1 Introduction

1.1   What is code-switching?

In Alsace, in eastern France, French is commonly mixed with the local dialect, Alsatian, which is a variety of German (or more precisely Alemannic). At a family gathering in Strasbourg on New Year's Eve, a discussion starts regarding the poor quality of Alsatian butchers compared to those in the rest of France. One of the guests, Mr Eder,[1] a jovial middle-aged man and a prolific talker, holds forth:

Example 1

1	MR EDER:	**du bekommsch do e fätze** … je sais pas dans quelle graisse
		you get some sort of scraps … *in goodness knows what sort of fat*
2		… avec quoi: avec **de de de was weiss denn de teiffel**
		*… with what: with **the the the the devil knows what***
3		**noh geh i anne un! putz diss ding**
		then I have to go and clean the thing up
4		parce que lorsque tu as un morceau de viande **im** … **im teller**
		*because when you have a piece of meat **on** … **on your plate***
5		**un noochher hesch eso gschnuddels un muesch abschniede diss ganze ding gell**.
		and then you find you have a sort of mess and you have to cut the whole thing off you see
6		**oder e so hoch fett uf'm ding … diss haw i halt schliesslich a nitt gere gell?**
		or fat this high on top of it … I really don't like that at all you see

(Gardner-Chloros, 1991:124)

Mr Eder is a fluent speaker of both French and the Alsatian dialect. His apparent hesitations, represented by dots or repetitions, are found in stretches *within* the same language (L2, L4, L6) just as often as *between* stretches in different languages – their purpose is dramatic effect. No rhyme or reason appears to govern the points at which he passes from one language to the other. This form of expression in bilinguals has been called "mixed discourse" – to say that there are two separate languages is more or less meaningless from the participants' point of view. In L2, he also uses a "bridge", i.e. a word which could come from

---

[1] All proper names have been changed.

1

2   Code-switching

either language, to facilitate the switch: *de*, which can be a partitive in French or
a definite article in Alsatian.

Example 2
A second generation Greek Cypriot teenager, brought up and living in London,
Olga, told this story about why her father emigrated to England. The interviewer
spoke the Greek Cypriot Dialect so as to encourage Olga to use that variety.

1   INTERVIEWER: **iati irthan stin Anglia?**
                   why did they come to England?
2   OLGA: this is a long story [laughs]. **Itun mesa se ena mikro χorio, tin
                   E..[name]**
                                                   They were in a little village,
                   E..[name].
3   **ke o pateras mu – afto mu ipen o pateras mu. δen eksero … an
                   ine sosto**
                   *and my father – this is what my father told me. I don't know… if it's
                   right.*
4   **Mu ipe pos mian niχta emethisen ke pien ke epiasen mian…
                   pos tin lene**
                   *he told me.. that one night they got drunk … and they went and
                   caught a … what's its name.*
5   INTERVIEWER: **boris   na to pis sta anglika**.
                   You can say it in English.
6   OLGA: chicken chicken. Chicken [laughing] **pu enan γitonon**….
                                                        *which (belonged) to a
                   neighbour.*
7   **ke ton ivrasi   ke ipen pola. Endrapiken pola**.
                   *and they found him and said a lot (had a big discussion). He was
                   very ashamed*.
8   **ke mu lei   ia afton ton skopon irten   stin Anglia**.
                   and he tells me (that it is) for that purpose [sic] *he came to England.*
                                                   (Gardner-Chloros, unpublished data)

In common with many other second/third generation London Greek Cypriots,
Olga is much more at ease speaking English. She hesitates and searches for her
words when she has to speak Greek. In L2, she plays for time by saying, "This is
a long story" in English. Later, she cannot find the simple everyday word
'chicken' in Greek (L4), a fact which is indirectly commented upon by the
interviewer (L5), and there are several other signs of her difficulties with Greek,
such as her omission of the verb 'belonged' in L6. Mixing the two languages *is*
the normal way to talk in her community, but speaking to a purely Greek-
speaking interlocutor clearly taxes her competence in Greek.

    In Example 3, it may seem to the observer that a single variety is being used,
but those familiar with in-group communication in this community would
recognize that speakers are in fact alternating between different varieties.

This extract is from Sebba's *London Jamaican* (1993). Two London teenage boys of Jamaican parents, Andrew and Barry, are discussing an incident involving another young man, which occurred while Andrew was serving in a shop. Although Andrew's entire description may appear to be in a variety of London English, the passages in bold are in fact in Creole. In some cases, as in the word *Lucozade* (L5), it is only the pronunciation which identifies the word as Creole (and so obviously we are taking the researcher's word that this is what happens). In other cases (e.g. the words in bold in L11), other features also tell us that this is Creole. The passage in L19–20 in capitals is described as being in a near-RP "posh" voice.[2]

Example 3

1	ANDREW:	yeah man, I was on the till on Saturday (1.2) and this this black man **come** in (1.0) and (0.6) you know our shop, right, (0.6) they u:m (0.2) give (.) refund on (0.3) Lucozade bottles (0.4)
	BARRY:	m:
5	ANDREW:	a black man **come in an' 'im b(h)u::y** a bottle (.) of **Lucozade** while 'e was in the shop [an'
	BARRY:	[free p- e's e got free pee off is it?
	ANDREW:	yeah
	BARRY:	small ones or big ones?
10	ANDREW:	big ones and 'e drank the bottle in fron% of us an then ask(d) for the money back (see man) **me want me money now**
	BARRY:	[heheh
	ANDREW:	he goes (pnk) (I'm on) the till guy (.) hhh (I jus) (0.6) I jus' look round at 'im (0.6) I said well you can't 'ave it (1.9)
15		I said I 'ave to open the till (w) wait till the next customer comes (1.0) **'now! open it now and give me the money'** (1.0)
		I said I can't (0.8) the man **just thump 'is fist down** an' (screw up dis for me) (.) (s no man) the manager just comes (.)
		WOULD YOU LEAVE THE SHOP
20		BEFORE I CALL THE SECURITY: hh the man **jus' take the bottle an' fling it at me an'** (I) jus' catch it at the (ground)

                                                            (Sebba, 1993:119–120)

Sebba suggests that code-switching is used here to "animate" the narrative by providing different "voices" for the participants in the incident which is described. Although both the customer and the narrator might be expected to speak the same variety, either London English or Creole, Andrew reserves Creole mainly to quote the customer and to describe his actions (L17 and 20–21).

These examples show that the behaviour of bilinguals can only be properly understood with some insider knowledge of the community and the circumstances

---

[2] RP is "Received Pronunciation". Pauses are indicated as in the original, as are brackets showing overlapping speech.

4   Code-switching

where it is displayed. First, the speakers' *competence* in the relevant varieties may
or may not be a determining factor in their choices. Second, an observer may or
may not be able to distinguish which shifts in accent, vocabulary or syntax are in
some way significant for the participants in the conversation.

1.2   Why study code-switching?

Such varied combinations of two or more linguistic varieties occur in countless
bilingual societies and communities, and are known as code-switching (CS).[3]
It refers to the use of several languages or dialects in the same conversation or
sentence by bilingual people. It affects practically everyone who is in contact
with more than one language or dialect, to a greater or lesser extent. Numerous
local names designate such mixed talk: Tex-Mex, Franglais, BBC Grenglish,
Chinglish, Spanglish, Tuti Futi, etc. In some earlier periods of history, CS was
equally common in writing (see the papers in Trotter, 2002). Apart from CS,
there are a number of other possible linguistic outcomes of language contact
including borrowing, convergence, pidginization, language death, etc. CS has
been found to occur alongside most of these, though it does not necessarily do
so. The various manifestations of contact are grouped here under the heading of
language interaction.

1.2.1   Code-switching as a window on speech and language

The study of *why* and *how* people code-switch provides insights about many
aspects of language as well as speech. This applies not only to how language
and languages are organized in the brain (the mechanisms of switching as
such are discussed in Chapter 6). At a functional level, bilinguals often switch
varieties in order to communicate something beyond the superficial meaning
of their words. Monolinguals can do this also, by switching between dialects,
registers, levels of formality, intonations etc. (Bell, 1984; Coupland, 1985;
Labov, 1971; Kerswill, 1994).[4] "I can do *aught* when you're with me, I can
do anything", said a male speaker in his sixties from Sheffield in a Radio 4
interview, talking to his wife in an aside. *Aught* is Northern dialect, which
he then repeats in Standard English, *anything*. Such a switch serves at least
two functions: by using a dialect word, he emphasizes the fact that he is talking
to his wife rather than the interviewer. At the same time, he reinforces his

---

[3] Code-switching is sometimes found in the literature written as two separate words, sometimes
with a hyphen and sometimes as one word. Diachronically speaking, the move from two words to
hyphenated words to a single word reflects the semantic acceptability and integration of the
concept. I have stuck here with the intermediate solution, hyphenation.

[4] Register variation is a cover term for "the full range of language varieties associated with
differences in communicative situation" (Biber & Finegan, 1993:316).

closeness to her by referring to their common heritage. This is despite the fact that such a switch is less obvious to an observer than a change of language. The associations of different varieties are sometimes consciously manipulated, as in the case of advertisements, which often use CS into English to sell their products (Chapter 4): see the German McDonald's advertisement in Box 1.

The characteristic ways in which bilinguals combine their languages in a particular community constitute a way of expressing their group identity – like a characteristic accent. Both the languages themselves *and* the sociolinguistic environment play a role in the patterns which emerge. Comparing CS across different communities and different language combinations can help reveal the *relative* role of linguistic and sociolinguistic factors – an important issue in Linguistics. *Within* particular societies, sub-groups can be identified by their characteristic CS patterns, as monolinguals can by discourse styles and registers. CS therefore helps us to understand identity formation and expression in bilinguals (Tabouret-Keller, 1997; Sebba and Wootton, 1998).

Third, switching between languages provides crucial material for our understanding of how language is both comprehended (processed) in the brain, and produced. What are the clues in the words and sentences we pronounce which allow others to decode our meaning, and which we assemble in order to put across that meaning? When we observe how this is done with two or more languages, some of those features are thrown into sharper relief.

Fourth, by analysing code-switched speech, we can find out which combinations of words or morphemes from different languages can easily be combined and which are more resistant, or perhaps even impossible. Since grammar consists of the rules regarding such combinations, CS acts as a signpost, pointing at where the difficult issues may arise, and paving the way towards a better understanding of grammar. Romaine, for example, has pointed out that code-switching research can help us to understand a key issue in Linguistics: the division of labour between grammar and lexicon (1995). Grammar specialists interested in CS try to discover whether the grammatical rules of the two varieties in contact are sufficient to explain the patterns in mixed language speech, or whether mixed codes have additional rules of their own.

All in all, CS is informative about language at a number of different levels. There are also good reasons to study it in its own right.

1.2.2 Studying code-switching for its own sake

It seems sensible that linguists should derive their data and evidence from the most typical speakers rather than from more exceptional ones. Numerous linguists have pointed out that most of the world is plurilingual. If you add together people who live in multilingual areas of the world (Africa, India, Singapore, Creole-speaking areas such as the Caribbean or Papua New

---

Box 1 Code-switching in advertising

Advertisements are a very common locus for CS, and examples can be found from round the world where English is combined with the local language in order to evoke a cosmopolitan – or American – lifestyle. This advertisement appeared as part of a code-switched series for McDonald's in Germany in the early 2000s. Note the use of many "bivalent" words (*croissant, warm, so, in*), reinforcing the other similarities between German and English; the use of German-style lower-case 'a' for the adjective *american*; and the use of the masculine pronoun *he*, as in German, to designate the ham.
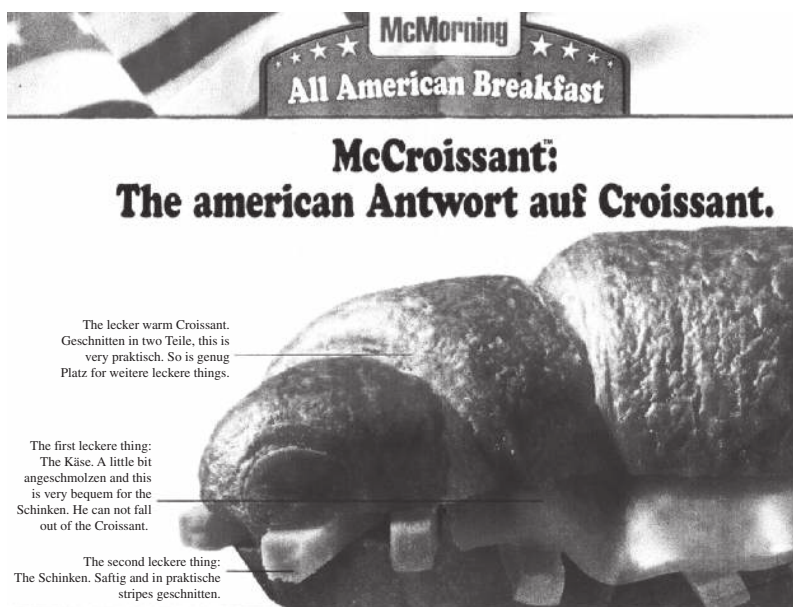
Translation

McCroissant: the American **answer to Croissant**.

The **delicious warm/**warm **Croissant**.
Cut in/in two **parts**, this is very **practical. So**/so is **enough room** for **further delicious** things.

The first **delicious** thing: the **cheese**. A little bit **melted** and this is very **comfortable** for the **ham**. He [sic] can not fall out of the **Croissant**.

The second **delicious** thing: the **ham**. **Juicy** and cut **in**/in **practical** stripes [sic].

Guinea, etc.); people who speak a regional language or dialect on top of a national language (from Basques to Chechens); and migrants and their descendants (Greeks in Australia, Punjabis in Britain, Spanish speakers in the USA, etc.), you are left with small islands of monolingualism in a multilingual sea. This is without counting people who learn a second/third language beyond a basic level at school (e.g. the Dutch or Scandinavians); those who have a different language for literacy from the one they speak (e.g. Gujerati and Punjabi speakers whose language of literacy is Hindi); those who become bilingual through changes in personal circumstances; and those whose mother tongue is not considered adequate for formal purposes (i.e. in diglossia) and who therefore have to master another variety in order to take part in official life (e.g. Flemish speakers in Belgium, speakers of dialectal Arabic in various Arab countries) (Baker and Prys Jones, 1998). Plurilingualism is still the norm in spite of the fact that a large number of the world's languages are under imminent threat of extinction owing to economic and globalizing forces (Crystal, 2000).

Most of these plurilingual speakers mix their languages in various ways in their daily lives. CS has been studied from Mexico to Kenya and from Finland to Australia – one of the main problems in writing this book has been the difficulty of doing justice to the profusion of work which has been done. In line with the eclectic approach to CS adopted here, a balance has been attempted between giving up-to-date, accessible references and older/less accessible, but seminal ones.

1.3   A common-sense approach

The approach to CS adopted here can be described as "common sense" or as pragmatic with a small "p". "Pragmatic" research on CS with a large "p" – discussed in Chapter 4 – focuses on the conversational functions of CS and its effects on conversational participants (Auer, 1998b). CS is presented here in a rounded manner, looking at work carried out from the Sociolinguistic, Psycholinguistic, Grammatical and Acquisitional perspectives as well as the Pragmatic. CS is taken at face value, rather than with a particular theory as the point of departure. It is important that CS be considered as the multifaceted phenomenon it is, rather than purely as a means of testing theoretical positions.

So far, research on CS has been fragmented within various sub-disciplines, yet there are considerable advantages in considering it as a whole. There is an analogy with the poem at the beginning of this book, which recounts an old anecdote about six blind men feeling different parts of an elephant, and being unable to gain an overall view as to what it was. The terminological discussion below (see 1.5) illustrates how little agreement there is about CS, its definition and limits. Until greater concensus emerges, we should continue to look at it from as many different angles as possible.

One reason for this is that each of the sub-disciplines in Linguistics uses different methodological approaches. Sociolinguists seek to record "natural" conversations but are subject to the Observer's Paradox; theoretical linguists use fallible intuitions as to which sentences are correct/acceptable; psycholinguists' experiments test isolable – but incomplete – skills. None of these methods on its own can provide a complete picture of behaviour as complex as CS.

Moreover, assumptions underlying tried and tested methodological para-digms are often insufficiently discussed in Linguistics. For example, variation-ist descriptions of CS are still current in the grammatical field (Poplack, 2000). These are based on the assumption that if we can account for the majority of cases of CS, then we are justified in ignoring the minority of cases which do not fit in, and which are dismissed as aberrations. But one of the most famous philosophers of science, Popper (1959), considered that a proposition which was not falsified by counter-evidence was not scientific. Although there have been a number of alternative proposals since then as to what constitutes proper scientific enquiry, Popper's test remains one of the most rigorous. At what point do counter-examples to a paradigm which has been put forward make it necessary to revise it? There has so far been a lack of discussion of this problem with regard to CS, despite the fact that several scholars have argued that, in the present state of knowledge, we should be trying to formulate grammatical tendencies rather than absolute rules (Jacobson, 1998b; Muysken, 2000). As Radford (1996:81) wrote: "Many advances in our understanding of adult syntax have come from probing the syntax of structures … which any computer corpus would show to be extremely 'rare' … 'Every example counts!'" Tracy has also pointed out, "What one counts as an exception is not just defined by some quantitative feature but by the fact that it lies in conflict with our theory" (1995:198).

Another example from the grammatical field is the widespread belief that all bilingual utterances have an underlying "Matrix Language", i.e. a grammatical template which can usually be identified with a particular language, such as Russian, and that CS consists in grafting material from another identifiable language, such as Spanish, onto such a base. This belief has been around at least since Weinreich wrote that "Each utterance is in a definite language" (1953:7). Such a belief makes the grammatical description of mixed language utterances much simpler, but fails to address the more fundamental question of what we mean by a language in the first place.

Le Page's view is that one of the principal tasks for linguists is to explain the formation of the concept of a homogeneous "language": "We set out how we saw such a concept evolving from observation of discourse, through the stereo-types denoted by such language-owning names as 'English' or 'French', to that of the most highly abstract and focused Chomskyan 'grammar'; and how actual linguistic behaviour was influenced by the stereotypes as progressively it was

named, formalized, standardized, institutionalized, and totemized by a society" (1997:31–32). "Languages" are often treated as if they were discrete, identifiable and internally consistent wholes, and we forget how historically recent and culturally selective such a view is (see Chapter 2).

The study of CS should force us to think "outside the box": to review methodologies, theoretical approaches and assumptions, often developed in a monolingual context, and see how they stand the test of being applied to the speech of bilinguals. A common sense approach involves recognizing exactly what are the limits to our ability to generalize at any given stage of the enquiry.

1.4 The study of code-switching

For a long time, CS was scarcely noticed by linguists writing about language contact. Milroy and Muysken (1995), who describe it as "perhaps the central issue in bilingualism research", point out that research on CS was slow to start compared with, say, research on borrowing or what used to be termed *interference*. In the seminal *Languages in Contact* (1953), Weinreich referred to the "transfer of words" from one language to another by bilinguals, but dismissed this as a "mere oversight" (1953:73–74). Haugen, writing at around the same time, also apparently overlooked the significance of CS, and wrote that: "The introduction of elements from one language into the other means merely an alteration of the second language, not a mixture of the two" (1950:211).

Over the last forty-odd years, there has been an explosion of interest in CS. CS had remained more or less "invisible" in research on bilingualism until the work of Gumperz and his associates in the 1960s and early 1970s (Gumperz, 1964, 1967; Gumperz and Wilson, 1971; Blom and Gumperz, 1972). Thereafter the subject took off – and there has been no sign of a downturn – as people realized that CS was not an isolated, quirky phenomenon but a widespread way of speaking. But research in this field is complicated by the multilayered significance of CS. Each new case which is documented can be looked at from multiple perspectives, so from the outset, a certain depth of engagement with the data is necessary.

Furthermore, by definition, studying CS implies dealing with several languages. Grasping the significance of a transcription where the reader or researcher is not familiar with one or both of the languages involved can be off-putting. This problem should be somewhat reduced in the future by various technical developments of use to the linguistic researcher, such as standardized transcription and coding systems, sound–text linking, and the possibility of collaborating on and sharing data over the Internet. Proposals for a system appropriate for CS are summarized in the Appendix (LIPPS Group, 2000; Gardner-Chloros, Moyer and Sebba, 2007). Because of the huge interest in CS on the one hand, and the difficulties of studying it on the other, a lot of work has crystallized around a few main approaches:

(1) *Sociolinguistic/ethnographic descriptions of CS situations*. These represent the majority of studies of CS. Although by their nature, they remain fairly fragmented, many important insights are derived from linking the manifestations of CS to aspects of the sociolinguistic situation (Chapter 3).

(2) *Pragmatic/conversation analytic approaches*. These rely on identifying the meanings brought about by CS in conversations, for example through following, or avoiding, the language choices of interlocutors (preference organization). This use of CS complements the exploitation of contrasting connotations of the two varieties (e.g. we-code/they-code). Such tactics may be used in the same conversation (e.g. Milroy and Gordon, 2003; McCormick, 2002) (Chapter 4).

(3) *Grammatical analyses of samples of CS and the search for underlying rules, models and explanations to explain the patterns found*. These have developed largely as a separate tradition from the sociolinguistic and the pragmatic. Although some authors have identified connections which deserve to be investigated (Muysken, 2000; Myers-Scotton, 1993b), this has not been a primary focus in the research so far (Chapter 5).

Each of these approaches is the subject of one chapter, as is the place of CS in language contact (Chapter 2); the implications of psycholinguistic work on bilinguals for our understanding of CS (Chapter 6); and CS in children and other learners (Chapter 7). The chapter pattern reflects the main research output and traditions, but within each chapter it is emphasized that there are no strict divisions between the questions which should be addressed in CS research – on the contrary, a major purpose of this book is to encourage the formulation of more holistic insights and research.

1.5   The vexed question of terminology

In the introduction to a volume on CS, Eastman wrote: "Efforts to distinguish code-switching, code-mixing and borrowing are doomed" (1992:1). Little has occurred since then to lighten this pessimistic view: terminology has been endlessly discussed in the CS literature without any real commonality of practice being achieved. Several good descriptions are, however, available of how the most important terms have been used in this field (Milroy and Muysken, 1995; Li Wei, 2000; Hamers and Blanc, 2000; Clyne, 2003), the key issues being described in sections 1.5.1 and 1.5.2.

1.5.1   A misleading term?

CS is not an entity which exists out there in the objective world, but a construct which linguists have developed to help them describe their data. It is therefore