occur uniformly across – or even within – the four domains that have been discussed here.

In phonetics, a large number of studies have found measurable differences between monolingual and bilingual native speakers. However, these differences do not seem to be perceptible for other speakers in many cases. In morphology and syntax, there is evidence for a gap between basic inflection, which seems to remain largely unaffected by attrition, and more complex syntactic phenomena, where bilinguals perform differently from monolinguals. Studies on the lexicon and semantics, while reputed to be one of the domains most strongly affected by attrition, are characterised by mixed results across methodologies. Attrition research
on speech fluency is still in its infancy, making it difficult to judge how strongly this skill is affected by attrition.

What are the reasons for these differences? In the following, three theoretical approaches will be presented that make predictions of the susceptibility to attrition with respect to one or several of the linguistic areas discussed here. These theories will be used, as far as possible, to generate hypotheses for the analyses in the main part.

Before this, one theory shall be mentioned that will not be treated in detail here: The regression hypothesis may well be the oldest and best-known attempt to explain the differential impact of L1 (and L2) attrition. It assumes that attrition passes through the same stages as acquisition, only in reverse order (last in, first out). It is inspired by the observation that language acquisition typically starts with unmarked elements, proceeding to marked ones, which are also assumed to be lost first in attrition (Olshtain, 1989). However, a critical examination of this hypothesis, based on data from L1 Dutch/L2 English bilinguals, has strongly suggested that, albeit intuitively appealing, it cannot be considered a valid explanatory framework for attrition (Keijzer, 2007, 2010). The hypothesis was found to be systematically unsuccessful in predicting the experimental findings and was deemed to be cognitively implausible. Earlier, but less detailed investigations came to similar conclusions (Jordens, de Bot, & Trapman, 1989). For this reason, the regression hypothesis will not be used as a theoretical approach in this thesis.

1.3.1 Generative approaches: The Interface Hypothesis
The oldest theoretical approach to attrition that is still applied to current research is based on the work of Noam Chomsky about language acquisition and the structure of natural languages. In this, the observation plays a crucial role that any healthy child, born into whatever linguistic environment, will become a competent native speaker of the language it happens to be immersed in (Chomsky, 1965). To explain this observation, theoretical assumptions have been made about the learner and the language learned: The fact that all natural languages seem equally learnable has been taken to suggest that these languages are shaped by linguistic universals, that is, a common set of structural rules, considered to be a superset of the rules of each individual language (Stechow & Sternefeld, 1988). However, native speakers of a language are not explicitly aware of the structure of their language and these rules are never explained to naturalistic learners; rather, children are simply exposed to
spoken language, whereby “no special care is taken to teach them and no special attention is given to their progress” (Chomsky, 1965, p. 200 f.). It has been argued that children would be unable to derive the specific rules of the language at hand from the linguistic input alone, which is judged to be sparse and unstructured (the “poverty of the stimulus” argument; Chomsky, 1965, 1980). Therefore it is assumed that a child has access to Universal Grammar, that is, innate and implicit knowledge of the set of rules that define human language.

The nativist theory of Universal Grammar – prominently represented by Chomsky and embraced by a large community of researchers – has been a mainstream linguistic framework from the late 1950s to the early 1990s. Its observations and assumptions have been disputed on empirical and theoretical grounds (for recent examples, see Dąbrowska, 2015; Evans & Levinson, 2009), but are still widely applied in present-day research. One of the reasons for this is that the concept of Universal Grammar has been closely linked with a descriptive apparatus of the rules that natural languages are assumed to conform to. The description focuses on the syntactic structure of language, which is generated – hence the term *Generative Grammar* – from a set of elements through two types of rules (Chomsky, 1957): Phrase structure rules, on the one hand, define the internal structure of these elements (e.g., a sentence being made up from a nominal phrase and a verbal phrase, which in turn are made up of several elements), whereas transformational rules, on the other, link the ‘deep’ logical form of a sentence with the ‘surface’ phonetic form (e.g., deriving a question like *Are you married?* from the declarative sentence *You are married*). The details of workings of the transformational rules as well as the definition of the two levels of representation – phonetic form (PF) and logical form (LF) – have seen some changes from the so-called *Standard Theory* (Chomsky, 1965) to the *Minimalist Program* (Chomsky, 1995), the latest revision of the framework. *Government and Binding*, developed in the 1980s (Chomsky, 1980, 1981), still seems to be the most widely applied instantiation of generative syntax.

Chomsky’s own writings have been mainly concerned with the competence of monolingual speakers. To make generative theory applicable to bilinguals, attribution research has drawn upon the concept of principles and parameters (Chomsky, 1981, 1995), which tries to define the relationship between universal principles (see above) and language-specific parameters, both shaping the linguistic features of a language. Parameters are considered to be binary ‘switches’ that can be fixed
to one of the values. It has been suggested that for each of these switches, there is a default setting or, in more structuralist terms, unmarked option (Snyder & Lil-lo-Martin, 2011). The setting of parameters to language-specific values in child L1 acquisition happens through positive evidence; once a parameter has been set, it cannot be changed any more. It has been pointed out early on that a straightforward transfer of this theory of L1 attrition would predict no loss at all, given that there is no positive evidence against the features of the L1 in the input that attrit-ers receive (Sharwood Smith, 1989).

Still, three different hypotheses for explaining attrition within a generative framework have been proposed: In the earliest of them, it has been suggested that language attrition may differentially affect language-specific parameters as a function of their markedness. The ‘unmarking’ of marked values, that is, of parameters in which the language in question deviates from the default value, was considered unlikely (but see Håkansson, 1995, for a study that is based on the possibility of unmarking). By contrast, an L2 with a marked value for a feature that is unmarked (i.e., keeps the default value) in the L1 does provide positive evidence for a conflicting parametrisation between the languages of a bilingual. In these cases, attrition may be expected (Sharwood Smith, 1989; Sharwood Smith & van Buren, 1991). Given that the distinction between marked and unmarked features is theoretical, rather than empirical, there is still room for disagreement on whether a feature is marked or unmarked in a language.

The second hypothesis would not require challenging the tenet that parameters cannot be reset. Rather, it emphasises the opposition between competence, that is, the internalised knowledge of an “ideal speaker-listener” (Chomsky, 1965, p. 3) and performance, that is, the application of this knowledge by individuals under chang-ing conditions. In Generative Grammar, the former is seen as the primary object of investigation, whereas attrition has been recognised to be a phenomenon in which both performance and competence play an important role (Gürel, 2004a; Seliger & Vago, 1991b). When attrition is seen as a process that mainly depends on a loss of fluent access (i.e., performance) to stable mental representations of L1 competence, this would amount to declaring that theoretical approaches in the generative tra-dition cannot be useful for studying attrition – an “undesirable starting point […] since one can never probe competence without involving performance” (Sharwood Smith & van Buren, 1991, p. 19).
The third and arguably most promising option also leaves the assumption unchallenged that parameter resetting is unlikely to occur, but suggests a possible gateway for L1 attrition within the generative framework. The focus of the Interface Hypothesis (IH) lies on the distinction between ‘narrow syntax’ and phenomena in which syntax interfaces with other cognitive and linguistic domains (Sorace, 2011; Sorace & Filiaci, 2006). The IH posits that features of syntax proper are eventually learnable in late L2 acquisition and will not undergo L1 attrition in bilingual speakers. The aforementioned availability of null subjects in languages like Italian or Turkish is a purely syntactic feature, for instance. When, however, such syntactic features are exploited for discourse effects, L2 acquisition is more likely to be unsuccessful and L1 attrition may ensue, according to the IH. The preferred interpretation of anaphoric null and overt subjects in Italian and Turkish, for example, would be such an interface phenomenon. The distinction between core grammar and interface phenomena can also be put in terms of formal features of lexical items and the way these are treated in the computational system. According to recent revisions of Generative Grammar (Chomsky, 1995), features can be interpretable or uninterpretable. Uninterpretable features of, for example, (pro)nouns must be deleted in the grammatical derivation before reaching the level of the Logical Form (LF), that is, they cannot contribute to the semantic representation of the sentence. Interpretable features, by contrast, have semantic content and can be used at LF, that is, at the interface with the conceptual-intentional system. In L1 attrition (and L2 acquisition), interpretable features are seen as the source of crosslinguistic influence (Montrul, 2008). This hypothesis has been tested with respect to interface phenomena in German (Wilson, 2009), Spanish (Perpiñán, 2011), Italian and Greek (Tsimpli, 2007; Tsimpli et al., 2004).

For the data that will be presented in this thesis, the predictions using approaches in the generative tradition are fairly straightforward: To begin with, linguists working within generative frameworks have been mainly concerned with (morpho)syntactic phenomena as well as some aspects of phonology (starting with Chomsky & Halle, 1968) and semantics (in a tradition opposed to Chomsky: Lakoff, 1971). Generative approaches therefore do not lend themselves easily to formulating hypotheses on attrition in the domains of phonetics, speech fluency and formulaic language. The study on the processing of grammatical gender (and non-finite verb forms) in Chapter 6 is the one for which the generative framework can be considered
immediately relevant. Grammatical gender, as a part of the lexical entry of each noun, is seen as an uninterpretable feature and not assumed to constitute a functional head of its own (Alexiadou, Haegeman, & Stavrou, 2007; Bernstein, 1993; but cf. Tsimpli, 2003). When it comes to grammatical gender agreement on definite determiners, the parametrisation of the two languages in question differs: The L1 of the speakers that were tested, German, has grammatical gender marking, whereas their L2, English, has not (it was lost in the 14th century; Faiß, 1993). Only the most liberal interpretations of the ‘unmarking’ of parameters suggest that marked L1 parameters can be overwritten by unmarked L2 parameters (Håkansson, 1995). This would, however, be required when predicting an influence of English, in which the parameter for gender is unmarked, on the marked setting in German. In the case that is considered more likely, the marked L1 setting would remain unaffected, in which case no attrition would be predicted (Sharwood Smith, 1989). The same prediction can be made with respect to the interpretability of the features: The use of grammatical gender in German, as it was tested in the experiment reported here, is a purely morphosyntactic phenomenon that is not used for discourse or other interface effects. Based on the Interface Hypothesis, no L1 attrition would therefore be expected (Sorace, 2011). In the non-finite verb forms, which are only used as a control structure in the experiment, no differences between the L1 and the L2 exist at all: In both languages, auxiliary verbs are combined with past participles (hat gefragt ‘has asked’), whereas modal verbs are combined with infinitives (kann fragen ‘can ask’).

1.3.2 Usage-based approaches: the Activation Threshold Hypothesis

The best known usage-based framework that has been applied to attrition, the Activation Threshold Hypothesis (ATH), stands in stark contrast to generative approaches, like the Interface Hypothesis (IH). As part of a neurolinguistic theory (Paradis, 2004) that attempts to elucidate the brain activity underlying bilingual production and comprehension, it is not limited to one linguistic level and focuses on cognitive, rather than structural aspects. The ATH tries to explain the mechanisms behind the volitional selection of languages in bilinguals and the automatic performance in the selected language. It suggests that the neurofunctional system of a multilingual is comprised of subsystems, each representing one of the speaker’s languages. These subsystems, while connected to a single system that contains representations of conceptual features, compete with one another for production,
which can lead to interference within the system. It is suggested that the activation of items or mental representations from each language requires a certain amount of impulses to reach the neural substrate of these items (Paradis, 1993). The amount of neural energy that is needed for the activation of an item is its activation threshold. Production and processing, while subserved by the same substrate, are assumed to require different amounts of energy; more impulses are required to self-activate an item for production than to activate it due to external impulses during processing. Each call of an item activates a trace, a term that is used in the ATH as a metaphor for the memory of previous calls, and contributes to lowering the activation threshold. When an item is inactive, its threshold gradually rises. This means that the value of the activation threshold of each item is determined by two factors: the frequency and the recency of previous activations. Interference between the subsystems, that is, the selection of an element from a non-selected language, happens when the activation threshold of an item in the non-selected subsystem is lower than that of the corresponding item in the other subsystem.

In language attriters, both the frequency and recency of L1 activations are subject to change, due to a prolonged reduction of L1 use. According to the ATH, this type of language disuse affects language production and processing for two reasons: On the one hand, neurocognitive systems are assumed to require stimulation to remain in good condition (Paradis, 2007). On the other, the activation of an item for selection is not only effected by lowering the activation threshold of the item itself; at the same time, competing items – including translation equivalents – have to be inhibited, that is, their threshold is raised during the activation of the target item (Paradis, 1993). In the long run, extensive L2 use will therefore contribute to generally raising the activation threshold of all types of L1 items. Comprehension in attriters should be less strongly affected than production, which is more effortful because it requires self-activation rather than being dependent on external stimulation. The ATH does not assume L1 threshold levels to rise across the board under L2 influence, but makes different predictions for two kinds of memory systems (Ullman, 2001b; Ullman et al., 1997): Declarative memory, subserved by the hippocampus and related areas, is a task-independent and informationally non-encapsulated system; it stores explicit metalinguistic knowledge, which is consciously learned and used, such as lexical items. Procedural memory, rooted in structures of the frontal and basal ganglia, is task-specific, by contrast; it represents implicit
linguistic knowledge that is learned incidentally and used automatically, such as processing routines (Ullman, 2001a). The ATH predicts that declarative knowledge, which is not modality-specific and therefore more susceptible to interference, will be affected more quickly than procedural competence (Paradis, 2007). An increase in the access time of words or multi-word sequences may be one of the first signs of attrition. This is reminiscent of the assumption that the lexicon, represented in declarative memory, will be affected first or most strongly in attrition, which is probably an overstatement (see section 1.2.3).

In the neurolinguistic theory that the ATH is part of, each language subsystem is modular, consisting of individual components for linguistic areas (phonology, morphology, syntax etc.) that can be differentially affected by attrition. However, there are no specific hypotheses with respect to these components beyond the general prediction that language use will be a primary determinant of activation thresholds and, as a consequence, of potential L2 interference. Paradis does note that attrition is “liable […] to follow the reverse of the developmental sequence found in L1 acquisition” (Paradis, 2007, p. 127), but in the light of recent insights that have cast doubt on the validity of the regression hypothesis (see above), this suggestion may have to be reconsidered.

Apart from the amount of use, a speaker’s motivation and attitude towards the L1 are also assumed in the ATH to have an impact on the strength of attrition. Both factors have been shown to influence the success of L2 acquisition; therefore, it is predicted that they are also able to contribute to lowering or raising activation thresholds in attrition.

A difficulty for the application of the ATH – and, by extension, of other usage-based approaches – lies in accurately quantifying the factors that are seen as the main determinants of attrition: individual language use and frequency of occurrence of items. It has been pointed out before that a linear relation between the amount of L1 use and proficiency is difficult to establish (Schmid, 2007), probably due to the fact that self-reported patterns of use might not be ideal predictors. In early attrition studies, participants were sometimes asked to provide global percentages of L1 use or even answers to binary questions like Do you speak language X? (Ammerlaan, 1996; Köpke, 1999). These coarse-grained variables have been replaced by multifactorial questionnaires that distinguish between interlocutors and contexts (Schmid, 2011). Still, the quality of the data seems to depend as much on asking
the right questions as on whether speakers are able at all to make a reliable estimate of their own language use. Self-reported data have mainly been found to be a valid predictor of attrition when L1 use for professional purposes was concerned (Schmid, 2007; Schmid & Dusseldorp, 2010). It is, however, unclear if professional language use itself is the factor that can be related to attrition. Alternatively, L1 use on the job in an L2 setting – that is, in a situation in which both languages are highly active, but code-switching is typically inappropriate – may strengthen other skills, such as domain-general cognitive control and executive function, that are considered to be beneficial to language maintenance in bilinguals (Steinkrauss & Schmid, 2016). In addition to that, it may be the case that these skills have been more highly developed in the first place in speakers who do work that requires seamless and complete switching between languages.

Frequency of occurrence in the input, as an index of the entrenchment of specific words, is probably even more difficult to quantify than language use. A study on bilingual L1 speakers of German has shown that lemma frequency in an independent corpus was a poor predictor for patterns of use. Frequency information that comes from the same corpus as the data in question seemed to be more strongly correlated with over- and underuse of lexical items with a specific frequency (Schmid & Jarvis, 2014).

An advantage of the ATH over generative models lies in its applicability to a wide range of data. For the analyses that will be presented in the main part of this thesis, the common thread of ATH-based hypotheses would be that the amount of L1 use should emerge as a significant predictor of the performance of attriters on various measures. In all cases, L1 use is expected to be positively correlated with native-likeness – although the abovementioned caveats about a less than linear relationship between L1 use and other variables as well as the obvious quantification difficulties have to be kept in mind. The frequency of items, which has been used as a proxy for entrenchment, is less likely to play a role in the present studies: The phonetic analyses (Chapter 3) only deal with sounds that are fairly frequent in German. Similarly, the sentences in the EEG experiment (Chapter 6), which was designed to be suitable for L2 speakers with less than native-like proficiency as well, only uses highly frequent target nouns and verbs. For practical reasons, the corpus frequency of the formulaic sequences (Chapter 5) was not assessed. It is a disadvantage that the ATH – except for the brief reference to the regression hypothesis – makes few predictions with respect
to how strongly different linguistic subskills will be affected by reduced L1 use. It is clear, however, that production should be more strongly affected than processing and that information should attrite more quickly when it is stored in the declarative memory system rather than the procedural system. We can therefore hypothesise that attriters will be fairly similar to monolinguals in the EEG experiment, which draws on declarative resources, but only requires passive processing, but less native-like in the disfluency analyses, which are based on production data and can be seen as a holistic measure of skills that require smooth access to both declarative and procedural memory. The production of formulaic sequences is also predicted to be subject to attrition: For one, such sequences, like single-word lexical items, are assumed to be stored in declarative memory; in addition to that, the frequency of such expressions is lower on average than that of the words they are composed of.

1.3.3 Dynamic approaches: the Dynamic Systems Theory
The Interface Hypothesis (IH) and the Activation Threshold Hypothesis (ATH), however disparate they may be in many respects, have one thing in common: Both theories highlight one single factor – L1/L2 parameter mismatches in the IH, L1 use in the ATH – as the most important predictor of the outcome of various attrition processes. Other factors – for instance, motivation in the ATH – are considered as well, but clearly have a subordinate role. Also, investigations founded on generative or usage-based theories typically use data that have been collected at only one point in time.

As an alternative, it has been suggested to adapt Dynamic Systems Theory (DST), originally used in applied mathematics, for the description and prediction of the linguistic development in individuals. DST-based models, such as the Dynamic Model of Multilingualism (DMM; Herdina & Jessner, 2002), try to provide a perspective that integrates fluctuation and complexity more readily than conventional ‘reductionist’ approaches (Herdina & Jessner, 2013; Larsen-Freeman, 1997). Four aspects of such dynamic frameworks can be considered most important in the context of attrition research:

First, instead of focusing on one single explanatory factor, dynamic models emphasise the complete interconnectedness of all variables (de Bot, Lowie, & Verspoor, 2007), that is, the assumption that each variable can potentially be influenced by each other variable. Also, a change in the value of one variable may only become relevant when another variable changes as well: An early example of such
a pattern in attrition research was found in a study of Dutch emigrants to France that used length of residence in the L2 environment and L1 use as predictors for L1 proficiency in written and oral tasks (de Bot, Gommans, & Rossing, 1991). Interestingly, stronger signs of attrition over time were only found in participants who made little use of their L1. If the length of residence and L1 use had only been investigated individually, the relationship between these two factors might not have been discovered. Studying interactions between variables is clearly not a new approach, but proponents of dynamic models have argued that monocausal explanations and attempts to find linear correlations in the data are still more widespread in multilingualism research than they should be (de Bot, 2007).

A second aspect that is also not exclusive to, but has been emphasised by dynamic theories is the temporal aspect of language. It has been suggested that the language system of an individual speaker is in constant flux with variables having a differential impact on its development at different points in time. In dynamic theories, changes are not expected to be linear or similar across individuals (de Bot et al., 2007). The idea that language acquisition or attrition ever reaches an end state is therefore alien to these frameworks. Rather, growth and decline are seen as normal phenomena that occur as long as the system exists (Herdina & Jessner, 2002). In cross-sectional research, however, nothing can be learned about the way a system develops across the lifespan. Therefore, it has been proposed that the complex changes to the linguistic system of a speaker can be captured most adequately in longitudinal (case) studies (Schmid, Köpke, & de Bot, 2013). In such studies, it has been suggested, similar patterns of change as in biological systems may be found, in which growth tends to follow a path that can be approximated by a sine curve: exponential increase in the beginning, less steep growth over time with the curve eventually becoming flat. Changes in attrition – as an example of negative ‘growth’ – are predicted to follow an inverse sine curve. In post-puberty attriters, the highest level of the growth curve has been reached, so initially, reduced input and use would only result in minimally reduced proficiency (mirroring the minimal increases in the late stages of growth). In such a model, noticeable language loss would only be expected to occur later in time (Herdina & Jessner, 2002; Opitz, 2011).

A third aspect that is central to dynamic theories touches upon a methodological detail of attrition research. In many studies, bilingual speakers (attriters) are compared to monolingual speakers (controls) under the assumption that the linguistic
behaviour of the group of attriters as a whole before emigration was identical to the current behaviour of the monolingual group. This assumption carries a particular weight in dynamic theories, which postulate that the development of a system crucially depends on its initial state (Larsen-Freeman, 1997). Seemingly minor differences in the beginning can account for major divergence further down the road (the butterfly effect, a term based on the metaphorical anecdote that the exact time course and path of a hurricane may depend on the flapping of the wings of a butterfly several weeks earlier). This means that nonlinearity also extends to the relationship between cause and effect. Even if the initial state of proficiency in a now bilingual group was in fact similar to that of a still monolingual group, it cannot be excluded that small differences between the two groups may have spiralled into much larger differences under the influence of additional factors (de Bot, 2007). Again, longitudinal studies – preferentially case studies focusing on well-defined aspects of language change – could be a remedy to the baseline problem that attrition research. It has to be pointed out, however, that the scarcity of longitudinal research in linguistics is due to practical limitations that are difficult to overcome, such as the availability of long-term funding and the willingness of participants to remain available over an extended period of time. Also, longitudinal studies suffer from observer effects more than cross-sectional research: When experimenters repeatedly interact with speakers in their L1, the testing sessions themselves may make up a major portion of the L1 contact of the participants (Jarvis, 2003; Stolberg & Münch, 2010).

A fourth constitutive element of dynamic theories appears to be particularly relevant to language attrition. Even if DST-based approaches stress that individual language systems are constantly changing, it is recognised that they occasionally settle in so-called attractor states. These states, which are not necessarily predictable, are preferred over other states and constitute episodes of stability in a fluctuating system. Clearly dispreferred states that the system never settles into are called repeller states. A classic, non-linguistic example for attractor states is the way horses run: These animals do not trot or gallop in every instant they move, but seem to ‘snap’ into one of these two ways of running sooner or later (de Bot et al., 2007). It has been suggested that a fully developed L1 system is a powerful attractor state for the linguistic system of a bilingual speaker. Attrition, in this view, is seen as a series of temporary small shifts away from this state, caused by prolonged L2 influence,
but it can be assumed that the system has the tendency (and ability) to return to its original state (Schmid et al., 2013).

Generating hypotheses on why and when these small shifts will occur is not one of the strongest points of current DST-based theoretical approaches. Proponents of these frameworks have conceded that it is a “problematic characteristic [...] that language development cannot really be explained” (de Bot, 2007, p. 62) and that the scarcity of testable hypotheses based on dynamic models is an unsatisfactory state of affairs. For the time being, it has been suggested to focus on the accurate description of processes involved in attrition. For the studies that will be presented in the main part, the DST-based approaches can only provide us with a general idea of which patterns to expect: First, as has already been mentioned, the DMM predicts attrition to roughly follow an inverse sine curve, in line with the idea of biological growth in language development (Herdina & Jessner, 2002). However, the time course of the process is unclear, that is, it remains to be empirically determined when (and if) the initial phase of relative stability ends. Second, it has been suggested in the same model that each stage of language attrition will be expressed in the form of increased scatter of performance (Herdina & Jessner, 2002). This scatter will likely go unnoticed in everyday language use, but can be detected by explicit performance measures in the form of optionality or variability (Opitz, 2011). When comparing groups of attriters with groups of monolingual speakers, we may expect the scatter, for instance, to translate to larger standard deviations in the data of individual speakers. Seeing scatter – or any kind of variation – as an “inherent property of a changing system” (de Bot et al., 2007, p. 14) rather than as uninformative noise is a perspective that is shared by most dynamic models of language development and that seems relevant for the interpretation of attrition data.

1.4 Research questions and hypotheses

One general question that will be asked in this dissertation is: Can we detect changes in the L1 production and processing among a group of German natives in North America, compared to L2 speakers of German, who are bilingual as well, and L1 speakers who have remained monolingual? Based on the empirical evidence and the theoretical approaches that have been reviewed in this chapter, we hypothesise that differences between bilingual and monolingual speakers will indeed be found.
The more interesting questions, however, do not revolve around whether there will be changes, but around the following details of the changes we expect to find:

First, the panoramic approach that will be taken in this dissertation allows us to determine if different linguistic subskills are affected differently by attrition in our participant group. It is not easy to formulate a hypothesis with respect to the vulnerability of linguistic domains in attrition, given that the Regression Hypothesis – as the major theoretical approach in which detailed statements were made about this – can be considered refuted. Following usage-based theories, we can, however, predict that production will show stronger signs of attrition than processing and that linguistic information stored in declarative memory will be more strongly affected than information represented in procedural memory. Based on these predictions, one may establish a hierarchy of attrition according to which the strongest effects of attrition would be found in production tasks that draw on declarative resources, whereas processing that is mainly based on procedural memory would remain relatively unaffected by attrition.

Second, the role of the L2 in the shaping of changes to the L1 will be discussed. On a very general level, there can be no doubt that living in an L2-dominant environment for extended periods of time is the source of L1 attrition. It is not quite clear, though, whether all changes in the L1 can be directly traced back to properties of the L2. At least for the domain of morphosyntax, older generative approaches indeed suggest that transfer is one of the mechanisms of L1 attrition and predict that L2 influence can lead to changes in the settings of (previously unmarked) parameters of the L1. More recent generative frameworks, by contrast, have emphasised the role of interfaces between ‘narrow syntax’ and other cognitive domains, which is, however, difficult to apply to the data presented in this thesis. Proponents of usage-based and dynamic frameworks, while not excluding the possibility of direct transfer, have pointed out that bilingualism brings about changes that reach farther and deeper than mere transfer from whatever happens to be the L2 and the L1. It is claimed that the permanent presence of any L2 fundamentally and permanently alters the way the L1 is represented in the brain. If this proves true, we would expect to find L1 attriters perform similarly to other bilinguals, such as highly proficient L2 speakers, who will be included in some of our studies.

Third, it will be investigated if attrition has an influence on data patterns that are found within the group of potential attriters. As we have seen, dynamic theories
claim that bilingualism leads to a more varied performance, that is, a wider range between, for instance, the average measurements of all speakers within one group. This variability is attributed to the state of flux that the linguistic system of a bilingual is assumed to be in. In monolingual speakers, by contrast, we expect to see more group- and speaker-internal consistency. This means that even if we find similar group means and no significant overall differences between monolingual and bilingual speaker groups, there may well be differences in standard deviations, in the number of outliers or in the number of speakers who fall outside the range represented by the reference group. Again, the question will be asked if in this respect, attriters behave similarly to other bilinguals, such as L2 speakers. In this latter group, the presence and importance of systematic and non-systematic variability was recognised earlier and studied more thoroughly than in attriters (R. Ellis, 1985; Verspoor, Lowie, & van Dijk, 2008).

Fourth, we will study the association between linguistic changes in the attriters, if we find any, and factors such as length of residence in the L2 environment and L1 use. While these factors are not given great importance in generative frameworks, they are crucial for usage-based approaches to bilingualism, such as the Activation Threshold Hypothesis. According to this hypothesis, the amount of energy that is required to retrieve the mental representation of an item in one of the languages of a bilingual depends on the frequency and recency of the last activations, which predicts a more or less linear relationship between L1 use and manifestations of attrition. This, however, is at odds with at least part of the empirical evidence on attrition, which has only found a tenuous association between these two variables. If a clear relationship cannot be established in our data, alternative explanations will have to be considered.

It is striking that the hypotheses which can be formulated on the basis of existing theoretical approaches to L1 attrition are rather general in nature. Partly, this may be due to the fact that language is not a monolithic whole and that different linguistic domains seem to be governed by different principles. This renders it difficult to make specific statements about the language system as a whole, especially when – as some of the very theories in question predict – bilingualism has created a state of flux and variation. In any case, there is some reason to doubt that the existing theories of language attrition can account for the full range of phenomena found in bilingual L1 speakers. Therefore, the research papers that will be presented
in the main part of this thesis do not limit themselves to testing the predictions of generative, usage-based and dynamic frameworks. Rather, they were conceived as independent investigations into various facets of attrition, partly based on other theoretical approaches and focused on also testing more concrete, domain-specific hypotheses. Each of the papers stands on its own, having been either published in or submitted for publication to different peer-reviewed journals. Irrespective of this, the results of these empirical studies are potentially relevant data for theoretical approaches to language attrition. In the General Discussion, we will therefore discuss the more general implications of the results with respect to said frameworks.