Thomason and Kaufman’s 1988 book *Language contact, creolization, and genetic linguistics* had a stimulating effect on the fields of comparative and descriptive linguistics and inspired a number of studies on various topics related to language contact: the relationship between typology and language contact; the effect of language contact on a language’s genetically inherited characteristics, and work on mixed and endangered languages. More generally speaking, the increased availability of data relating to language contact has enabled wider-ranging discussion on the nature of language contact and its consequences (see Hickey 2010 for a more detailed account of these subjects).

Within this landscape, our book lies at the crossroads of the following themes:

1. vulnerable and endangered languages, since some of the languages described here are minority languages losing ground under the linguistic influence of dominant neighbouring languages (see chapters on Cham, Wa); 2. areal typology, since our book is concerned with one area in particular: Mainland Southeast Asia (hereafter MSEA); 3. language contact and genetic affiliation, since the various grammatical sketches lay emphasis on characteristics shared by unrelated languages. This in turn raises the question of how such traits are acquired and how they spread, though neither of these two issues is addressed in this volume, at least from a typological perspective.

Specifically, we address here the issue of linguistic area or *Sprachbund*. This issue is closely associated with language contact and has been discussed extensively over the past fifteen years (see Thomason 2000, 2001; Muysken 2000; Stolz 2002; Heine & Kuteva 2005; Campbell 2006; Bisang 2006a; 2006b; 2006c; Matras et al. 2006; Matras & Sakel 2007; Muysken 2008; and Bisang 2010, *inter*
Linguistic area, or Sprachbund in German, is a concept which was introduced to linguistics in the 1930s by NS Trubetzkoy, in response to the need to account for the linguistic situation observed initially in the Balkans (Sandfeld 1930) and in India (Emeneau 1956; Masica 1976 etc.). In each of these regions both languages with a close genetic relationship and languages from different language families were found. However, despite lacking a common origin, these languages had surprising structural similarities, apparently acquired in part through contact with structural linguistic features that they did not originally possess.

This concept of linguistic area has triggered much debate among linguists interested in language contact. For some authors (e.g. Aikhenvald & Dixon 2001) the difficulty in distinguishing what is inherited through parentage from what is diffused through contact casts doubt on the possibility of establishing genetic parentage in language contact situations, for it may be impossible to determine whether shared traits arise through inheritance, distribution, independent parallel development or by accident.

For other linguists, the concept of linguistic area is viewed in the light of more general work on language change and the constraints on language change associated with language contact (e.g. Gilbers et al. 2000; Heine & Kuteva 2005; Aikhenvald & Dixon 2006b; Matras et al. 2006; Siemund & Kintana 2008). Studies in recent years concur on the idea that a linguistic area is not a purely linguistic phenomenon, but also brings into play history and culture (Aikhenvald & Dixon 2001: 11–13; Dahl 2001: 1458, Thomason 2001: 104).

“[Linguistic areas] arise in any of several ways—through social networks established by such interactions as trade and exogamy, through the shift by indigenous peoples in a region to the language(s) of invaders, through repeated instances of movement by small groups to different places within the area.” (Thomason 2001: 104)

In other words, the conditions which give rise to language contact (bilingualism, diglossia) are not sufficient in themselves to cause the emergence of a linguistic area. The linguistic communities involved must also share a common culture or a common history—and they are generally aware of this fact. Thus, even if it is not possible to determine when changes in languages occur, it may be possible to identify the factors which favour the emergence of a linguistic area, namely extralinguistic socio-linguistic factors such as culture or social organization, community type, history, politics, geography (for instance ‘spread zone’ vs. ‘residual zone’ – cf. Dahl 2001: 1461), population density and diversity, etc.

Finally, there is much published literature on the status of certain geographical areas where languages come to share common properties without being closely related to one another, such as the Balkans (van der Auwera 1998),
Meso-America (Stolz & Stolz 2001), the ‘Baltic area’ (Koptjevskaja-Tamm 2006), Africa (Heine et Nurse 2008), as well as South Asia (Masica 1994, Ebert 2001).

In recent decades a number of definitions of a linguistic area or Sprachbund have been proposed (see Campbell 2006; Stolz 2006: 33), all of which aim to describe the phenomena of linguistic convergence, common innovations or common retentions, or to identify the specific properties which set a Sprachbund apart from other language-contact situations. In general, such accounts invoke the same key concepts, namely (1) a geographical area; (2) the involvement of a number of languages (at least three); (3) shared linguistic characteristics; (4) convergence as a result of contact; (5) convergence not by accident; (6) convergence not as a result of shared heritage (Muysken 2008: 3). However, despite numerous attempts to define the concept precisely, a consensus emerges on the impossibility of identifying universal criteria (Stolz 2002, 2006; Bisang 2006c). Some linguists go so far as to suggest the outright abandonment of the term of Sprachbund, referring as it does not to a real object but to a projection by linguists (Stoltz, 2002: 260), an a posteriori construction “based on the accumulation of residue and borrowed traits, regardless of how and when they came to be shared among the languages involved” (Campbell 2006: 14).

Meanwhile, faced with an unsatisfactory definition which is unlikely to produce a concrete generalization (see Dahl 2001: 1457–8), and because of the impossibility of distinguishing between a situation of linguistic borrowing and one of formation of a linguistic area (Sakel & Matras 2008), other linguists have favoured replacing the notion of linguistic area with less constraining concepts such as ‘contact superposition zone’ (Koptjevskaja-Tamm & Wälchli 2001) or ‘zones of contact-induced structural convergence’ (Bisang 2006c).

The problems encountered in seeking to define a linguistic area, however, do not detract from the relevance of studies of changes induced by contact between the languages spoken within a particular geographical area, i.e. areas of linguistic convergence. In the case of Southeast Asia, a region characterized by the presence of five language families and several millennia of contact between the area’s linguistic communities, an areal approach is fruitful both for the description of undescribed languages and for typological studies.

2 Establishing the existence of the Mainland Southeast Asian linguistic area

Assessing the extent of our knowledge of the languages of Mainland Southeast Asia, we notice that recent publications on Asian languages are often geneti-
cally oriented, with the exception of Goddard (2005), a nice overview of the linguistic situation in Asia with partial information on each language. A number of recent publications do not address the MSEA Sprachbund, for instance Thurgood & LaPolla (2003) on Sino-Tibetan languages; Adelaar & Himmelmann (2005) on Austronesian languages; Diller, Edmondson & Luo (2008) on Tai-Kadai languages, and Jenny & Sidwell (2015) on Austroasiatic languages.

The Mainland Southeast Asian Sprachbund inhabits a geographical area stretching from the easternmost fringes of India in the west to China in the east, encompassing the peninsular Southeast Asian states of Burma, Thailand, Laos, Cambodia and Vietnam, as well as peninsular Malaysia. Five different language families are present in the area (Austroasiatic/Mon-Khmer, Tai-Kadai, Hmong-Mien, Sino-Tibetan and Austronesian). It is only relatively recently that this area has been described as a linguistic area or Sprachbund (Matisoff 1991, Bisang 1996, etc.), even though now the idea is firmly attested (and uncontroversial) since Enfield 2005.1 Thus, new general publications on linguistic areas do cite the (M)SEA area as one of the recognized Sprachbünde (see for instance Heine & Kuteva 2005: 203), although it was not mentioned in previous works (see Campbell 1994, Feuillet 2001).

In the final analysis, we can say of a linguistic area that “in the absence of a universally valid numerical value of shared isoglosses, language contact situations lend themselves to a classification as a Sprachbund if the absolute number of shared isoglosses with no genetic basis among the members of the Sprachbund significantly exceeds the number of such isoglosses they have in common with languages outside the Sprachbund” (Stolz 2006: 36).

We have adopted an empirical and – we hope – more promising approach towards the study of the MSEA linguistic area. Previous studies such as Matisoff (1986), Migliazza (1996) and Enfield (2005) draw together features that cut across the genealogical phyla in many domains. These works cast light on phenomena or structural properties that cluster around this geographical area, offering an explanation for the distributional asymmetry that cannot be accounted for in terms of linguistic genealogy.

By adopting an approach that is similar to an isogloss approach rather than a geographical, historico-cultural or communicative approach (Stolz 2006: 36), we aim to determine the maximal distribution of single features. To put it in other words, one task of this book is to provide new information about the limits

---

1 Enfield (2001, 2005) provides a very well-documented argumentation even though he is not alone in having pointed out the strange similarities between unrelated languages in the area. See also Henderson (1965), Matisoff (1986: 75–80), Matisoff (1991), Bisang (1996), Migliazza (1996).
of the contact-induced convergence area (see for instance the chapters on Yongning Na (Mosuo), Khumi or Southern Min) in Southeast Asia, examining the geolinguistic distribution of the given features (or zone of overlap of several such features) that shape the linguistic area.

3 Structure of the book

The structure of this book is intended to be in line with existing typological studies of specific grammatical phenomena such as Kahrel and van den Berg's (1994) work on negation, or cross-linguistic studies such as Aikhenvald & Dixon (2003, 2006), Zúñiga & Kittilä (2010), and Kopecka & Narasimhan (2012) inter alia. Each of these publications is a collection of chapters which adopt a common format, structure and/or theoretical approach.

The present volume on languages of MSEA has been conceived in a similar way, to allow researchers to do cross-comparisons and to facilitate such comparisons by ensuring that all the chapters have a broadly similar organization and structure and use similar terminology. Our aim in adopting this common approach is to allow specific linguistic phenomena to be studied across a range of languages.

Each contributor was asked to compile a grammatical sketch of a MSEA language following the same guidelines, to allow the reader to navigate easily between and across chapters and languages. A copy of the guidelines which each author was asked to follow is included as an appendix at the back of the book.

Thus, the typological descriptions of the languages are intended to have a common structure to facilitate comparison and to highlight, on the one hand, the unique typological features of the language and, on the other, the features shared with other languages in the area.

Beginning with general information about the language — its affiliation, its geographical location, relevant ethno-linguistic information—the chapters then provide information on (I) phonology, (II) morphology, (III) grammar and syntax, (IV) semantics and pragmatics. Each section examines in turn the features known to be shared by the languages of the Sprachbund. Thus, in the section on Phonology, the stereotypical phonological features observed in languages of the area are as follows: complex vowel systems (diphthongs, a large number of vowels, contrastive vowel length), tone or register systems (or a combination of both), a restricted set of final consonants; a restriction on consonant clusters (see Enfield 2005: 186ff). Each contributor was asked to address the question: “To what extent does the language conform to this stereotype (or not)?” by pro-
viding a description of the phonological system and syllable structure, keeping in mind (as a secondary guideline) the correlation between the two as highlighted by Henderson (1965).

In the second section of each chapter, which is on Morphology, the contributors were asked to examine the morphological structure of words, the generally observed tendency toward monosyllabicity, the presence of sesquisyllables, the lack of extensive inflectional morphology, and the use of compounding (rather than derivation). Expressives (psycho-collocation, elaborate expressions) and reduplication are also phenomena examined as shared features across the MSEA area.

The third section, on Grammar and Syntax, is divided into three parts devoted, respectively, to (1) the nominal domain, (2) the verbal domain and (3) the clausal domain. Besides the basic structure of the NP, we also asked for a close investigation of classifier systems (see Grinevald 1999, 2000). Next, MSEA languages being famous for their lack of inflection, the description of the verbal domain relies on notional categories expressed in the verbal phrase, grammaticalization (see Matisoff 1991), and serial verb constructions (see Bisang 1996, Durie 1997, Aikhenvald & Dixon 2006, Vittrant 2006). In terms of sentence organization, the basic word order is (S)VO for MSEA languages, except for Tibeto-Burman languages in the area. All languages in the area show a characteristic information structure, being ‘topic-prominent’ rather than ‘subject-prominent’. All languages in the area are also known for widespread ellipsis of definite arguments, regardless of grammatical role. These phenomena are investigated in each language.

Marybeth Clark (1985) also noticed that many MSEA languages share a similar device for asking questions that request an affirmative or negative response rather than other information, that is to say for wh-questions or yes-no-questions. Matisoff (1986: 78) noticed that MSEA languages have a penchant for nominalizing whole sentences without embedding them into any larger unit, typically via a particle, which is also used in citation-form verbs, and which has a relativizing/genitive function in other constructions. These two phenomena are examined by some contributors to the present volume (Chappell, Lidz, Peterson).

Regarding the last section on Semantics and Pragmatics (IV), MSEA languages seem to share basically similar conceptual frameworks about humans and nature (Matisoff 1986: 79). Thus the comparative semantics domain needs to be investigated. As for pragmatics, MSEA languages have systems of sentence-final particles as a basic mode of distinguishing illocutionary force (such as requesting, questioning, persuading, advising, reminding, instructing, etc.), but also for expressing ‘propositional attitudes’, that is to say emotions of the speaker (such as surprise, doubt, impatience, reluctance, hesitation, etc.). Lastly, some languages may have developed systems for encoding politeness (i.e. formality vs. intimacy with an interlocutor) and honorific systems (to ex-
press respect or reverence towards the referent). All these phenomena are explored in each of the thirteen languages described.

Each grammatical sketch ends with two appendices: a glossed text, and a table summarizing the features examined in the chapter. This glossed text offers a glimpse of the language used in more natural context and a demonstration of the broader expertise of the linguists who compiled them. A series of maps at the end of the book show the geographical distribution (in terms of the location of the language described in this book) of a number of the more linguistic features explored in the tables.

In conclusion, thirteen languages of MSEA are described in this collection. As with any linguistic survey, some explanation of why we chose that number and that particular selection of languages is required.

The languages in our sample were selected for a number of reasons. In part, languages such as Khmer, Wa, Vietnamese, Thai and Hmong are representative of the Mainland Southeast Asian peninsula, which is known to be the geographical core of this linguistic area, from which certain features diffused outwards. In contrast, languages such as Burmese, Khumi, Mosuo and Min are spoken on the periphery of the same area.

Any examination of the traits shared by the languages of Southeast Asia must also include some examples of languages which are representative of other contrasting phenomena, such as:
(a) size of speaker population – i.e. major or national languages (Malay) vs. minor or marginal languages (Cham);
(b) dominant lowland languages (Vietnamese or Khmer) vs. languages of highland minorities (Wa);
(c) languages with a long-established literary tradition (Burmese) vs. languages with a predominantly oral tradition or which are unwritten or seldom written (Mosuo).

These are also factors which have determined the choice of languages in this collection.

Finally, we would have preferred to be able to include, for each language family, at least one major and one minor language, which has unfortunately not been possible for two of the language families, namely Tai-Kadai or Hmong-Mien (a family whose very size makes an objectively major language difficult to identify).

The chapters are organized according to language family: five Sino-Tibetan languages; four Mon-Khmer (Austroasiatic) languages; two Austronesian and one each of Tai-Kadai and Hmong-Mien.

We might reasonably be accused of arbitrariness in our selection of languages; in fact our choice was intended to be neither comprehensive nor sys-
tematically representative of the area. Rather, we wanted to give an impression of the colours and flavour of the region, based on linguistic commonalities and differences. Given that MSEA is a recognized linguistic area (or, a zone of contact-induced structural convergence) in the literature on language contact and Asian languages more generally, our goal was not to prove its existence. Rather, it was an attempt to further explore the boundaries of the area and the path of diffusion of shared linguistic features.

At the same time, the diversity represented in our selection of languages entails a concomitant variability in the depth of the descriptions in the chapters of this book: for certain languages, the authors have the benefit of a wealth of previous linguistic investigations, while for others the authors are amongst the first linguists to describe the language in question. In all cases, however, the authors are actively involved in research on the languages they have described and the data is not taken from secondary literature.

On that note, we are very grateful to the consultants who have collaborated with the authors of each chapter. Without such collaboration or the readiness of linguists to engage in fieldwork on this kind, a book such as this would not have been possible. Lastly, we hope to inspire further work on the many languages of Southeast Asia which have yet to be described.

References


Emeneau, Murray B. 1956. India as a linguistic area. Language 32. 3–16.


