1   Introduction

1.1   The study of language contact

Manifestations of language contact are found in a great variety of domains, including language acquisition, language processing and production, conversation and discourse, social functions of language and language policy, typology and language change, and more. This makes it a special challenge to compile an overview of the subject. Most introductory works devoted to contact linguistics have hitherto chosen to specialise either in the individual-synchronic aspects of bilingualism, or in structural-diachronic aspects of contact-induced language change. This book introduces an integrated theory of language contact, within which the study of these various domains can be bound together.

Since the launch of modern contract linguistics through the works of Weinreich (1953) and Haugen (1953), the study of individual bilingualism and of societal multilingualism has occupied a centre-stage position in the field. A testimony to this position is provided by a series of introductory textbooks that focus on one or both these areas, covering topics such as the acquisition of two languages from birth, bilingual language processing, diglossia and societal bilingualism, and language policy in multilingual communities (see Grosjean 1982, Hamers and Blanc 1989, Romaine 1989, Hoffmann 1991). Appel and Muysken's (1987) textbook was one of the first introductory works to take into account diachronic aspects of contact-induced language change. It was soon followed by Thomason and Kaufman's (1988) monograph, which remains one of the most influential and frequently cited works on language contact in the context of historical linguistics. Both these books put a spotlight on grammatical borrowing, and on the emergence of areal language clusters and of new 'contact' languages. Two further domains of investigation within contact linguistics have received attention in specialised introductory textbooks: the study of pidgins and creoles (Holm 1988–89, Arends, Muysken, and Smith 1995), and the study of code switching (Muysken 2000a, Milroy and Muysken 1995).

In recent years, several valuable contributions have appeared that aim to provide a state-of-the-art description of the field of contact linguistics or parts of it. Thomason's (2001) introduction to language contact emphasises historical linguistic aspects, including linguistic areas, language maintenance and shift, and contact languages. Winford (2003) is one of the first to combine a discussion

1

of codeswitching with an overview of historical aspects of contact, and Clyne (2003) combines a synthesis of other works on codeswitching and individual and societal bilingualism with a detailed discussion of the Melbourne corpus of immigrant languages. Myers-Scotton (2002b) outlines the Matrix Language Frame model of codeswitching and applies it to further phenomena such as language attrition, lexical borrowing, and the emergence of contact languages. The book is one of few attempts at a comprehensive discussion of contact phenomena within a specific theoretical framework. Further aspects of bilingualism, such as second-language acquisition and child bilingualism, societal multilingualism and language policy, and language processing and intercultural communication are covered in Myers-Scotton (2005).

1.2   Toward an integrated, functional approach to language contact

It is difficult to follow in the footsteps of the authors of these many insightful and inspiring works. My reason for wanting to add yet another book to the list of these fine introductions derives from a wish to strengthen the focus on a number of aspects in the discussion of language contact:

First, with the exception of Winford (2003) and Myers-Scotton (2002b), most textbooks continue to specialise in either synchronic (individual and societal) or diachronic aspects of language contact. Winford's book is an exception, as it devotes a chapter to a comprehensive and thorough discussion of codeswitching as well as accommodating a discussion of second-language acquisition. Missing from Winford's discussion are aspects of bilingual first-language acquisition and language processing, as well as an integrated theoretical approach that links the various domains. Myers-Scotton's (2002b) book is by contrast devoted entirely to introducing the Matrix Language Frame model, and does not pretend to cover the state-of-the-art in the individual fields to which the model is applied. The present book attempts to do both: To present the state-of-the-art in a wide range of sub-fields in contact linguistics, both synchronic and diachronic, and at the same time to offer a number of theoretical principles through which contact can be interpreted and appreciated in an integrated manner.

Second, it is my impression that much work has tended to focus on the implications of language contact to the inner coherence of language 'systems', while the perspective of the bilingual individual, which had stood so much in the foreground of Weinreich's (1953) work, seems to have been demoted. To be sure, this perspective is given much coverage in both conversation-analytical and 'rational choice' models of codeswitching (e.g. Auer 1984, Maschler 1994, Li Wei 2002, Myers-Scotton and Bolonyai 2001), in the recent direction in the study of child bilingualism (cf. Lanza 1997), and in some models of bilingual language processing (Grosjean 1998, Paradis 2004, Green 1998). But speakers' communicative

goals and intentions, their discourse strategies, and their language processing capacities are at the core of any speech production and so also of the structural innovations that constitute the seeds of potential language change. They therefore merit consideration when we set out to interpret processes of contact-induced change. 'Contact' is, of course, a metaphor: language 'systems' do not genuinely touch or even influence one another. The relevant locus of contact is the language processing apparatus of the individual multilingual speaker and the employment of this apparatus in communicative interaction. It is therefore the multilingual speaker's interaction and the factors and motivations that shape it that deserve our attention in the study of language contact.

Third, while interest in language contact has been on the rise among language typologists, and while a series of generalisations on the structural outcomes of contact have been proposed, tested, and discussed in the past, a typologically oriented framework of contact is still missing. The discussion in this book is informed by recent sampling of cases of contact-induced change. Taking into account processes observed in other domains of contact such as second-language acquisition and bilingual language processing, I propose some generalisations about the degree to which different structural components of language and different grammatical categories are 'vulnerable' in contact situations. The underlying assumption is that the language faculty is stratified and that the hierarchical behaviour of categories will reflect this stratification. In this respect, the study of language contact is of value toward an understanding of the inner functions and the inner structure of 'grammar' and the language faculty itself.

Fourth, no integrated approach to language contact has yet been formulated from a functionalist perspective. Such a perspective rests on a view of language as social activity and of communication as goal-driven. Consequently, it views speakers as actors who use language in order to achieve goals, and it attributes the selection of entire codes and of individual structures of language – constructions, word-forms, intonation, and so on – to goal-oriented activity. Moreover, it regards the structures, categories, and forms of language as triggers of linguistic-mental processing tasks that engage the hearer in communication. The dimension of the 'hearer' is therefore crucial to our analysis of linguistic 'categories' and their function; and the function of 'categories' is in turn regarded as central to their fate in various processes, from acquisition to codemixing in conversation, and on to structural borrowing and deliberate manipulation of language.

The theory of contact that is explicated in the following chapters is not a 'Theory' in the sense of a formal, self-contained, finite set of rules and principles that label, and pretend to be able to predict, each and every outcome of language contact. Rather, it is a theoretical approach that seeks to make generalisations about various manifestations of language contact, informed by, and embedded within, a broader understanding of language and communication. Such an understanding draws on a variety of sources, ideas, and works in linguistics. It includes a view of communication as part of a repertoire of social activities (Labov 1972a and 1972b, Hymes 1974, Gumperz 1980, Schiffrin 1987, Saville-Troike

1989), of communicative interaction as a repetitive form of human behaviour (Sacks, Schegloff, and Jefferson 1974, Sperber and Wilson 1986, Ehlich and Rehbein 1986),[1] and of grammar as the packaging of information in discourse (Givón 1984, 1990). It is based on the assumption that speakers' linguistic repertoires consist not just of formal rules and a lexicon, but of constructions (Goldberg 1995, Croft 2001), that language change is the product of innovation by individuals (Labov 1994, Croft 2000), and that speakers are creative and able to exploit meanings in new contexts, leading to the formation of new categories through 'grammaticalisation' (Heine, Claudi, and Hünnemeyer 1991, Hopper and Traugott 1993). It is also guided by an appreciation of 'pragmatics' as a method to uncover the very purpose and the inner function of structural categories, and not just to describe their casual employment.

My principal assumption in this book is that bilingual (or multilingual) speakers have a complex repertoire of linguistic structures at their disposal. This repertoire is not organised in the form of 'languages' or 'language systems'; the latter is a meta-linguistic construct and a label which speakers learn to apply to their patterns of linguistic behaviour as part of a process of linguistic socialisation. Rather, elements of the repertoire (word-forms, phonological rules, constructions, and so on) gradually become associated, through a process of linguistic socialisation, with a range of social activities, including factors such as sets of interlocutors, topics, and institutional settings. Mature multilingual speakers face a constant challenge to maintain control over their complex repertoire of forms and structures and to select those forms that are context-appropriate. Context-appropriate selection does not necessarily conform to a separation of 'languages': In some contexts, certain types of cross-linguistic 'mixing' and 'inserting' may be socially acceptable and may constitute effective goal-oriented communication.

Speakers' awareness of, and ability to implement social norms on the selection of elements within the linguistic repertoire is a central aspect of communication in multilingual settings. Awareness and the ability to control the repertoire may receive support from institutions and overtly articulated social norms and values concerning language. On the other hand, the language faculty presents itself as uneven with respect to the ease of control and selection of structures, as some language processing operations may escape the speaker's control more easily than others.

Communication in a language contact setting is the product of the interplay of two primary factors (Figure 1.1): Loyalty to a set of norms that regulate the context-bound selection of elements from the repertoire, and a wish to be able to exploit the repertoire in its entirety irrespective of situational constraints. The balance between these two factors is determined by a need to remove hurdles that stand in the way of efficient communication.

When loyalty prevails in a strict manner, then 'interference' or compromises are likely to be minimal. But when the wish to exploit the full repertoire is given some leeway, then strict context-bound separation of repertoire components might be compromised. Individual words that are usually reserved for interaction
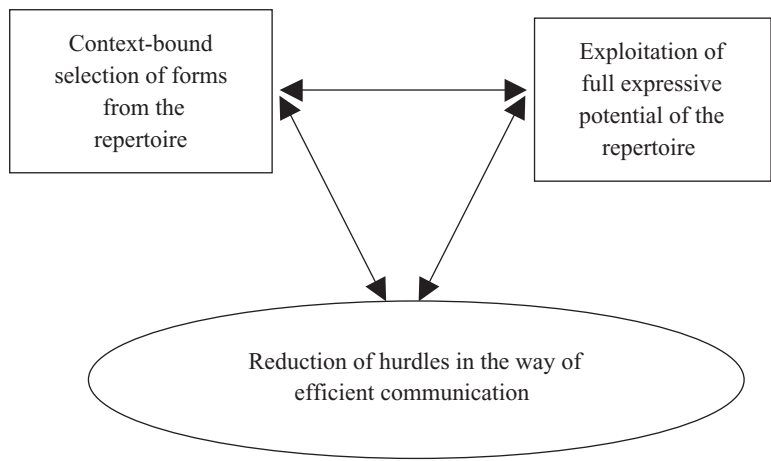
Figure 1.1 *The interplay of factors in communication in language contact settings.*

in Context set A might, for example, be employed ('inserted') also in interaction in Context set B. Second-language learners might draw on the phonology of their native language while communicating in a second language, bilingual children might employ constructions from one language that are not usually used in the chosen language of conversation, and adult bilinguals might insert discourse markers from one language when communicating in another. All this suggests that multilingual speakers do not 'block' or 'switch off' one of their languages when communicating in another, but that they have the full, complex linguistic repertoire at their disposal at all times.

The interplay of factors displayed in Figure 1.1 may lead to language change when a particular pattern of linguistic behaviour becomes widespread and accepted within a relevant sector of the speech community. Thus, an inserted word-form from another language may become a loanword, collective language-learning may show substrate influences in phonology (as well as in other domains of structure), the morpho-syntactic constructions of languages in bilingual communities may undergo convergence, and discourse markers from one language may be borrowed into another language. Contact-induced language change is thus ultimately the product of innovations that individual multilingual speakers introduce into discourse in a multilingual setting. Such innovations are, in turn, strategies that allow speakers to navigate between the two push-and-pull factors that we have identified: complying with social norms and expectations on context-appropriate selection of structures, on the one hand, and exploiting the full potential of the linguistic repertoire, on the other. From the point of view of their functionality, synchronic and diachronic manifestations of contact are therefore inseparable. Consequently, contact is not regarded here as an 'external' factor that triggers change, but as one that is internal to the processing and use of language itself in the multilingual speaker's repertoire of linguistic structures.

A final note is in order on the position of grammatical categories in the inves-
tigation of structural manifestations of contact. I follow a functionalist perspec-
tive that regards 'words' as more than just 'words': linguistic expressions and
constructions – whether bound or unbound morphemes, or morpho-syntactic
organisation patterns – trigger distinct types of mental processing operations.
These operations may be organised and retrieved in different ways, allowing
counterpart structures in the multilingual repertoire to be more or less easily
distinguished from one another, selected and controlled. The inner function of
grammatical categories therefore has a key role in explaining the behaviour of
that category in language contact situations. In typological perspective, sampling
reveals that there are some noteworthy differences between categories in respect
of their likelihood to undergo change as a result of contact. Hierarchical dif-
ferences among categories also appear in language acquisition and in language
mixing in conversation. In line with explanatory accounts in linguistic typology,
I take the view that re-occurring patterns of structural change and structural cat-
egorisation are not accidental, but that they are based in part on shared human
conceptualisations of reality, and more specifically on the foundations of manag-
ing and engaging in human communicative interaction. I will be paying special
attention to the role of categories and their functions, assigning similar outcomes
of contact not just to similar social settings and similar processes of identity
negotiation, but also to the role of categories in triggering and regulating distinct
language processing operations. In this respect, I will be assuming that contact
has not just a social dimension, but a communicative dimension, and that the
structure of grammar and the changes that it undergoes are a reflection of this
communicative dimension.

1.3   The structure of this book

Following from the interests and principles just described, this book
is intended to deliver an integrated discussion of individual and social aspects of
bilingualism as well as of processes of language change. As noted above, one of
the aims of the book is to restore the centre-stage position of the bilingual speaker
as a creative communicator in the perspective that we take when investigating lan-
guage contact. I therefore open with a chapter that introduces the 'Preliminaries'
of the emergence of a multilingual repertoire (Chapter 2). This chapter is a case
study of the early acquisition of language in a trilingual child. Its main purpose is
to illustrate how bilinguals develop a command of the repertoire along with the
skills to manage the interplay of factors depicted in Figure 1.1: compliance with
interlocutor expectations on the context-bound selection of elements from the
repertoire, on the one hand, and exploitation of the full repertoire, on the other.
Already at the early stages of language acquisition, the child speaker develops a
sensitivity toward expectations on context-appropriateness of word-forms. At the

same time, the compromises that are made are tightly connected to the referential and later to the language-processing function of the relevant word-form, and we obtain our first glimpse into the hierarchical nature of functional categories at work. As the child speaker matures, a balance between compliance with norms and exploitation of the full repertoire is struck through the application of a series of strategies, among them various types of language mixing as well as creative procedures such as the replication of constructions using context-appropriate word-forms. These strategies will accompany the speaker in adulthood and will constitute the core for individual innovations, which in turn constitute the foundation pool for potential processes of language change. On the whole, Chapter 2 thus gives us an overview of the potential effects of bilingualism in conversation, in language processing, and on language change.

Chapter 3 provides a brief examination of the principles by which components of the linguistic repertoire are mapped onto social activities, the triggers behind changes in this mapping arrangement, and the social and collective processes of intervention with the mapping of repertoire and activities. Chapter 4 returns to the principles of acquiring and maintaining a multilingual repertoire. I review current approaches to bilingual first-language acquisition, second-language acquisition, and bilingual aphasia, and discuss models of bilingual language processing and data on bilingual speech production errors. Three of the central themes of the book are strengthened in this chapter: The first is the proposal that multilingual speakers have at their disposal not 'language systems' that can be switched on and off, but an integrated repertoire from which elements are selected during each and every communicative task-schema. The second is the assumption that speakers are creative communicators, who will draw, if necessary, on the wide range of their repertoire in order to make communication more efficient. Evidence for this claim is found in the creative strategies employed by second-language learners. The third is the suggestion that the functional value of linguistic categories is a factor in speakers' ability to select structures within their repertoire; this makes certain categories more vulnerable during a failure of the 'selection mechanism' and so more likely to appear in non-voluntary language choices.

Chapter 5 examines the alternation of languages in conversation – codeswitching. We begin by interpreting the difference between codeswitching and borrowing as a gradient. Following a review of structural and conversational approaches to codeswitching, attention is paid again to the relevance of category affiliation to switching behaviour. The fact that speakers may exploit the contrast between languages or the complementary functions that they may have for a range of expression purposes is seen once again as confirmation that multilingual speakers have at their disposal an integrated, complex repertoire of linguistic structures.

Chapters 6, 7, 8, and 9 are all devoted to the effects of contact-induced language change that are most commonly known as 'borrowing'. In all four chapters, 'borrowing' is viewed as a form of levelling of structures across the multilingual repertoire, with the outcome that a single structure is employed, irrespective of interaction context and so irrespective of choice of 'language'. I start Chapter

6 by defining what I call the replication of linguistic Matter – concrete word-forms and morphs. I discuss the factors that motivate the replication of matter, and general statements that have been made in connection with the likelihood of borrowing to affect various grammatical categories. Chapter 7 discusses lexical borrowing and strategies to accommodate nouns, adjectives, and verbs, while Chapter 8 is devoted to the borrowing of grammatical structures and phonology. Chapter 9 focuses on the replication of linguistic Patterns – the arrangement, meaning, and combination of units of matter. I propose the model of 'pivot-matching' as an explanation for the creative procedure by which speakers avail themselves of a construction from within their repertoire, thereby exploiting the repertoire's full potential, and replicate it drawing on a constrained selection of linguistic matter or word-forms and morphs. Examples are discussed in a survey of categories and word classes, and space is given to the implications for language convergence and the formation of so-called linguistic areas.

Chapter 10, finally, is devoted to the emergence of new speech forms – 'contact languages' – in situations of language contact. I review current approaches to pidgins and creoles as well as mixed languages, and attempt to relate these products of language contact to the theme of multilinguals' creative use and exploitation of their repertoire: In the case of pidgins and resulting creoles, learners make efficient use of a variety of elements in a multilingual repertoire consisting of impressions of a target language, and a selection of structures from a substrate language, coupled with creative processes such as word-composition and grammaticalisation, in order to sustain communication in a new set of interaction contexts. Mixed languages are the product of a purpose-bound, conscious, and deliberate re-negotiation of the multilingual repertoire in a limited set of interaction contexts – usually those involving in-group re-affirmation of identity. The Outlook (Chapter 11) revisits the three central themes of the book: The constant availability and presence of a complex repertoire, the role of speakers as creative communicators and innovators, and the relevance of the inner stratification of the grammatical apparatus to processes of speech production and language change in language contact situations.

2 An emerging multilingual repertoire

2.1 A case study

The present chapter examines the emergence of the linguistic repertoire in an individual speaker in a multilingual setting. It traces the gradual development of constraints on the selection of structures within the repertoire and the acquisition of strategies to manage that repertoire. These strategies constitute the foundations on which bilinguals draw when alternating between languages. They also form the background and the pre-requisite for any contact-induced change. By surveying the bilingual child's strategies of managing the linguistic repertoire, we obtain a picture of the potential effects of language contact on speakers, on language use, and on the shape and structure of language.

I base this chapter on informal observations of the language acquisition process of a trilingual child, whom we shall call 'Ben'.[1] Born and raised in England in the late 1990s, Ben is exposed to two languages in the home: German, which he hears from his mother, and Hebrew, which he hears from his father. Both parents speak their respective languages consistently to Ben, consciously trying to avoid mixing. Between the ages of 0:4 and 4:4, input is balanced: During the first two years of his life Ben spends four days a week with an English-speaking child minder. He is cared for at home during roughly half of the working week primarily by his father, and during the other half primarily by his mother, while weekends are spent with both parents. At the age of 1:11, Ben's parents move into separate households, in separate towns. Ben stays primarily with his mother, spending three to four working days at an English-speaking nursery, while six days out of a fourteen-day cycle are spent with his father. Holiday time is spent equally with each of the parents. Most of the holiday time with the mother is spent in Germany, and around half the holiday time with the father is spent in Israel – in both countries with family and relations. On the whole, between the age of 0:4 and 4:4, Ben spends roughly equal amounts of time with each of the two parents (each speaking his/her language consistently) and at the English-speaking nursery, with exposure during holidays to monolingual contexts of German and Hebrew.

2.2   Lexical development

Ben's active language acquisition history begins at the age of 1:3. His first words are typically direct repetitions of words directed at him by a parent. For example, ['bada] follows the father's offer in Hebrew of *banána* 'banana', and ['ɛtɛ] follows the question in Hebrew *et-zé?* 'this one.ACC?' ('Do you want this one?'). At 1:4, words begin to appear on a more regular basis, and are no longer limited to direct repetitions. A number of onomatopoetic items are used, such as [baː] for 'sheep' and [ʔufʔuf] for 'dog', [pək] for 'toaster' and ['tita] for 'clock' (*tick-tack*). Just like the lexical words, they too are introduced by the parents. While some of these sound-symbols – *tick-tack* for instance – might be regarded as universal, others actually differ from more conventional language-specific baby-talk, for instance from German *mäh* for 'sheep' or Hebrew *hau-hau* for 'dog'. The sound-symbol [baː] for 'sheep' in fact originates in direct imitation of sheep during a countryside holiday. Some forms, such as [pək] for 'toaster', are entirely improvised. The onomatopoetic set thus constitutes a kind of 'family speech'. Significantly, both parents continue to use the same onomatopoetic sound sequence once it is established for a particular referent. The 'multilingual' child is thus exposed at this early stage in the development of his linguistic resources to a set of labels – 'words' – that are used by both parents,[2] alongside another inventory of words that are specific to each of the parents (i.e. 'proper' words belonging to each of the two languages). The child's active repertoire of 'proper' words contains from the very beginning items from German – e.g. [da] for German *da* 'there', [bal] for *Ball* 'ball', ['bada] for *Badewanne* 'bathtub' – and from Hebrew – e.g. ['baji] for Hebrew *garbáyim* 'socks', ['ʔəxa] for *yaréax* 'moon', [bajt] for *naaléy báyit* 'slippers'. Significantly, at this stage, only a single word is used actively per referent/object. There are, in other words, no active 'bilingual synonyms' – different words, deriving from different languages, which are used alternately to represent the same referent/object. This is well in line with general observations on early stages in bilingual first-language acquisition (see Chapter 4), as well as with a more general assumption that infants show a 'mutual exclusivity bias' in acquiring labels for referents (cf. discussion in Bloom 2000).

At the age of 1:6, Ben's repertoire already consists of an active vocabulary of around 40 words that are used regularly. In addition, the child is familiar with some ten names of persons (three of which are mainly used to refer to persons that appear on photos contained in a family photo album). By and large, lexical tokens that have been 'acquired' – meaning that they have been used actively by the child in communicative interaction and not just as one-off, on-the-spot repetition of adult utterances – continue to be used consistently by the child, irrespective of the identity of the parent-interlocutor. The child's active use of vocabulary tokens in any given situation thus consists a mixture of German- and Hebrew-derived items: we find for example [man] for German *Mann* 'man',