1 Introduction

Bernard Comrie and Greville G. Corbett

The Slavonic languages (usually called the Slavic languages in the United States) are the major languages spoken over most of eastern and much of central Europe, as indicated in map 1.1 on page 2. The Slavonic language with the greatest number of speakers, Russian, has spread, as a result of gradual expansion, from its original heartland in eastern Europe across most of northern Asia to the Pacific coast. The parts of eastern and central Europe where Slavonic languages are spoken are areas of great current political interest, with the emergence of new experiments in democracy, economic organization and artistic expression in societies whose recent history has been primarily one of tight centralized control.

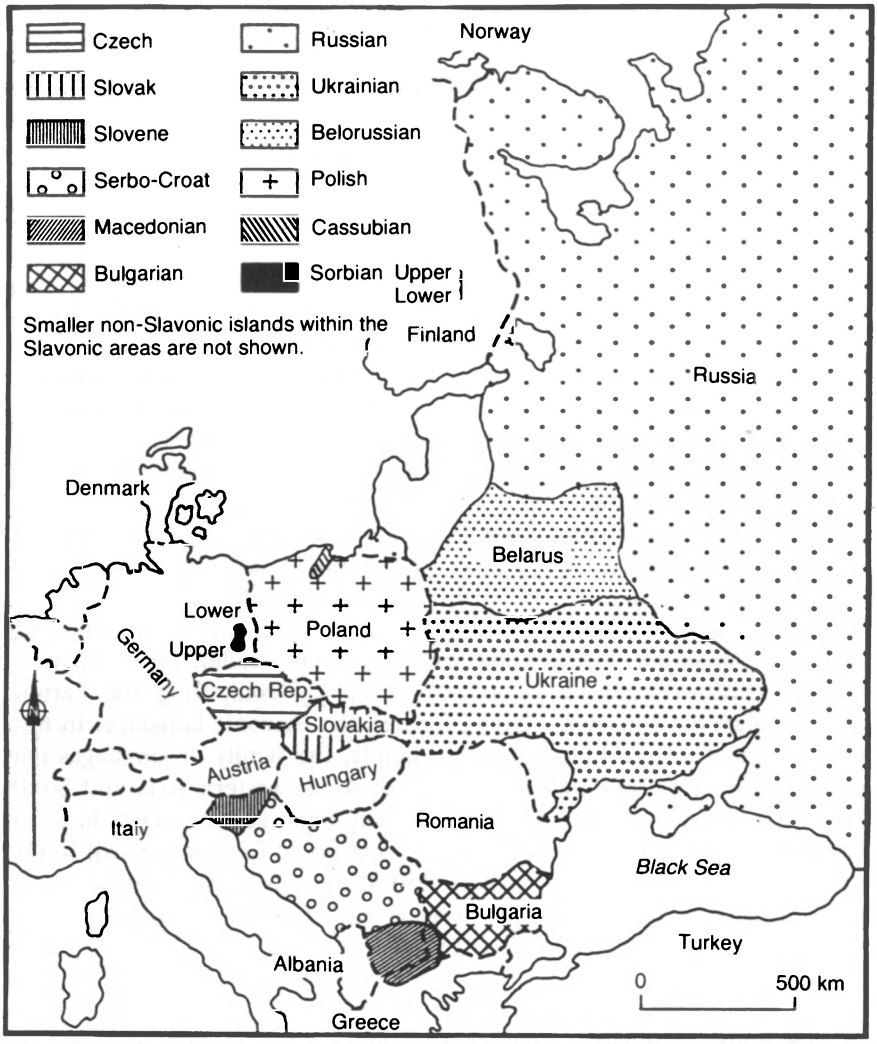
The Slavonic languages form a genetic unit, that is they are all descendants of a single ancestor language, conventionally called Proto-Slavonic, whose characteristics can be reconstructed by comparing the various attested Slavonic languages. Going further, the Slavonic languages in turn form a branch of the Indo-European family, the family of languages that covers most of Europe and large parts of south-western Asia and South Asia and which includes English: the ultimate genetic relatedness of English and Russian, while perhaps not apparent at first glance, can still be seen in such similar items as Russian три/tri, English *three*, Russian сын/syn, English *son*, Russian свинья́/svin'já 'pig', English *swine*.

1 The structure and scope of the book

In this book, a separate chapter is devoted to each of the following languages: Bulgarian, Macedonian, Serbo-Croat, Slovene, Czech, Slovak, Upper and Lower Sorbian (one chapter devoted to these two closely related languages), Polish, Russian, Belorussian, Ukrainian. For the present geographical location of these languages, reference should be made to map 1.1. These are the generally recognized contemporary standard literary Slavonic languages, each of which is either the (at least *de facto*) official language of an independent country or countries (Belorussian, Bulgarian, Czech, Macedonian, Polish, Russian, Serbo-Croat, Slovak,

Map 1.1 Approximate present-day distribution of Slavonic languages in Europe

Source: Adapted from Jakobson, 1955



More detail can be found on the maps in the individual chapters; this is particularly relevant for the area of the former Yugoslavia.

Slovene, Ukrainian) or is used locally for some official purposes (Upper Sorbian, Lower Sorbian) – this reflects the political situation in early 1993. Although the official status of Cassubian is as a dialect of Polish rather than as a separate language, the distinctiveness of this variety in comparison to the bulk of Polish dialects has led us, following the practice of many other

Slavists, to treat it as a separate language. In addition, two extinct languages have been included: Old Church Slavonic, as the language of the oldest attested writings in a Slavonic language, of great importance for our understanding of the early history of Slavonic languages, and Polabian, which is structurally clearly to be considered a distinct Slavonic language. A further chapter has been devoted to Proto-Slavonic, the reconstructed ancestor of the Slavonic languages, which provides the necessary bridge between the Indo-European family and its Slavonic branch. Finally, two chapters do not deal with individual languages: that on alphabets and transliteration (chapter 2) discusses material particularly important in the case of Slavonic because of the variety of alphabets, orthographic conventions and scientific and non-scientific transliterations of non-Roman alphabets that are current for Slavonic languages, while the chapter on the Slavonic languages in emigration (chapter 18) emphasizes that many Slavonic languages are spoken in lands far beyond the Slavonic heartland in eastern and central Europe and shows the importance of these far-flung communities for socio-linguistic research.

This volume differs from previous surveys of the Slavonic family in several significant respects. First, each chapter is written by an acknowledged specialist in the particular language. The expansion of work in the field means that it is now impossible for an individual to cover the whole family with the necessary level of expertise. Second, the chapters are highly structured, with each author providing detailed information on the same important topics. Thus the reader interested in a specific topic, whether it be verbal aspect, clitics or numerals, can easily find comparable information on each of the Slavonic languages. And the comparison is further facilitated by treating the languages as of equal linguistic interest; the relative standing of the different languages in terms of number of speakers and political importance does not influence the attention accorded to each. Third, within the descriptions syntax is given its due place. Modern linguistics has put syntax in the centre of the stage; this means that much more is known about the syntax of the Slavonic languages than was the case even a few years ago. Fourth, the book is accessible to a wide readership. To assist non-Slavists, all the examples from languages which use the Cyrillic alphabet are given in transliteration (as well as in the Cyrillic form; see Transliteration from Cyrillic and notes there; pages xii–xiii). Terms which are likely to be less widely known are explained. Thus, besides providing an up-to-date survey of current knowledge for Slavists, the volume is also a source of reference for all others with an interest in the Slavonic family; indications of further sources in English and other widely read languages are provided where possible in the bibliographies. Given the major input from Slavists into mainstream linguistics in the past (see section 2) and from linguistics into Slavonic studies, it is natural to seek to maintain this relationship.

As was mentioned above, each of the individual-language chapters is written according to a single general plan. The structure of each chapter is as follows. An introductory section 1 provides a brief account of the current status of the language and of its historical development in social terms (including the development of the literary standard). The section on phonology (section 2) deals with the sounds of the language in question and relations among them, in particular the inventory of phonemes (section 2.1), that is which sounds can be used in the language to distinguish words (in the way that the phonemic opposition between /p/ and /b/ in English enables one to distinguish between *pin* and *bin*); the subsections on morphophonemics (sections 2.2 and 2.3) deal with the ways in which the phonemic shape of an item can change in different morphological forms, as in the way that the English morpheme (minimal grammatical unit) *wife* appears in phonemically different shapes in the words *wife* /waif/ and *wives* /waivz/, that is, /waif/ versus /waiv/. The section on morphology (section 3) deals with the details of how morphemes are combined into words, such as how the English morphemes *pen* and *-s* (the plural suffix) combine to give *pens* or, to take a more complex example, how the morpheme *sing* combines with the morpheme for past tense to give the word *sang*. Morphology can be further divided into inflectional morphology (sections 3.1 and 3.2), which deals with relations among different forms of a single lexical item, such as the relations among *walk*, *walks* and *walked* as different forms of the lexical item WALK in English, and derivational morphology (section 3.3), which deals with the relations among distinct but formally related lexical items, such as among English *observe*, *observer*, *observation* and *observational*. Within inflectional morphology, Slavonic languages, like most Indo-European languages, make a clear distinction between nominal morphology (section 3.1) and verbal morphology (section 3.2); linguistic terminology is not entirely standardized in this area, so readers are asked to take particular care in noting the senses in which we use the following terms, especially *nominal* and *noun*: *nominal* is a cover term subsuming nouns (see below), adjectives, pronouns and numerals; *noun* refers to nouns in their narrow sense, the traditional 'name of a person, place or thing'; *adjective*, *pronoun* and *numeral* are used in their usual senses, as is *verb*. In general, the same paradigm items are given for each language to make comparison easier. Section 4 of each chapter deals with syntax, the various patterns of combining words into phrases and sentences. Section 5 of each chapter deals briefly with the lexis (vocabulary) of that language, including in particular the relative weight of lexical items inherited from Proto-Slavonic (or created using morphemes of Proto-Slavonic origin) and those borrowed from other languages; for comparative purposes, lexical items are given from three well-defined lexical fields that have been important in recent linguistic and anthropological studies of lexis, namely colour terms (follow-

ing Berlin and Kay 1969), body parts and kinship terms. Finally, section 6 discusses the most salient characteristics of the main dialects of the language.

Our emphasis on the innovative nature of the present volume should not be interpreted as a lack of gratitude towards the pioneering work of our predecessors, which has indeed made this book possible. The scientific study of the Slavonic language family has a history of well over a century, the initial work generally being considered Miklosich (1852–75), a detailed and compendious comparison of the individual Slavonic languages and of Slavonic with other Indo-European languages; the inclusion of a volume on syntax set an example that only too many successors have failed to heed. The next major landmark is Vondrák (1906–8); it is chastening to see how many of the problems that remain at the forefront of Slavonic linguistics are already treated in these early works, such as the positioning of clitic pronouns and the use of different cases after the copula. The intervening years have seen the appearance of the detailed comparative grammar of the Slavonic languages by Vaillant (1950–77), in addition to the first two volumes of the more concise work by Bräuer (1961–). Scholarly (as opposed to pedagogical) introductions to the Slavonic language family are available in various Slavonic languages, such as Бернштейн/Bernštejn (1961) in Russian, Horálek (1962) in Czech, Lehr-Spławiński, Kuraszkiewicz and Sławski (1954) in Polish and Nahtigal (1952) in Slovene (also available in a Russian translation). In Continental Western European languages there is van Wijk (1956) in French, in addition to the German translation of Nahtigal (1952) and most recently Panzer (1991). In English such works range from the concise introduction of Jakobson (1955) via the medium-sized Entwistle and Morison (1949) to the detailed survey of the individual Slavonic languages of De Bray (1951). Finally, important recent contributions to the social and cultural development of the Slavonic languages have appeared in the publications of the Yale Concilium on International and Area Studies: Schenker and Stankiewicz (1980) and Picchio and Goldblatt (1984), and in Stone and Worth (1985).

In many ways the Slavonic languages form a homogeneous group within Indo-European. They are therefore an ideal area for comparative and typological work. A very positive aspect of this research has been the concern to consider data from each of the Slavonic languages, rather than just from the most easily accessible. This concern to give the comprehensive picture was initiated by Rudolf Růžička: see, for example, his account of reflexives (1973).

The perceived strategic importance of Russian from the 1950s on meant that American work on machine translation concentrated on translation from Russian to English, as in the Georgetown GAT system and the well-known SYSTRAN work (Hutchins 1986: 70–8, 209–18). There has also

been a good deal of research in the former Soviet Union. While machine translation has had a chequered history, the linking of interest in Slavonic languages with computer technology has had several spin-offs. The Slavonic family is probably uniquely well provided with research tools in the form of morphological dictionaries (whether giving roots and derived forms or concentrating on inflectional information), frequency dictionaries and reverse dictionaries. Russian is particularly well covered, but researchers working on certain other Slavonic languages also have useful grammatical and lexicological reference works at their disposal.

2 Some salient characteristics of Slavonic languages

In this section, we list some of the most important typological characteristics of the Slavonic languages, in particular those that have provided important material for the development of general linguistic theory (sections 2.1–2.4).

In phonology, one of the most distinctive features of Slavonic languages is the presence of a substantial number of palatal and palatalized consonants, in many Slavonic languages forming pairs of palatalized (soft) and non-palatalized (hard) consonants; perhaps the extreme case is Russian, where almost every consonant participates in this palatalization opposition. Another characteristic of Slavonic languages is the presence of an extensive set of morphophonemic alternations within inflectional and, especially, derivational morphology, as in the *k*:*č* alternation in Russian крик/krik 'shout' versus кричáть/kričáť 'to shout'; see further section 2.1.

All Slavonic languages have a rich morphology, including a rich inflectional morphology, and in this respect can be characterized as conservative Indo-European languages. While some languages have lost some of the inflectional categories found in Proto-Slavonic (perhaps most strikingly the near-complete loss of case in Bulgarian and Macedonian), all Slavonic languages retain a rich set of morphological categories; often there are even a few innovations relative to Proto-Slavonic. Typologically, Slavonic morphology is primarily fusional, that is a given affix frequently combines the expression of a number of grammatical categories, for example in Russian столу́/stolú 'table' (DAT SG), the inflection *-u* encodes simultaneously dative case (compare NOM SG стол/stol) and singular number (compare DAT PL столáм/stolám). Morphologically, verbs and, especially, nouns fall into a number of distinct conjugational/declensional classes, so that while Russian стол/stol 'table' has its dative singular in *-u*, жена́/žená 'wife' has жене́/žené with the affix *-e*, and кость/kost′ 'bone' has ко́ст-и/kóst-i with the affix *-i*. Most of the morphological categories found in Slavonic languages are those familiar from other Indo-European languages, but one verbal category that is particularly richly developed in Slavonic languages is aspect (section 2.2).

Perhaps the most salient syntactic characteristic of Slavonic languages is their so-called free word order, whereby the order of major constituents is determined not so much by syntactic factors (grammatical relations, such as subject, object) as by pragmatic factors (such as topic, focus; see further section 2.3); this makes it hard to characterize individual Slavonic languages in terms of such typologies as Subject–Verb–Object versus Subject–Object–Verb. This freedom of word order is particularly clear in the case of the major constituents of the clause (such as subject, verb/predicate, direct object, indirect object), while the order within individual constituents tends to be more fixed (though by no means always absolutely so): thus genitives usually follow their head noun, while demonstratives, numerals and adjectives usually precede; all Slavonic languages make extensive use of prepositions, with postpositions having at best marginal status.

Slavonic languages have extensive agreement systems, for instance between adjectives and their noun or between verbs and their subject, and the intersection of agreement with the rich morphology already alluded to gives rise to a number of complications with theoretically interesting resolutions (see further section 2.4). The fact that finite verbs usually encode the person–number of their subject leads to the possibility of omitting unstressed subject pronouns, although the extent to which such omission is favoured differs from language to language: in Serbo-Croat, for instance, it is normal to omit unstressed subject pronouns, while in Russian their inclusion is usual.

Subordination in Slavonic languages in general follows patterns familiar in other European languages, with a strong preference for finite subordinate clauses with clause-initial conjunctions and, in most languages (the exceptions are Bulgarian, Macedonian and the eastern variant of Serbo-Croat), an infinitive used in certain constructions where its understood subject can be retrieved from the syntactic context. The written Slavonic languages also make extensive use of other non-finite constructions, such as participles substituting for relative clauses and gerunds (verbal adverbs) substituting for adverbial clauses, although such non-finite constructions are not characteristic of the spoken languages.

2.1 Morphophonemics (Morphophonology)

One characteristic of all Slavonic languages is a rich set of morphophonemic (morphophonological, morphonological) alternations. Indeed, it is perhaps not surprising that much of the fundamental work in morphophonemics, including generative phonology, has been done by linguists who worked largely with Slavonic material: Jan Baudouin de Courtenay, Nikolaj Trubeckoj (Trubetzkoy), Roman Jakobson and Morris Halle; see, for instance, Anderson (1985: 56–139, 318–22), Jakobson (1948) and Halle (1959).

By a morphophonemic alternation we understand a situation where a given morpheme (minimal grammatical unit) has more than one phonemic representation in different words into which that morpheme enters. In English, for instance, the alternation between /ei/ and /ou/ in *break* /breik/ versus *broken* /broukən/ is an instance of morphophonemic alternation involving the two allomorphs (/breik/, /brouk/) of the morpheme *break.*

A few of the morphophonemic alternations found in Slavonic languages continue alternations found in Proto-Indo-European: for instance, the vowel alternation found in Russian теку́/tekú 'I flow' versus ток/tok 'current' is a direct reflex of the Indo-European ablaut alternations that also show up in, for example, English *break* versus *broken.* For the most part, however, the morphophonemic alternations of Slavonic languages represent either Proto-Slavonic innovations or the innovations of individual Slavonic languages, since a propensity for generating new morphophonemic alternations seems to be a characteristic of Slavonic languages.

In the Proto-Slavonic period, for instance, major new morphophonemic alternations arose as the result of the various palatalizations (see further sections 2.9–2.10 of chapter 3). Thus, the alternation that shows up in Russian пеку́/pekú 'I bake' versus печёшь/pečëš′ 'you bake' derives from the first palatalization of *k* to *č* before a front vowel (in this case, Proto-Slavonic *e*). Another set of morphophonemic alternations that arose in the Proto-Slavonic period was that between back and front vowels, depending on whether the preceding consonant was hard (non-palatalized) or soft (palatalized), as the result of a sound change whereby vowels were fronted after soft consonants. Thus the ending of the nominative–vocative–accusative singular of *o*-stem neuter nouns remained *-o* after hard consonants, but became *-e* after soft consonants, as can still be seen in Russian ме́сто/mésto 'place' versus по́ле/póle 'field'.

In the late Proto-Slavonic period, new morphophonemic alternations between a vowel and zero arose; the sign for zero is ∅. These alternations came about through the loss of the reduced vowels (symbolized ъ and ь deriving from Proto-Indo-European *u* and *i*, respectively), which are known as *jers. Jers* in strong positions developed into full vowels (the actual vowels are different in different Slavonic languages) while those in weak positions were lost, thus producing alternations like Russian рот/rot 'mouth', GEN SG рта/rta (that is, *rot-* alternates with *r∅t-*), for Old Russian ръть/гьть, GEN SG рьта/гьта. (See further section 2.25 of chapter 3.) Vowels that alternate with zero in this way are known as 'mobile', 'fugitive' or 'fleeting' vowels.

The phenomenon of akan′e in Russian (see chapter 15, sections 2.2 and 6), whereby unstressed *a* and *o* became ʌ or ə in unstressed syllables, provides an example of a language-specific sound change that has given rise to morphophonemic alternations, as can be seen from comparing the

vowels in the different inflectional forms of Russian голова́/golová 'head': NOM SG голова́/golová [gəlʌvá], NOM PL го́ловы/gólovy [góləvɨ], GEN PL голо́в/golóv [gʌlóf], where the shifting stress gives rise to alternations between *o* on the one hand (under stress) and *ʌ* or *ə* on the other (no stress). Note that we mark stress by ˈ on the stressed vowel; this symbol is chosen to avoid confusion with other diacritics.

The existence of morphophonemic alternations led linguists investigating Slavonic languages to posit a level of morphophonemic representation at which a given morpheme would be given a constant representation; one convention for indicating that a representation is morphophonemic is to enclose it in braces, that is { }. Thus, the stem of the Russian word for 'head' would be {golov-}, which would then be related, by the operation of rules, to more phonetic representations such as those given above for individual inflectional forms. Because of different morphophonemic behaviour, segments that are phonetically and phonemically identical may receive different morphophonemic representations. Thus the vowels of the first syllable of Russian сова́/sová [sʌvá] 'owl' and страна́/straná [strʌná] 'country' are phonetically identical, but are differentiated in other inflectional forms of the words when the first syllable is stressed, as in nominative plural со́вы/sóvy [sóvɨ], стра́ны/stránу [stránɨ], that is, the morphophonemic representations of these stems would be {sov-} and {stran-}, respectively. In describing the morphology of Old Church Slavonic, it is necessary to distinguish morphophonemically between two kinds of *y*, since {y_1} shows up after soft consonants as *i*, while {y_2} shows up as *ę* after soft consonants; compare the following forms of the masculine *o*-stem nouns *rabъ* 'slave' and *mǫžь* 'man': INST PL *raby* (morphophonemically {raby$_1$}), *mǫži*, ACC PL *raby* (morphophonemically {raby$_2$}), *mǫžę*. In this, one can see the origin of abstract levels of phonological representation in generative phonology. The precise degree of abstractness that should be allowed in morphophonemic alternations has proved to be controversial; while probably most linguists would be happy with the morphophonemic representations proposed in the preceding paragraph, many would be less happy with Jakobson's attempt to account for the alternation found in Russian examples like жать/žatʹ 'to press', first person singular present tense жму/žmu, by positing a morphophonemic representation {žm-} and a rule that drops the nasal and inserts *a* in the infinitive stem; instead, this latter example would probably be treated most simply as a morphological irregularity.

Although morphophonemic alternations typically arise as the result of conditioned sound changes, there is a tendency, well reflected in Slavonic languages, for the original phonological conditioning to be lost, that is for morphophonemic alternations to become increasingly morphologized. In Proto-Slavonic, the difference in endings selected by hard- and soft-stem nouns was transparently phonological, so that in Old Russian, for instance,

we find the ending for hard stems -ъ in genitive plural рабъ/rabъ 'slave' and the corresponding ending for soft stems -ь in дъждь/dъždь 'rain'. In Modern Russian, however, the relevant forms are рабо́в/rabóv and дожде́й/doždéj, where the suffixes *-ov* and *-ej* are morphological, rather than morphophonemic, alternants; the expected correspondent of *-ov* for soft stems, namely *-ev*, occurs only with nouns ending in *j*, for example геро́ев/geróev from геро́й/gerój 'hero'. Finally, in every Slavonic language at least some instances of some inherited morphophonemic alternations have been lost by analogy. In Old Russian, for instance, the dative–locative singular of рука/ruka 'hand' was руцѣ/rucě, with *c* as a result of the second palatalization; in Modern Russian, however, we have simply NOM SG рука́/ruká, DAT-LOC SG руке́/ruké. In Proto-Slavonic, and still in Old Church Slavonic, the morphophonemic opposition of hard versus soft consonants corresponds exactly to the phonetic opposition of non-palatalized versus palatalized consonants. During the history of several individual Slavonic languages, however, some of these consonants have become phonetically non-palatalized but none the less retain their earlier morphophonemic behaviour. In such languages, the morphophonemic class of soft consonants thus no longer corresponds exactly to the phonetic class of palatalized consonants, as when Russian masculine nouns ending in (synchronically non-palatalized) *ž* require the genitive plural ending *-ej* characteristic of soft stems, rather than *-ov* as with hard stems, as in ежей/ežéj, genitive plural of ёж/ёž 'hedgehog'.

2.2 Aspect

One of the major contributions of Slavonic linguistics to general linguistic theory has been the notion of verbal aspect; indeed the very term 'aspect' in this sense is a direct translation of Russian вид/vid (compare ви́деть/vídet′ 'to see'). In every Slavonic language, with Russian used here as an illustration, most verbs occur as a pair, one member of the pair being of the imperfective aspect (such as писа́ть/pisát′ 'to write'), the other being of the perfective aspect (such as написа́ть/napisát′). Like tense, aspect is concerned with the general notion of time, but whereas tense is concerned with locating the situation described by the clause in time, relative to other time points (most commonly the present moment), aspect is concerned with the internal temporal structure of situations. The essential content of the perfective/imperfective opposition is that between bounded and unbounded situations, or rather between the presentation of situations as bounded or unbounded, respectively. In a sentence like Ко́ля написа́л (PRFV) письмо́/Kólja napisál (PRFV) pis′mó 'Kolja wrote a letter', Kolja's writing of the letter is presented as a bounded event, that is as an event that is complete – from which we can deduce that Kolja did indeed finish writing the letter. By contrast, Ко́ля писа́л (IMPFV) письмо́/Kólja pisál (IMPFV) pis′mó 'Kolja was writing the letter', makes no explicit reference to

the boundedness or completion of the writing; indeed, it is quite possible that Kolja gave up writing the letter before completing it. As this example illustrates, the distinction sometimes corresponds to that between simple and progressive verb forms in English, namely with dynamic verbs referring to a single action, but this is not a general equation. Stative verbs are typically in the simple form in English, for example *the book lay on the table*, whereas in Russian they are typically in the imperfective, since a state is by definition unbounded (to begin or to end a state is an action, not part of the state), whence Russian кни́га лежа́ла (IMPFV) на столе́/kni̇ga ležála (IMPFV) na stolé. In English, habitual situations are usually in the simple form, for example *Kolja wrote a letter every day*, whereas Russian uses the imperfective, since the habit (as opposed to any individual act of letter writing) is not bounded, that is Ко́ля писа́л (IMPFV) письмо́ ка́ждый день/Kólja pisál (IMPFV) pis'mó káždyj den'.

Aspect is particularly salient in the Slavonic languages because the perfective/imperfective opposition characterizes virtually all verb forms, usually covering all moods and tenses (though the present/future opposition is typically neutralized in the perfective) and both finite and non-finite forms. The pervasiveness of aspectual oppositions in Slavonic languages is, no doubt, one reason for the extent to which aspectology has leant on the Slavonic opposition; see, for instance, Comrie (1976).

Although the perfective/imperfective opposition is the basic opposition in all Slavonic languages, most Slavonic languages also have some other, typically more restricted, aspectual oppositions (called sub-aspects). Most Slavonic languages, for instance, have an opposition, restricted to verbs of motion, between a determinate sub-aspect (essentially, motion in a single direction) and an indeterminate sub-aspect (motion in various directions). Russian contrasts determinate идти́/idtí with indeterminate ходи́ть/xodít' 'to go', for example он идёт в шко́лу/on idët v škólu 'he is going to school' versus он хо́дит по по́лю/on xódit po pólju 'he is walking about the field'; both are imperfective – the perfective is пойти́/pojtí 'to go, set out', as in он пошёл в шко́лу/on pošël v škólu 'he has gone to (set out for) school'.

It is worth noting briefly the basic principles of the formation of imperfective–perfective verb pairs in Slavonic languages, using Russian examples. In general, simple unprefixed verbs are imperfective (such as писа́ть/pisát' 'to write'), with only a handful of exceptions being perfective (like дать/dat' 'to give'). Perfective verbs are formed from simple unprefixed imperfective verbs primarily by prefixation. Prefixation also normally changes the lexical meaning, so that, for instance, the perfective verb описа́ть/opisát' means 'to describe', though for a given simple unprefixed imperfective verb there is typically one (lexically determined) prefix that is most neutral; in the case of писа́ть/pisát' this is на-/na-, so that написа́ть/napisát' can be glossed simply as 'to write' (PRFV) –

whether such prefixes are ever truly lexically neutral is one of the controversies in current Slavonic aspectology. Less commonly, perfectives are formed by suffixation, as in Russian исчéзнуть/iščéznut′ 'to disappear', the perfective of исчезáть/iščezát′. Prefixed perfectives (in particular, those where the prefix carries a difference in lexical meaning), and also simple unprefixed perfectives, form corresponding imperfectives by suffixation: the imperfective of описáть/opisát′ 'to describe' is опúсывать/opìsyvat′, that of дать/dat′ 'to give' is давáть/davát′. In addition, most languages have some idiosyncratic pairs, including suppletive pairs, for example, Russian брать/brat′ (IMPFV) versus взять/vzjat′ (PRFV) 'to take'.

2.3 Functional Sentence Perspective

In English, word order plays an important role in carrying the basis syntactic relations within a sentence. In *John saw Mary*, for instance, only this particular order of words is possible, and any change in the order of words either changes the meaning (as in *Mary saw John*) or leads to a nonsentence (for example, *John Mary saw, saw Mary John*). In Slavonic languages, however, the word order is not tied to the expression of syntactic relations in this way. In Russian, for instance, any of the six logically possible word-order permutations of the sentence Кóля (NOM) вúдел Тáню (ACC)/Kólja (NOM) vìdel Tánju (ACC) 'Kolja saw Tanja' is grammatical and has the same basic meaning, that is refers to a situation in which Kolja saw Tanja; thus Тáню (ACC) вúдел Кóля (NOM)/Tánju (ACC) vìdel Kólja (NOM) still means 'Kolja saw Tanja', and not, for instance, 'Tanja saw Kolja'. This freedom of word order goes hand in hand with the richer morphology of Slavonic languages. In the Russian example just cited, the prime indication of who did the seeing is the nominative case of the noun phrase Кóля/Kólja, while the prime indication of who was seen is the accusative case of the noun phrase Тáню/Tánju, thus freeing word order to express other distinctions.

Just what does order express in Slavonic languages? There is no uniformity of terminology, and even some disparity of conceptual basis, in answering this question, but the following represents something approaching a consensus. Word order in Slavonic languages is determined primarily by the arrangement of given and new information, more specifically placing towards the beginning of the sentence information that is given (that is, already shared by speaker and hearer) and placing towards the end of the sentence information that is new (that is, the new information that the speaker wants to convey to the hearer). If one interlocutor asks 'Who did Kolja see?', then the fact of Kolja's seeing is given information and will come first in the answer in a Slavonic language, while the fact that the one seen is Tanja is new information, so that this will come last, giving, for instance, the version Кóля (NOM) вúдел Тáню (ACC)/Kólja (NOM) vìdel

Tánju (ACC) in Russian. Conversely, if the question is 'Who saw Tanja?', then the fact of Tanja's being seen is given information, while the identification of the one who did the seeing as Kolja is new information, giving, for instance, the Russian version Тáню (ACC) вúдел Кóля (NOM)/Tánju (ACC) vídel Kólja (NOM) as an appropriate answer. The new information is also referred to as the focus of the sentence.

Sometimes, the structure of the discourse will force or suggest a particular constituent of the sentence as what that sentence is about. For instance, if someone asks 'What about Tanja?', then an appropriate reply must be about Tanja. The item that the sentence is about is called its topic (or theme), the rest of the sentence is the comment (or rheme). In Slavonic languages, the topic usually occurs at the beginning of the sentence. Imagine the following conversation between A and B:

A: Vanja saw Vera.
B: What about Kolja? Who(m) did he see?
A: Kolja saw Tanja.

In the last turn of the conversation, Kolja has been established as topic, and the rest of the sentence is the comment. Furthermore, the fact that the person seen is Tanja is the new information or focus, so we have: topic *Kolja*, comment *saw Tanja*, focus *Tanja*, giving in Russian the word order Кóля (NOM) вúдел Тáню (ACC)/Kólja (NOM) vídel Tánju (ACC). Thus, in a sense the basic word order in most Slavonic languages can be said to be Topic–X–Focus, where X represents material other than the topic and focus (non-focus comment material); deviations from this order serve primarily to indicate emotional expressiveness in spoken registers.

These major differences between the function of word order in English and in Slavonic languages were first studied in detail by linguists of the Prague School, such as Vilém Mathesius (Mathesius 1939 and, more generally, 1947), who were interested in comparing and contrasting English and Czech syntax. The general area of study that covers such notions as topic, comment, focus, is referred to variously as functional sentence perspective, communicative dynamism, topic–comment (theme–rheme) structure; the Prague School used the Czech term *aktuální členění*. They noticed another distinction that ties in closely with those already mentioned (freedom of word order, richness of morphology). English has a number of productive syntactic processes that enable one to change grammatical relations (such as which noun phrase is subject of a sentence); the most evident is the passive, which enables one to rephrase *Kolja saw Tanja* as *Tanja was seen by Kolja*. One function of such syntactic processes in English is to bring a noun phrase to sentence-initial position, thus marking it overtly as topic of the sentence. While passives are possible in Slavonic languages, they tend not to be particularly idiomatic, especially in spoken

registers, and are usually much more heavily restricted than in English, with, for instance, no possibility of a literal translation of *Kolja was given a book by Tanja.* In functional terms, the equivalent of English *Tanja was seen by Kolja* in a Slavonic language is not a passive, but rather an active sentence with the object preposed, such as Russian Та́ню (ACC) ви́дел Ко́ля (NOM)/Та́nju (ACC) vídel Ko̍lja (NOM). To a large extent, the functional equivalent of English rules that change grammatical relations is the possibility of word-order permutations in Slavonic languages.

While it is reasonably clear that for English the basic word order is Subject–Verb–Object (SVO), the question of the basic word order in Slavonic languages is not so clear in syntactic terms (though one can say that the basic order is Topic–X–Focus). For some languages, such as Russian, there is consensus that the basic order is SVO, for instance on the basis of the greater textual frequency of SVO over other word orders and on the basis of the preferred interpretation of potentially ambiguous sentences like мать лю́бит дочь/mat′ ljúbit doč′ 'the mother loves the daughter' (rather than 'the daughter loves the mother'), where both nouns happen not to distinguish nominative from accusative. The grammatical traditions of some other Slavonic languages, however, either suggest other basic orders or no basic order in syntactic terms, and this is reflected in the chapters on individual Slavonic languages.

2.4 Agreement and agreement categories

Slavonic languages preserve a rich inflectional morphology and have made innovations in the categories involved in agreement; these two facts result in complex agreement systems, which have attracted considerable interest. Typically, we find agreement within the noun phrase in case, number and gender. In Russian интере́сн-ая кни́га/interésn-aja kníga 'interesting book', the adjective интере́сн-ая/interésn-aja stands in the nominative singular feminine form, these features matching those of the head noun. Finite verbs typically agree with their subject in person and number; Russian мы пи́ш-ем/my píš-em 'we write' as opposed to, say, они́ пи́ш-ут/oní píš-ut 'they write'. Past tenses are frequently formed with the so-called *l*-participle, which creates a more interesting situation, as in Serbo-Croat Snèžana je dòšla/Снѐжана је дòшла 'Snežana came (literally: Snežana is come)'. Here the auxiliary verb je/је 'is' shows agreement in person and number (third person singular), while the participle shows agreement in number and gender (singular and feminine). Some Slavonic languages, such as Russian, use a null form for the verb 'be' in the present tense, so that we find: Та́ня пришла́/Tánja prišlá 'Tanja came'. The former participle is the sole form in the past tense, so it may be said that Russian verbs agree in person and number in the present, but in number and gender in the past. Various types of pronoun also show agreement with their antecedents, in number and gender. The description given so far

covers a large proportion of the instances of agreement in Slavonic. There are, however, many examples where additional factors are involved, which require elaboration of our account. We will consider these first in terms of the constructions where complications occur, and then by looking at the agreement categories affected.

There are several constructions where more than one agreement form may be found. Consider this example of agreement with conjoined noun phrases in Russian: преподава́лась матема́тика и фи́зика/ prepodaválas' matemátika i fizika 'was taught mathematics and physics', that is, 'mathematics and physics were taught'. We find agreement (feminine singular) with just the nearer conjunct матема́тика/ matemátika. But the plural form преподава́лись/prepodavális' is also possible, showing agreement with both conjuncts. A similar option occurs in comitative constructions (such as Ива́н с бра́том/Iván s brátom 'Ivan with brother', that is 'Ivan and his brother'). Quantified expressions too are a complex area: given a phrase like пять де́вушек/pjat' dévušek 'five girls' we may find plural agreement, but alternatively also neuter singular agreement. Subject–verb agreement in Slavonic is normally controlled by a noun phrase in the nominative case. Here, however, the part of the phrase which is in the nominative case, пять/pjat' 'five', lacks gender and number features, while де́вушек/dévušek 'girls' is in the genitive plural; one possibility, therefore, is that agreement fails, and so the verb takes the default form, the neuter singular. Alternatively, the plural may be used, given that the quantified expression, though not formally nominative plural, nevertheless denotes a plurality. This gives rise to variants пришло́ (SG) пять де́вушек/prišló (SG) pjat' dévušek and пришли́ (PL) пять де́вушек/prišlí (PL) pjat' dévušek 'five girls came'. The constructions discussed have in common a choice between agreement determined by the form, 'syntactic agreement', or by the meaning, 'semantic agreement'. Several factors bear on the choice in individual instances. Let us look at the problem from the viewpoint of the item which determines the agreement, the 'agreement controller' (for instance, the conjoined noun phrases in our first example). We find that controllers which precede the agreeing element and controllers which denote animates are more likely to give rise to semantic agreement (plural in this case) than those which do not. If we start from the agreeing element or 'target' we find that predicates are more likely to show semantic agreement than are attributive modifiers, relative pronouns more so than predicates, and personal pronouns more so than relative pronouns. This is the Agreement Hierarchy (attributive < predicate < relative pronoun < personal pronoun). Further patterns have been established, in addition to the interaction of these two major types of factor, so that the picture is indeed complex; see, for instance, Corbett (1983) for further details.

Let us now move on to consider the agreement system in terms of the

categories involved. Person is perhaps the least controversial since Slavonic has the three persons found widely within and beyond Indo-European. It is worth noting, however, that Polish uses third-person forms for polite address. Number is more complex; Slavonic inherited a three-number system, singular/dual/plural, but the dual has been lost in almost all the modern Slavonic languages. However, this loss has caused considerable complications in agreement within numeral phrases involving the number 'two', and often 'three' and 'four' as well.

Slavonic languages are particularly helpful for coming to a clearer understanding of the category of gender (for a general survey see Corbett 1991). The standard languages preserve three genders, masculine, feminine and neuter, though the neuter is under pressure in several languages and is being lost in some dialects. In addition to the three main genders, a new subgender of animacy has arisen. In the accusative case, animates take different agreements from inanimates, for example Serbo-Croat òvāj prózor/òvāj прóзор 'this window' (inanimate) as opposed to òvog sîna/òvог сйна 'this son' (animate). There are no separate accusative forms involved, but always syncretism with the genitive; thus òvog sîna/òvог сйна 'this son' is also a genitive case form. There is considerable variation of two sorts, first in the forms affected. In the south-west, as in the case of Serbo-Croat, only the masculine gender is subdivided into animate and inanimate, and that in the singular only. Russian, in the north-east, is at the other end of the spectrum since animacy affects the masculine singular and all genders in the plural. The other type of variation is in the categories of nouns treated as animate. First male humans of certain types were included, and then the boundary spread 'downwards'. The animate subgender is still much more firmly based on semantic classifications than are the three older genders, but in some languages various inanimates may be treated as animate and the semantic basis is becoming less clear.

In addition to this elaboration of the gender system, some West Slavonic languages have further introduced special agreements for nouns denoting male persons (thus a subset of the masculine animate subgender). In Polish, for example, we have in the plural an opposition between predicate agreement forms such as *byli* 'were', for subjects denoting male persons, and *były* 'were' for all other plural subjects. Here again, the assignment of nouns to the masculine personal category is much more clearly based on semantics than are the traditional genders.

Finally, we look at case (though recognizing that the matching of case forms need not necessarily be treated as agreement). Most of the Slavonic languages preserve a vital case system, with minor weakenings (several have lost the vocative, for example). However, Bulgarian and Macedonian have dramatically reduced the inherited case system. For the languages which retain a substantial case system, a particularly interesting problem of case agreement is the question of whether nominal complements of copular

verbs agree in case with their subject or not: in the equivalent of sentences like 'he was a fine poet', the complement may be nominative or instrumental, depending on the language, with both possibilities acceptable in some languages.

3 Suggestions for using this book

The book has been designed to meet the differing requirements of a variety of readers. Some need a straightforward reference work, and for them information on particular languages can be found through the contents page, while more specific data on particular topics is to be located through the index. Then there are linguists of various types, who may require a general introduction to the Slavonic family. Such readers might start by working on any one of the chapters devoted to a contemporary Slavonic language, and then by branching out from there. Typologists can begin from a particular problem, whether in phonology, morphology, syntax or lexis, and move from language to language concentrating on the relevant section. Historical linguists with a grounding in Indo-European philology will no doubt prefer to start with the chapters on Proto-Slavonic, Old Church Slavonic and the alphabets and transliteration, and then progress to the modern languages. Sociolinguists should begin at the end, with the chapter on the Slavonic languages in exile, and then refer particularly to the introductory and dialect sections of the chapters on the modern languages.

Slavists too can approach the book in various ways. The Slavist could look first at the Slavonic language he or she knows best, since this will make it clear how the familiar information is organized. Alternatively, it makes sense to explore the family either by looking at a new Slavonic language which is closely related to a familiar one, or to take the opposite view and to look at a language which is as different as possible from the language or languages already known. After looking at one or more of the contemporary languages, the Slavist might then take in the chapter on alphabets and transliteration, followed by the historical perspective in the chapters on Proto-Slavonic and Old Church Slavonic, and also the chapter on the Slavonic languages in exile.

References

Anderson, Stephen R. (1985) *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*, Chicago and London: University of Chicago Press.

Berlin, Brent and Kay, Paul (1969) *Basic Color Terms: their Universality and Evolution*, Berkeley: University of California Press.

Bräuer, Herbert (1961–) *Slavische Sprachwissenschaft*, Berlin: Walter de Gruyter.

Comrie, Bernard (1976) *Aspect*, Cambridge: Cambridge University Press.

Corbett, Greville G. (1983) *Hierarchies, Targets and Controllers: Agreement Patterns in Slavonic*, London: Croom Helm.

—— (1991) *Gender*, Cambridge: Cambridge University Press.

De Bray, R.G.A. (1951) *Guide to the Slavonic Languages*, London: Dent. (Third edition 1980 in three volumes: *Guide to the South Slavonic Languages*; *Guide to the West Slavonic Languages*; *Guide to the East Slavonic Languages*, Columbus, Oh.: Slavica.)

Entwistle, W.J. and Morison, W.A. (1949) *Russian and the Slavonic Languages*, London: Faber & Faber.

Halle, Morris (1959) *The Sound Pattern of Russian*, The Hague: Mouton.

Horálek, Karel (1962) *Úvod do studia slovanských jazyků*, 2nd edn, Prague: Nakladatelství Československé akademie věd.

Hutchins, W.J. (1986) *Machine Translation: Past, Present, Future*, Chichester: Ellis Horwood, and New York: John Wiley.

Jakobson, Roman (1948) 'Russian conjugation', *Word* 4: 155–67. Reprinted 1984 in Roman Jakobson, *Russian and Slavic Grammar: Studies 1931–1981*, Berlin: Mouton, 15–26.

—— (1955) *Slavic Languages: a Condensed Survey*, New York: King's Crown Press.

Lehr-Spławiński, Tadeusz, Kuraszkiewicz, Władysław and Sławski, Franciszek (1954) *Przegląd i charakterystyka języków słowiańskich*, Warsaw: Państwowe wydawnictwo naukowe.

Mathesius, Vilém (1939) 'O tak zvaném aktuálním členění věty', *Slovo a slovesnost* 5: 171–4.

—— (1947) *Čeština a obecný jazykozpyt: soubor statí*, Prague: Melantrich.

Miklosich, Franz (1862–75) *Vergleichende Grammatik der slavischen Sprachen*, vol. I *Lautlehre*, vol. II *Stammbildungslehre*, vol. III *Formenlehre*, vol. IV *Syntax*, Vienna: Wilhelm Braumüller.

Nahtigal, Rajko (1952) *Slovanski jeziki*, Ljubljana: Univerza v Ljubljani, Filozofska fakulteta. (German translation 1961: *Die slavischen Sprachen, Abriß der vergleichenden Grammatik*, Wiesbaden: Harrassowitz; Russian translation 1963: *Славянские языки*, Москва: Издательство иностранной литературы.)

Panzer, Baldur (1991) *Die slavischen Sprachen in Gegenwart und Geschichte: Sprachstrukturen und Verwandtschaft* (Heidelberger Publikationen zur Slavistik. A: Linguistische Reihe, 3). Frankfurt am Main: Peter Lang.

Picchio, Riccardo and Goldblatt, Harvey (eds) (1984) *Aspects of the Slavic Language Question*, vol. I *Church Slavonic – South Slavic – West Slavic*, vol. II *East Slavic*, New Haven, Conn.: Yale Concilium on International and Area Studies.

Růžička, Rudolf (1973) 'Reflexive versus nonreflexive pronominalization in Modern Russian and other Slavic languages', in F. Kiefer and N. Ruwet (eds) *Generative Grammar in Europe* (Foundations of Language Supplementary Series 13), Dordrecht-Holland: Reidel, 445–81.

Schenker, Alexander M. and Stankiewicz, Edward (eds) (1980) *The Slavic Literary Languages: Formation and Development*, New Haven, Conn.: Yale Concilium on International and Area Studies.

Stone, Gerald and Worth, Dean (eds) (1985) *The Formation of the Slavonic Literary Languages* (UCLA Slavic Studies 11), Columbus, Oh.: Slavica.

Vaillant, André (1950–77) *Grammaire comparée des langues slaves*, 5 vols, Paris: IAC (vols I–II), Klincksieck (vols III–V).

van Wijk, Nicolaas (1956) *Les Langues slaves. De l'unité à la pluralité*, 2nd edn, The Hague: Mouton. (First edition 1937 in *Le Monde slave*.)

Vondrák, Václav (1906–8) *Vergleichende slavische Grammatik*, vol. I *Lautlehre und Stammbildungslehre*, vol. II *Formenlehre und Syntax*, Göttingen: Vandenhoeck & Ruprecht. (Second edition 1924–28.)

Бернштейн, С.Б. (1961) *Очерк сравнительной грамматики славянских языков*, Москва: АН СССР.

2 Alphabets and Transliteration

Paul Cubberley

ALPHABETS

Many alphabets have been used at one time or another to represent the Slavonic languages. The most commonly used, to be looked at in detail, are Glagolitic, Cyrillic and Latin (which we will hereafter call by the Slavonic name 'Latinica', for lack of a useful parallel English term like 'Latinic'); sporadically also the Greek, Arabic and even Hebrew alphabets have been used, and we will also deal briefly with these.

The distribution by location and period is roughly as follows:

Glagolitic: Moravia ninth century; Macedonia ninth to eleventh centuries; Bulgaria ninth to twelfth centuries; Croatia tenth to sixteenth centuries, then in Church usage until the nineteenth century, and sporadically into the twentieth century; Slovenia fifteenth to sixteenth centuries; Bohemia and Poland fourteenth to sixteenth centuries;

Cyrillic: Bulgaria ninth century to present; all the East Slavonic area (Russia, Ukraine, Belorussia/Belarus), Macedonia and Serbia (also Bosnia, Montenegro) tenth century to present;

Latinica: the West Slavonic area in general tenth century to present; Croatia and Slovenia tenth century to present; Serbia (but always secondary to Cyrillic), also Bosnia, the same period; Belorussia (and part of Ukraine) sixteenth to twentieth centuries;

Greek: Macedonia (especially Aegean) fifteenth to nineteenth centuries;

Arabic: Belorussia sixteenth to eighteenth centuries; Bosnia fifteenth to twentieth centuries;

Hebrew: Belorussia (fragmentary) sixteenth to eighteenth centuries.

Generally, there has always been a close correlation between alphabet and religion, though not necessarily one of cause and effect. The main correlations are shown in Table 2.1.

In the following we shall consider the early period as a whole, to indicate the establishment of the three main alphabets, and then follow the development of each separately. For the later periods only major reforms will be dealt with, and the details of reforms in each language should be sought in the relevant chapter.

Table 2.1: Slavs: alphabet and religion

Poland: Cath. – Lat.	Belarus: Orth. – Cyr. (Lat.) (Cath. – Lat.) (Musl. – Arab.) (Jud. – Hebr.)	Russia: Orth. – Cyr.
Lusatia (Sorbs): Cath. (Prot.) – Lat.		
Bohemia/Moravia: Cath. – Lat.	Slovakia: Cath. – Lat.	Ukraine: Orth. – Cyr. (Cath. – Lat.)
Slovenia: Cath. (Prot.) – Lat.	Croatia: Cath. – Lat. (Glag.)	Serbia: Orth. – Cyr. (Glag./Lat.)
Bosnia: Musl. – Lat. (Arab.) Orth. – Cyr. Cath. – Lat.	Macedonia: Orth. – Cyr. (Glag./Greek.)	Bulgaria: Orth. – Cyr. (Glag.)

Note: Parentheses indicate former or coexisting situations; layout is roughly geographical. Abbreviations used: Cath. Catholic, Orth. Orthodox, Prot. Protestant, Musl. Muslim, Jud. Judaic, Lat. Latinica, Cyr. Cyrillic, Glag. Glagolitic, Arab. Arabic, Hebr. Hebrew.

1 Early history

Traditionally, the start of Slavonic writing is credited to Constantine (also known as Cyril, the name he took on becoming a monk), who with his brother Methodius led a mission from Byzantium to the Moravian Slavs in the early 860s, in preparation reputedly having created an alphabet in which to write Slavonic speech.

1.1 Pre-Constantine period

The question of the extent to which any Slavonic language was written before the time of Constantine and Methodius remains unanswered, but of course not without hypotheses. The fact is that there are few facts! There is no hard evidence of any such written form for that period. In a general way, one can hypothesize that there must have been some cases of a written Slavonic language if only because the Slavs were active on many fronts well before 860: in the west, they lived next to, and usually, but not always, under the domination of, various Germanic peoples; in the south they had been fighting and living next to Greeks since the sixth century; in the east, they had been trading with the Greeks and the Scandinavians since at least the early ninth century. All of these peoples already had a writing system, whether Latinica or Greek, and it would be logically surprising if no attempt was ever made to use these alphabets to write some Slavonic

language, if only for trade or treaty purposes. Would a victorious side be content to see a treaty supposedly favouring them written only in the language of the conquered side? At best only if many of the victors' leaders were fluent in the language of the losers.

The fact remains, however, that there are no concrete examples of such writing, but only elusive pieces of a suggestive sort, the most famous being that occurring in the work of the monk Chrabr in his *O pismenech* (On the Letters) – believed to have been written in the 890s, or possibly even earlier and by Constantine himself – in which he enigmatically describes the pre-Constantine Slavs as having used *čerty i rězy* 'lines and cuts' with which to count and predict (*čьtaaxǫ i gadaaxǫ*). Speculation on the meaning of this has, not surprisingly, been wide, the most popular view being that the lines and cuts are no more than just that – counting signs, and not an alphabet; another, once common, view was that they might have been a runic alphabet, borrowed presumably from the Scandinavians.

A similar enigma surrounds the mention in the *Pannonian Life of Constantine* that during a mission to the Crimea in 860 he was shown a Gospel and Psalter written in *rousskymi pismeny*, on the face of it meaning 'in Russian letters' (or rather Rus′-ian, since the adjective would at that time have referred to the ethnonym 'Rus′', thus to all the East Slavs and not just the north-eastern branch later referred to as 'Russians'); were these indeed some local Slavonic version of, say, a Greek alphabet (whether uncial or minuscule)? The only other hint is that Constantine is reported not to have seen these before, but to have learnt to read them surprisingly quickly. Does this suggest that he recognized the language beneath them and simply not the letters? Or that, as the legend implies, he was simply brilliant at learning a whole new language? The most popular view has been that the name of the letters has been corrupted, and originally read *sourskymi pismeny*, that is Syriac, the only problem being that Constantine is elsewhere credited with already knowing, or at least being familiar with Syriac, so that his effort at learning these 'new' letters is somewhat diminished. In the circumstances of the eulogic way in which his character is built up in this 'Life', the 'Syriac' explanation is unsatisfactory. By the same token, there is no other explanation for this phrase which fits either: the letters could not be Greek, Hebrew or Armenian since Constantine would likewise have recognized these; it remains possible that what he saw were corrupted forms of one of these alphabets, and that his 'feat' was to decipher the corrupted letters and reconstruct their original Greek (or whatever) form. Given that the document was allegedly a Gospel, he would already know the content, and thus be able to give the impression of understanding the new language. This interpretation is tempting alsoin that it could suggest the origin, at least in principle, of the alphabet which Constantine is credited with creating a few years later, and it brings us to the central question of the appearance of an entirely new alphabet applied

to a Slavonic language. (Another view places the creation earlier, in the mid 850s; see below.)

1.2 Constantine

While there is taken to be no doubt that Constantine was the prime mover in the 'creation' of a Slavonic alphabet, there is an immense amount of doubt about every detail of this business, especially on the formal questions like: did he 'create' an alphabet, in the sense that he dreamt up the forms from scratch? If so, which alphabet did he 'create'? Did he 'create' two alphabets, or was one 'created' by someone else? Did he adapt some existing alphabet to Slavonic needs? If he adapted some other one, which was it? And so on.

Questions of principle (Why? Who for?) are less crucial, but merit a brief review before we take up the formal ones. The traditional view is that the alphabet (whichever it might have been) was created specifically in response to the Moravian request to Byzantium for a mission. Scepticism about this has centred around the speed with which everything was done, apparently no more than a year having passed between the request and the mission, a short time for the creation of an excellent alphabet plus the translation into a Slavonic language, using this new alphabet, of at least the Gospels. The only response has been that Constantine's philological interest might have led him to 'play' with an alphabet before this.

Two further points have been made: (1) only a native Slav, and not even a Greek philologist, would be likely to engage so seriously in the alphabet creation, which was hardly 'play'; (2) the translation of the Gospels into a vernacular was dubious, only the three 'sacred' languages being fully acceptable, and it is doubtful that a highly placed Greek would lightly consider it. (True, there were precedents in the Eastern Church in the Coptic and Armenian rites.) A recent work (Hoffer Edle and Margaritoff 1989) takes this further; assuming on the basis of the above that Constantine was Slav, the claim is made that he was in fact ethnically a Bulgarian, though born in Byzantium (Salonica), and this heritage led him to be interested in the cultural freeing of Bulgaria from Byzantine influence, complementing its political and military independence. The later return of the missionaries to Bulgaria and not Constantinople is taken to support this. In sum, it is claimed that the alphabet was created (as early as 855) as the first step in allowing Bulgaria to develop its own culture. The general idea that Constantine was Slav, and had been working on an alphabet for the Slavs, has been common amongst Bulgarian historians.

The most persuasive aspect of this view is that it offers an answer to the awkward question of the acceptability of the vernacular at this date: while the use of the vernacular was fine for missionary activity in general, its use in translating the Church books was quite another matter; the subsequent

Table 2.2: Old Slavonic alphabets

Cyrillic (OCS)	*Glagolitic* (OCS)	*Name* (*in transliteration*)	*Transcription* (*transliteration, if different*)	*Modern Cyrillic*
а	ⰰ	azъ	a	а
б	ⰱ	buky	b	б
в	ⰲ	vědi/vědě	v	в
г	ⰳ	glagoli/glagolь	g	г
д	ⰴ	dobro	d	д
є	ⰵ	jestь/estъ	e	е
ж	ⰶ	živěte	ž′ (ž)	ж
ѕ ʒ	ⰷ	(d)zělo	dz′ (dz)	Mac. s
з	ⰸ	zemlja	z	з
и	ⰹ	i, ižei	i	и
ї	ⰺ ⱅ	iže	i	Ukr. i
ћ	ⰻ	g′ervь/d′ervь	g′/j (ǵ/j)	SCr. ћ
к	ⰼ	kako	k	к
л	ⰽ	ljudьje/ljudije	l	л
м	ⰾ	myslite/myslěte	m	м
н	ⰿ	našь	n	н
о	ⱀ	onъ	o	о
п	ⱁ	pokoi	p	п
р	ⱂ	rьci	r	р
с	ⱃ	slovo	s	с
т	ⱄ	tvrdo/tverdo	t	т
оу ȣ	ⱅ	ukъ/ikъ	u	у
ф	ⱆ	frtъ	f	ф
х	ⱇ	xěrъ/xerъ	x	х
ѡ	ⱈ	otъ	o	(о)
ц	ⱉ	ci	c′ (c)	ц
ч	ⱊ	črvь	č′ (č)	ч
ш	ⱋ	ša	š′ (š)	ш
щ	ⱌ	št′a	št′ (št)	щ
ъ	ⱍ	jerъ	ŭ/ə (ъ)	ъ
ы ъи	ⱍⱏ ⱍⱅ	jery	y	ы
ь	ⱍ	jerь	ĭ (ь)	ь
ѣ	ⱎ	ětь/jatь	ä (ě)	(е)
ю	ⱏ	ju	ju	ю
ꙗ	–	ja	ja	я
ѥ	–	je	je	е
ѧ	ⱑ	jusъ malyj	ę	(я)
ѩ	ⱒ	jusъ malyj jotirovannyj	ę/ję	(я)
ѫ	ⱓ	jųsъ bolьšij	ǫ	(у)
ѭ	ⱔⱑ	jusъ bolьšij jotirovannyj	jǫ	(ю)
ѯ	–	ksi	ks	(кс)
ѱ	–	psi	ps	(пс)
ѳ	ⱘ	thita/fita	f	(ф)
ѵ	ⱙ	ižica	i/v	(и/в)

Note: Double forms in columns 1 and 2 are free variants. Variant names in column 3 (separated by /) reflect local differences. The transliteration is given in parentheses in column 4 only where different from the transcription, and corresponds to that used throughout this book. Transcription and transliteration are discussed later in this chapter. Column 5 gives in parentheses the most frequent equivalents where the actual form no longer occurs anywhere.

trouble in Moravia was clearly centred on that problem, the Slavonic liturgy being seen as heretical. Of course, practical Church (and state) politics could and did bend the rules, as is apparently the case in the Byzantine acceptance in 863 of Constantine's alphabet and translations (with some precedents as noted above), and in the Roman acceptance in 869 (with no precedents), but it is interesting to contemplate the notion that Constantine – who was as yet only a scholar and official, and not religious, becoming a monk only in 869 – might have *secretly and unofficially* applied his existing alphabet and (effectively illegal) translations to the Moravian business, though they were intended to be used for non-religious purposes in Bulgaria. This would not necessarily conflict with the early historical records of what happened, as these could easily (prefer to) recognize the later official acceptance as having preceded the unofficial use, even if they were aware of this unofficial use. General awareness of the problem at the time is seen in the report in the *Life of Constantine* of the debate in which Constantine became engaged in Venice (while on the way to Rome) precisely on this matter; or in Chrabr's *On the Letters*, equally clearly a defence of the use of a Slavonic alphabet and language for liturgical purposes.

The formal problems are many, and they all centre round the one fact, that there are *two* alphabets both clearly 'created' to fit Slavonic needs: Glagolitic and Cyrillic. If only Cyrillic existed, there would have been little trouble: it is clearly based on uncial (capital) Greek, and the problems would have been reduced to determining the origin of the letters which could not have come from Greek, like those representing the sounds /ž, š, č, c/, which Greek did not have. It should be said, incidentally, that even determining the origin of these letters would not have been as simple as it looks, as their origins in either alphabet are by no means unequivocally proved, and this is especially the case if one tries to omit Glagolitic from the equation. We shall return to this question after some discussion of Glagolitic. Table 2.2 gives a parallel list of forms and values. The letter-names are included partly out of simple interest, but mainly because some letters will be referred to by name later in the text. Discussion of the origin and development of the names may be found in Cubberley (1988).

The search for the formal origins of Glagolitic has occupied Slavists for well over a century, and remains unsolved. There are still attempts at new

solutions (the present author being guilty of one in 1982), and still no fully accepted view. One very popular view is that Glagolitic is a *totally individual creation* 'from scratch', the corollary, of course, being that the creator was Constantine; the advantage of this view is that it appears to obviate the need to find a formal model in some other alphabet – I say 'appears' because I, amongst many others, find it difficult to believe that any normal person with no ulterior motive would rather work from scratch than adapt some existing system; and even if there is an ulterior motive, say the need to disguise the source or the application, it is still more natural to work from a real base. Moreover, it is said that as Constantine was a philologist this approach would have appealed to him, but it is precisely as a philologist that he would have had access to many existing systems to use as, at least, a notional base. Thus the 'genuine creation' theory seems to me formally somewhat unsatisfactory.

Most popular is the view that Glagolitic is based on Greek cursive forms. This view has been around since the last century and, while there is much uncertainty about many of the derivations, the general principle seems provable, that is most Glagolitic letters can be derived from Greek cursive forms in a way that is formally satisfying. Moreover, there is the circumstantial evidence that such a use of Greek forms is logical, first because Constantine was either Greek or bilingual in Greek, second, if the Slavs in closest contact with the Greeks were writing their languages, it would logically be in Greek letters, and finally, we have the enigmatic 'Russian letters' mentioned above, possibly suggesting that Constantine observed some such use of Greek letters and noted it as a model. In this view Constantine's role is mainly that of formalizer: he would have settled on formal variants for each sound needed; his creative contribution would have been in the addition of letters needed for non-Greek sounds, especially the palatals. This scenario also answers the protest that Constantine would not have used *cursive* Greek for the Gospels, as this was unacceptable Greek usage (in that for Church books they used either the uncial or the minuscule forms): he was formalizing an *existing* usage, and not establishing his own. Indeed, it is when one turns to an explanation of the appearance of *Cyrillic* that this argument becomes pertinent, and may even be used to support the above scenario: if Constantine's disciples were to start thinking in terms of the *dignity* of Church books rather than the more philological question of a Slavonic alphabet, and if they *knew* that Glagolitic was based on cursive forms, then they would logically have opted for a 'new' form based on the 'more dignified' uncial forms. Of course, as indicated above, any argument of dignity or acceptability militates against the whole idea of using a Slavonic vernacular or alphabet in the Church books anyway, but at least by the time of the Bulgarian period some official recognition had been given by Rome, albeit fleetingly.

Many other 'sources' have been suggested for Glagolitic, in fact just

about any other alphabet which was around at the time (for a list see Истрин/Istrin 1963: 65); however, none of these has been as generally accepted as the cursive Greek theory. The most interesting variation is that the source for Glagolitic was Cyrillic. This found many important supporters earlier this century, the most notable being Karskij (Карский 1928/1979: 249–50), and more recently has been put by Istrin (1963: 147; 1988). It is generally now rejected on various grounds, both formal and logical: (1) while one can see Cyrillic as only indirectly derived from Glagolitic, in that the model here was basically the existing Greek uncial variant, the opposite derivation of converting to a cursive form is not acceptable, so one is stuck with a formal derivation, which requires considerable inventiveness on the part of the interpreter, and implies the same inventiveness on the part of the creator; and (2) what could be the motivation for rejecting Cyrillic in favour of a much less 'dignified' script? The only answer which makes any sense here is 'as a code': holders of this view (including Karskij and Istrin) argue that this was deemed necessary in Moravia after the proscription of the Slavonic liturgy (after the death of Methodius in 885), where the new script would have served to disguise the Slavonic content. However, it does not seem likely that the marauding German clergy would be fooled for long by a script that no one had seen before; what else could they think it was? And how many of the underground Slavonic brethren could have learnt to use this script in the circumstances? Altogether, this seems an unlikely sequence of events, and one is left with the conclusion that Cyrillic cannot have predated Glagolitic.

Other circumstantial arguments put forward to support the Glagolitic before Cyrillic order include: the existence of palimpsests (reused manuscripts) with Cyrillic superimposed on Glagolitic, but none in the other direction; the identification of local features which unite the Macedonian area with Glagolitic (for example no Turkisms) and the Bulgarian area with Cyrillic (Turkisms); and the (supposed) superiority of Glagolitic as representative of the early Slavonic (Macedonian) phonological system. None of these features is really of any clear significance, and all have been challenged even factually.

Let us assume, then, that the order of events is: Glagolitic is formed by the adaptation of cursive Greek by some Slavs during the preceding couple of centuries (Istrin accepts the possibility of such a 'Proto-Glagolitic', which at best would have helped in the conversion of Cyrillic to Glagolitic – 1963: 147); it is formalized by Constantine, who also adds letters for the non-Greek sounds; Constantine's disciples in Bulgaria perceive Glagolitic as unsuitable for Church books and make a new Slavonic alphabet based on uncial Greek. The only remaining formal questions are then: where did the added letters come from? And can we satisfactorily relate the Glagolitic and Cyrillic versions of these (that is, can we derive the Cyrillic ones from the Glagolitic)?

Of the many Slavonic sounds not existing in Greek, the most obvious are the palatals – /š, ž, č/ – but also /c/ and /b/, and of course many vowels, especially the nasals, *jers* (mid-high reduced) and *jat'* (low front) (see table 2.2 for names and symbols). Very little attempt has been made at finding sources for the vowel letters; most attempts at finding sources for the palatals and /c/ offer multiple sources, for example Coptic for /ž/ and Hebrew for /š/, /č/ and /c/. In the belief, mentioned above, that Constantine, as a philologist, would have used some consistency in his choice of sources, and as far as possible seek a single source for all of these sounds, I have argued elsewhere for Armenian as such a source for the consonants, and Greek variants for the vowels. The details may be found in Cubberley (1982), and here I mention only the three main issues relating to Armenian: (1) Constantine would almost certainly have known at least the Armenian alphabet (if not some of the language), there being many highly educated Armenians living in Constantinople, possibly including some of his colleagues at the 'university', where he taught philosophy, like John the Grammarian and Leo the Philosopher, who was also head of the institution, and at least one recent emperor having been Armenian (see Charanis 1961: 211); (2) Armenian had a wealth of palatal sounds, more than Slavonic, with letters to represent them all; (3) it is possible to make formal associations between these and the corresponding Glagolitic letters (except for one – /š/, for which one must argue via some confusion over the letter representing the reflex of PSl. **tj*, resulting in the later formation, in Cyrillic, of the form **ш**, which was then borrowed back into Glagolitic (discussion in Cubberley 1982: 299–302)).

As to the Cyrillic versions of these Slavonic sounds, it is not too difficult to see enough similarities to manage a derivation from Glagolitic. Of course, one could argue in either direction, but for the reasons listed above we are now assuming Glagolitic primacy. Thus for example we can derive Cyrillic **ж**, **ч**, **ц** from Glagolitic ⰶ, ⱍ, ⱌ respectively; **ш** is the same in both; for the vowels we derive ъ, ь, ѫ, ѧ from ⱏ, ⱐ, ⱘ, ⱔ respectively, while the symbol originally used for /ě/ (Cyrillic ѣ, Glagolitic ⱑ) has been confused through the many changes and local reflexes of this Proto-Slavonic sound (see Cubberley 1984: 284–5).

1.3 End of the ninth century

I thus favour the view that at the end of the ninth century Constantine's disciples, many of them nameable, such as Kliment Oxridskij, Naum Preslavskij, Konstantin Preslavskij, 'created' the alphabet now known as Cyrillic on the basis of the (more dignified) Greek uncial script, using Glagolitic as the model for the Slavonic-only sounds, to some extent giving them a 'square' look to match the uncial style of the rest. This period was that of the First Bulgarian Empire, with the strong Car' Symeon in charge, and a generally pro-Greek attitude, at least in matters cultural.

The question of the naming of the two alphabets is really a minor one, and is probably most simply explained by a confusion in the reporting of the creation of 'the alphabet', since no early source talks of two alphabets; both are referred to, if at all, as 'bukvica', 'azbuka', etc., with no further qualification. Only much later did either name, whether that of Cyril (*kirillica*), from Constantine's adopted monastic name, or Glagolitic (*glagolica*), from *glagol-* ('word, say'), become attached to one or the other alphabet. The name *glagolica* appears to have developed in the Croatian area – probably in the fourteenth century – from the name *glagolity*, applied to adherents of the Slavonic liturgy. In the South Slavonic area in general, but especially where Glagolitic remained active, there arose in the seventeenth century the legend that Glagolitic had been created by St Jerome (Hieronymus), and the two alphabets are typically listed side by side as the alphabets of St Hieronymus and St Cyril. This may well be the start of the association of Cyril with Cyrillic, which then spread to the East Slavonic area. The names *glagolica* and *kirillica* are attested there only in the nineteenth century.

1.4 Tenth–eleventh centuries

In the First Bulgarian Empire, which lasted until 1018, when it was militarily defeated by the Byzantine Empire, Cyrillic and Glagolitic must both have flourished, though we have, in fact, no original documents from that period; the number of later copies, however, testifies to the strong tradition which was established there. There is also some epigraphical evidence for both (see Велчева/Velčeva 1989). This tradition spread first into Serbia, possibly during the tenth century, or perhaps only in the next – the eleventh-century Glagolitic *Codex Marianus* is thought to be of Serbian origin (Ивић/Ivić 1986: 111), and a recently found pottery inscription from Kosovo is claimed to be from the tenth century (*Borba* 6.3.90); also into Bosnia, where a Glagolitic tradition lasting into the thirteenth century has been claimed (Kuna 1977) and even Croatia, where the two coexisted until the Reformation (Jurančič 1977); and then, more importantly for its subsequent fate, to Rus′, officially in 988, when Prince Vladimir formally adopted Orthodox Christianity as the state religion (though there were certainly conversions before that date, for example Princess Ol′ga, daughter of Igor′, is supposed to have been baptized around 957). For Rus′, the eleventh century was one of intense 'literary' activity, in the sense of large-scale copying of the Bulgarian books (as well as some original translation from Greek). *Ostromir's Gospel* (dated 1056) is one of the earliest and finest examples of what might already be called Russian (Rus′-ian) Church Slavonic, using a classic square uncial Cyrillic (see figure 2.2(a) on page 34).

On Old Church Slavonic as such, see chapter 4 of this book; on local versions and the role played by Church Slavonic in various areas, see the

collections by Schenker and Stankiewicz (1980) and Stone and Worth (1985). For our purposes, I would like only to mention that all three major alphabets were used even for early Old Church Slavonic documents, for example the following, none of which is later than the eleventh century:

Glagolitic: *Codex Zographensis* (Bulgaria), *Kiev Fragments* (Moravia)
Cyrillic: *Ostromir's Gospel* (Rus′), *Savvina Kniga* (Macedonia)
Latinica: *Freising Fragments* (Slovenia)

2 Glagolitic: later history

After the initial period, what one might call the Constantinian period (to the end of the ninth century), there is some evidence (mainly epigraphic) of the continued existence of Glagolitic, alongside Cyrillic, in the Bulgarian/Macedonian area, around the centres of Preslav and Ohrid, until the beginning of the thirteenth century (Велчева/Velčeva 1989: 21). However, Cyrillic steadily became dominant throughout the twelfth century. As noted above, Glagolitic also survived briefly in Serbia, probably into the twelfth century, and in Bosnia possibly into the thirteenth. Its subsequent history, however, belongs almost exclusively to the Croatian area.

2.1 Rus′

In Rus′ there are a few early (eleventh-century) examples of Glagolitic graffiti in Novgorod (Vajs 1937; Медынцева/Medynceva 1969), indicating only that it was known there, but apparently not much used; possibly it travelled north with some of the Bulgarians who were brought, especially by Jaroslav the Wise in the eleventh century, as we know from the *Laurentian Chronicle*, to undertake translation and teaching activity, but another view is that the source was Bohemia (Štefanić 1963: 29). It may have been used or passed on in Rus′ as a curiosity or even as a cipher.

2.2 Bohemia

Glagolitic may have arrived in Bohemia even before the death of Methodius, and probably continued to be used till the late eleventh century (Štefanić 1963: 28); later, there was a period of Glagolitic activity at the Emmaus Monastery from the fourteenth to the early seventeenth century, the active period lasting only till the mid-fifteenth century (Mareš 1971: 187–90). The source of this was Croatian in any case; that is, it was a secondary development of Croatian Glagolitic. However, it did produce important texts like the *Reims Evangelistary* and the *Czech Bible* of the fifteenth century, and served to convey the views of the Hussites back to the Croatian area (Hamm 1974: 41–2).

2.3 Poland

From Bohemia, Glagolitic was taken to one monastery in Poland (Kraków) at the end of the fourteenth century and may have survived there too till the sixteenth century; however, only fragments remain as evidence of this sojourn (Hamm 1974: 41–2).

2.4 Slovenia

Slovenia too had a flirtation with Glagolitic, probably likewise of a secondary nature, through Croatian influence, in the fifteenth–sixteenth centuries (Kolarič 1970); some believe that Glagolitic remained known in this area from the original (ninth century) Pannonian period, and is thus not a secondary product (Zor 1977).

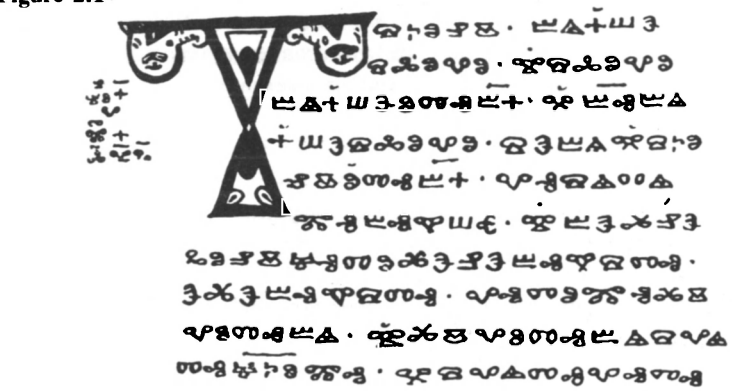
2.5 Croatia, Dalmatia

Croatia and Dalmatia were the areas in which Glagolitic not only survived, but flourished for many centuries, having arrived in Croatia probably by the tenth century (Štefanić 1963: 31), though others have dated the arrival later (Велчева/Velčeva 1989: 18). The apparent reasons are somewhat paradoxical, in that these were the areas dominated from early on by the Roman Church (especially after the Schism of 1054, when contact with Byzantium became more restricted), so that one would expect Latinica to have been *de rigueur*. In fact, Glagolitic became the symbol of (partial or nominal) independence from Rome; it was tolerated by Rome as a small concession permitting its continued influence where it mattered (in this case in the otherwise Byzantine-dominated Balkans), and finally made official in the mid-thirteenth century.

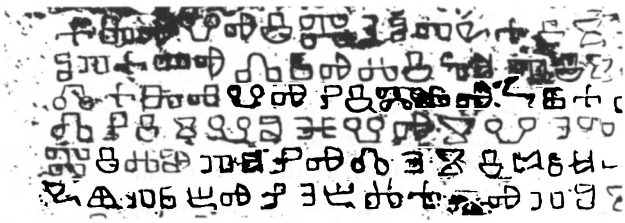
Formally, there was a gradual but marked change in the letter shapes: from the original round style, there was a shift first to a slightly more square shape, and finally the typical Croatian very square shape. (See figure 2.1 (a/b/c) for examples.)

Glagolitic continued to be used in Croatia until the early nineteenth century, especially on the Adriatic islands, during that time having acquired a cursive form, as it was used in administrative functions also (see figure 2.1 (d)), and having been printed in several major centres, like Venice, Tübingen and Rome (the earliest is a Missal of 1483 (place unknown), and other important examples are a Primer of 1527 (Venecija), a Testament of 1562 (Tübingen) and a Missal of 1631 (Rome)). As late as 1893 a Missal was printed in Rome, and Glagolitic was still used within the Church until the 1920s (Ивић/Ivić 1986: 117 places the end point at 1927, when a Latinica edition of the Glagolitic Missal was produced). However, Glagolitic ceased to be very active outside the church from the seventeenth century.

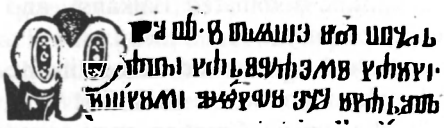
Figure 2.1



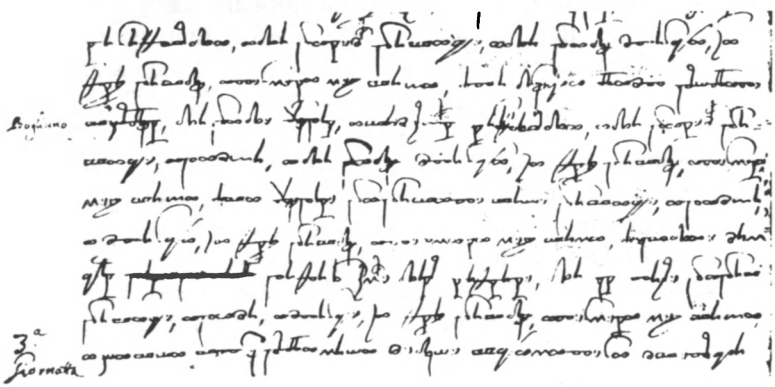
a. Round Glagolitic: eleventh century, Bulgaria



b. Transitional Glagolitic: twelfth century, Croatia



c. Square Glagolitic: thirteenth century, Croatia



d. Cursive Glagolitic: sixteenth century, Croatia

3 Cyrillic: later history

Cyrillic remained ensconced in the three basic areas in which it first developed: Rus', Bulgaria and Serbia. Its further development is of a marginal nature – partly related to local phonological changes and partly to purely graphic ones. Only in the case of its application to non-Slavonic languages were there any major changes in its form. After looking at the main Slavonic developments in the three basic areas, we shall briefly consider the non-Slavonic situation. In each case, at issue are both stylistic developments, affecting the alphabet as a whole, and purely formal ones, related to graphic and orthographic reforms (whether official or spontaneous), that is the exclusion of letters or introduction of new ones and changes to shapes.

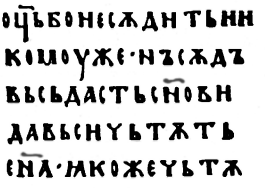
3.1 Style

In all areas, the initial square *uncial* style (Ru./Bg./SCr. устав/ustav), as seen in *Ostromir's Gospel*, gave way by the fourteenth century to a less square and slightly more irregular style referred to as *semi-uncial* (полуустав/poluustav), at first in less religious contexts, but eventually even in fully religious ones. (See figure 2.2 for examples.)

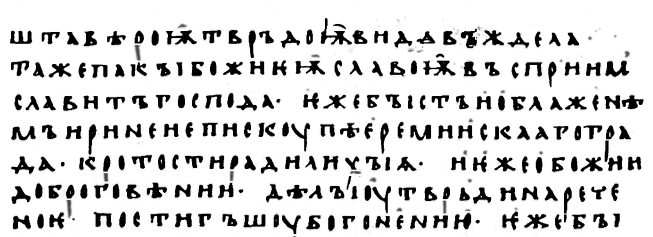
As more and more completely non-religious works came to be written, notably administrative documents of one sort or another, we observe a steady 'corruption' of the letter shapes: a new *cursive* form gradually takes shape. The period of development of this form relates directly to the status of the administration in given areas, thus it is earlier in the Grand Duchy of Lithuania (where Belorussian was used in this role) and the Kingdom of Serbia (both from the fourteenth century), but later in the Russian and Bulgarian areas. The Bulgarian area is the one in which the cursive forms never really developed, as their language was rarely used in the administrative role during the Byzantine and Turkish periods; it is only in the nineteenth century that we find large numbers of examples. In the Russian area, it is not until the fifteenth century that we can identify such a style (Ru. скоропись/skoropis'). For the East Slavonic area it may be thought somewhat paradoxical that it was in the non-Russian area of the Grand Duchy of Lithuania that this cursive style was first developed, but the reason is that the use of Belorussian and its alphabet as the administrative language was not tied in the same way to the more serious, religious applications of Cyrillic, as was the case in Muscovite Rus', as well as the fact that the latter's administration itself was developing more slowly. At any rate by the end of the fifteenth century we have a more or less united cursive throughout the East Slavonic and Serbian areas, with only one or two local features, such as flourishes on particular letters. This similarity should not surprise us, as both the starting point and the motivation for change were the same in all areas (see figure 2.3).

Apart from the Bulgarian area, where, as mentioned, the cursive forms

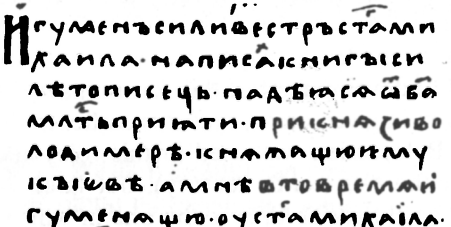
Figure 2.2



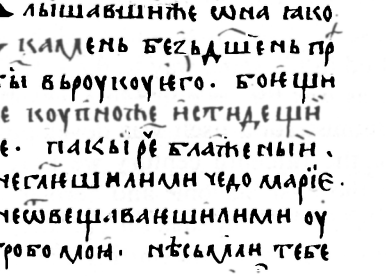
a. Ustav: eleventh century, Rus′



b. Ustav: eleventh century, Bulgaria

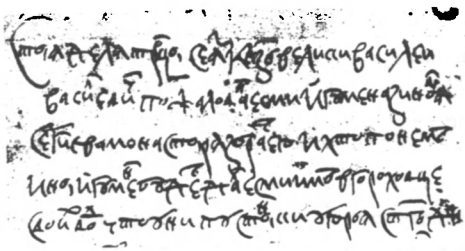


c. Poluustav: fourteenth century, Russia

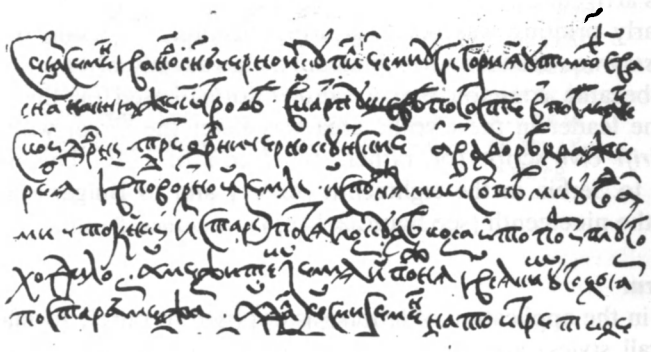


d. Poluustav: fourteenth century, Bulgaria

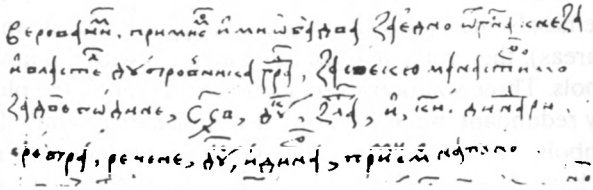
Figure 2.3



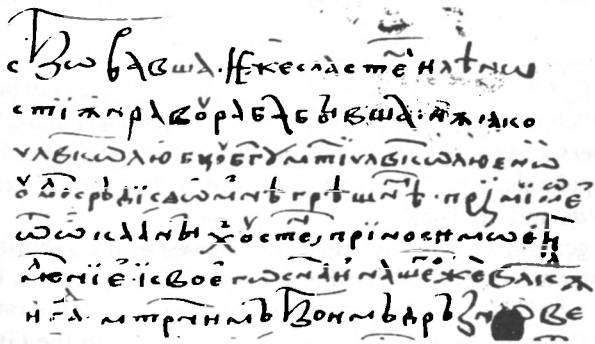
a. Cursive Cyrillic: fifteenth century, Belorussian



b. Cursive Cyrillic: sixteenth century, Russia



c. Cursive Cyrillic: sixteenth century, Dubrovnik



d. Cursive Cyrillic: sixteenth–seventeenth century, Bulgaria

did not have a chance to develop in any formal way (though at many levels of written interaction the same process must have occurred – see Figure 2.3 (d)), there is one area where the administration adhered to a much more formal semi-uncial style – Bosnia. This peculiar style, which, however, is very similar to the early printing style which was developed in Belorussia by such as Francišak Skaryna, is referred to as *bosančica* (see figure 2.4 (a/b)).

Cursive style is by definition handwritten and rather idiosyncratic, in spite of the degree of normativization which still occurs at the levels of teaching and formal application, and so it was not appropriate for printing when this arrived. For this more formal shapes were necessary, and as most of the early printing was religious in application, a version of the most formal *ustav* was the first norm in all areas. It was only when administrative printing became common that a simpler form, based on the *poluustav*, arose. The leader in the Cyrillic area was Peter the Great with his *graždanskij šrift* 'civil script' (or, rather, 'typeface'), of 1710. This in turn was exported to Serbia in the eighteenth century and to Bulgaria and Macedonia in the nineteenth (see figure 2.4 (c)).

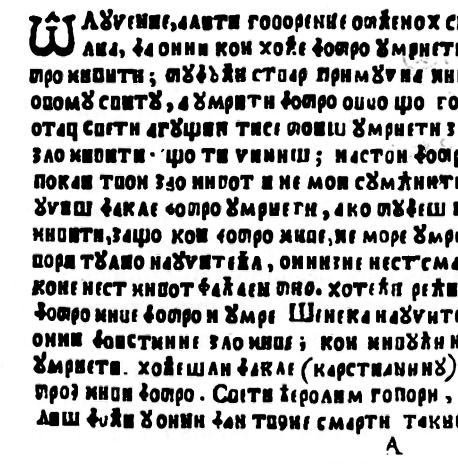
3.2 Form

Changes in the application of letters (as opposed to purely formal changes and overall styles) are usually the result of phonological changes which cause either redundancies or inadequacies in the symbolic representation of the sounds. Thus, for example, when the nasal vowels of Proto-Slavonic ceased to be nasals and fused with other, existing vowels (as was the case in all Cyrillic areas), the problem arose of what to do with the now redundant 'nasal' symbols. Three solutions were possible: (1) retain the old symbols as phonetically redundant, but etymologically informative, symbols; (2) delete the old symbols, replacing them with the ones used for the sounds with which they have fused (e.g. ѫ > у (u), ѧ > ѩ/я (ja); (3) apply the old symbols to some new function for which a new need has arisen (for this there will usually need to be some association between the old and new functions).

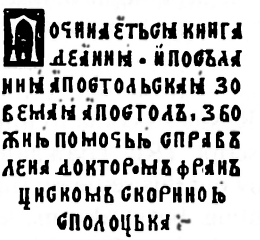
All of these approaches were applied at some stage in the various reforms of Cyrillic in every area. For all areas the following principal phonological changes caused such problems: for vowels, denasalization, the loss of the *jers* and the development of /ě/; for consonants, the development of the palatals, especially the complex ones, like the reflexes of Proto-Slavonic **tj/dj*, and /j/ itself.

Another general cause of trouble, of course, is the inheritance of an alphabet that was created for another language system, as was the case for Cyrillic in general (and to a lesser extent Glagolitic), with the rather serious problems of the multiple vowel symbols brought in from the Greek system (where they were by medieval times equally redundant!): the three letters

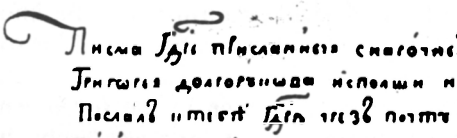
Figure 2.4



a. Bosnian Cyrillic (Bosančica), seventeenth century



b. Printed Cyrillic: religious, sixteenth century, Belorussia



c. Printed Cyrillic: civil, eighteenth century, Russia

for /i/, the two for /o/ and the variants for /u/. There were similar problems with the Greek consonantal letters, with two each for /z/ and /f/ (see table 2.2).

For all areas the solution to all these problems was a very long time coming: the earliest solution by far was that of Serbia (or rather of the whole area where Cyrillic was subsequently used for Serbo-Croat), namely that of Vuk Karadžić in 1814 and 1818; for Ukrainian the final (modern) version arrived only in 1917; for Russian in 1918; Bulgarian in 1925 and 1945; Belorussian in 1933; Macedonian in 1944. Up to the eighteenth century there was very little in the way of 'conscious' reform. One can identify only isolated conscious attempts at rationalization at earlier periods, but even the 'grammarians' of those periods, for example Konstantin Kostenečki (a Bulgarian living and writing in Serbia) in the fifteenth century or Maksim Grek (a Greek living and writing in Russia) in the sixteenth (both of whose 'grammars' were well known throughout the Cyrillic Slavonic world), in principle accepted problems like redundant letters as normal, based as much on the sacredness of the written form as on what might have been seen as important etymological information. It is rather in the practice of religious copyists and then secular clerks that we find rationalizations like the separation of ѩ and ѧ by phonetic context: ѩ initially and post-vocalic (= [ja]), ѧ post-consonantal (= [a] after soft consonant).

The first real attempt in Russia to come to grips with these problems was Peter the Great's 'civil script' (Ru. гражданский шрифт/graždanskij šrift) of 1708/10: not only did this settle on simpler forms of all letters for use in administrative printing, but it also made a start on deleting redundant letters which were marked as 'Church' variants, and also introduced some new forms, notably э and я. Э was to serve the needs of foreign words, where the sound [e] occurred initially without any prothetic [j], and was done by employing a little-used variant of е, which some regard as the Glagolitic letter ⰵ; however, no Glagolitic form would be known at this stage in Russia; the variant involved had in fact been used for some time in the southern (Ukrainian) area, and before that in the fourteenth century for Greek words; the shape of я had arisen in the seventeenth-century cursive style, especially for ѧ, but also for ѩ, and was probably seen as a useful unified form of these two letters, which by now represented effectively the same vowel sound ([a]), though varying by position, as indicated above. At the same time, by no means all the redundancies were removed in this reform: double symbols continued to exist for /i/, /f/, /z/; also this system was by no means a universally accepted one, tied as it was at first to administrative usage. The debate about shapes and variants was to be conducted fairly freely until 1918, basically because in Russia, as elsewhere, there was as yet no mechanism for the enforcing of a norm.

The Russian Academy of Sciences, founded in 1725, offered potentially such a mechanism, and indeed had such a role in view for itself, but its attempts at reform were few and not particularly effective, especially in the sense that individual writers, whether as theorists or creative writers, went their own way, not recognizing the dictates of the Academy. The 'final' formal reform came only in 1918, when the new regime included orthography in its range of general reforms: only then were redundant letters like і, ѵ, ѣ, and ѳ completely removed, and ъ in its redundant final position.

This Russian scenario is typical of the whole Cyrillic area, in that orthographic reform really relied on the acceptance of a literary norm for the whole language, and generally the latter was worked out only during the nineteenth century. This was certainly the pattern also for Ukrainian, Belorussian, Bulgarian and Macedonian.

For Ukrainian and Belorussian the main differences from Russian reflect the different behaviour of the vowel /i/ and its variants: since Ukrainian early fused the back and front variants [y] and [i], it had no need for two letters for those, and they became written both with и; however, it then developed a new high front phoneme /i/, and used for this the 'redundant' і rather than recast everything in the Russian style (ideologically not a serious option). For Belorussian the phonemic situation is the same as Russian; however, they opted for the і letter rather than и for the front [i]. Examples are the words for 'son', 'blue', and 'hay':

	'son'	*'blue'*	*'hay'*
Russian	сын [sy-]	синий [s'i-]	сено [s'e-]
Ukrainian	син [sy-]	синій [sy-]	сіно [s'i-]
Belorussian	сын [sy-]	сіні [s'i-]	сена [s'e-]

A second area of difference lies in the presence in both Ukrainian and Belorussian, but not Russian, of the semi-vowel [u̯]: in both it is a positional variant of either /v/, /l/ or /u/. While Ukrainian makes use of the existing letters for those sounds, Belorussian shows the innovation of the letter ў; to some extent this reflects Belorussian's different attitude towards orthographic principles (see below, section 8); for details refer to the relevant language chapters.

For Bulgarian one of the central issues of the orthography debate has been the phoneme /ă/: etymologically, it comes from Proto-Slavonic *ъ, *ǫ and even *ь, so that all of the three Old Church Slavonic letters ꙗ, ѫ or ь have at some time been used for it, and not always etymologically either. In the early nineteenth century the letter ă was used for this sound from whatever source, then ѫ became the popular letter, though ъ was still in contention. From the middle of the century the etymological principle was applied,

with ѫ for etymological *ǫ and ъ for etymological *ъ/ь, and this usage survived effectively until 1945. In addition ъ and ь were still written in final position, as in pre-1918 Russian, though they no longer had any phonetic value. Only in 1945 was ѫ removed entirely, ъ written for /ă/ when internal, and value-less final ъ dropped; the last problem, that of final /ă/, was then handled by the use of а in that position. Examples are, of Proto-Slavonic, pre-1945 and post-1945 Bulgarian:

	'hand'	*'sleep'*	*'I can'*
Proto-Slavonic	*rǫka	*sъnъ	*mogǫ
Pre-1945 Bulgarian	рѫка	сънъ	могѫ
Post-1945 Bulgarian	ръка	сън	мога

(See Мирчев/Mirčev 1963: 25–33.)

Another issue of some difficulty has been the use of ѣ in its etymological place in competition with the dialectal reflexes (sometimes /a/, sometimes /e/, both by position and by dialect). The 1945 reform removed ѣ and settled for я/е as they occur in the new standard.

Macedonian had fewer problems of this sort, its only real phonetic peculiarity being the reflexes of Proto-Slavonic **tj* and **dj*, namely /ќ/ and /ѓ/ respectively. The nineteenth-century efforts at sorting out the orthography were influenced greatly by Vuk Karadžić's reforms (see below), and it was only the larger problem of the codification of the language which slowed things up, so that it was not until 1903 that a system virtually that of Vuk was made as formal as the situation allowed (that is, not very), through the efforts of writer and publicist Krste Misirkov. The main differences from Vuk's system lay in the use of the apostrophe to indicate palatal consonants rather than a ligature with ь, thus н', л', к', г'; he also used і rather than ј for /j/. Misirkov's system was not widely known or applied, but its principles were the ones used in the orthography officially adopted as part of the new standard language in 1945; at this point, the Serbian ligatures њ and љ were adopted, as well as ј, and the acute rather than the apostrophe in ќ and ѓ (see Koneski 1983: 111–15).

The case of the Serbian–Croatian area was different, probably, in that the question of orthography was actually at the forefront of the larger business of literary norms, and this was because of the parallel existence of at least two alphabets – Cyrillic and Latinica, but indeed even of three, since Glagolitic was still functioning in the Croatian area. It was Vuk Karadžić who led and carried out the crucial reforms of the Cyrillic alphabet. These reforms, described and put into practice in his grammar of 1814 (*Писменица сербскога језика/Pismenica serbskoga jezika*) and, especially, in his dictionary of 1818 (*Српски рјечник/Srpski rječnik*), were mainly concerned with the writing of the 'new' palatals /č'/, /dž'/

(from PS1. **tj*, **dj*) and /dž/ (foreign, from Turkish), and also of /j/ itself. For the first Vuk used a form which had been around for a long time in Serbian Cyrillic – ђ (see table 2.2, th letter called *g'ervь*, originally representing Greek /g'/ or /j/, then Serbian /j/). He used the recently invented longer-tailed variant ђ to mark the voiced equivalent and the form џ (inverted h or a variant of ц?) for the foreign sound (Младеновић/ Mladenović 1989: 156–7). The letter for /j/ – j – was, not surprisingly, taken from Latinica in its Croatian usage. Most subsequent discussion of orthography in the Serbian–Croatian area centred around the relationship between the two major alphabets and their application to the literary language, whose variants were the main point of issue in the establishment of norms. We will have some more to say on this in the discussion of Latinica developments in the area.

3.3 Non-Slavonic use of Cyrillic

While this book deals with the Slavonic languages, it may also be of interest to note that some non-Slavonic areas have used or still use Cyrillic, through Slavonic influence in the religious, cultural or political domains:

1 Rumania for all its early literary history used not only Cyrillic, but the entire Old Church Slavonic language, like the use of Latin in western Europe. The first examples of written Rumanian date from the sixteenth century, but the alphabet remains Cyrillic. It was only around 1860 that Cyrillic was replaced by Latinica in non-religious writing, and in the Church only in 1890.
2 Of the sixty-two non-Slavonic written languages of the former Soviet Union listed by Gilyarevsky and Grivnin (1970: 9), fifty-five use Cyrillic; many of the fifty-five have additional letters in their alphabet, sometimes from Latinica, often variations of Cyrillic, and often by the use of diacritics. The Turkic languages were written first in Latinica during the early Soviet period (many having previously used Arabic), but converted to Cyrillic in the late 1930s.
3 Mongolian has been written in Cyrillic in the Mongolian People's Republic since 1941 (though Mongolian in Inner Mongolia (in China) still uses the traditional Old Mongolian vertical script).

4 Latinica

4.1 Early history

If Glagolitic and Cyrillic, both based on Greek, had problems of redundancy, they nevertheless represented fairly early alphabets dedicated to representing Slavonic speech, and especially through the early creation of extra symbols for specifically Slavonic sounds, they both ended up being

rather good in this function. Latinica, on the other hand, in the early stages at least, had no one attempting to make it fit a Slavonic system; in all the Slavonic areas in which it became established early, there was no early attempt at writing the given Slavonic language, it being assumed that written documents, even secular ones, used Latin (language and therefore alphabet). When we do find examples of Slavonic words written in Latinica at this stage, the specifically Slavonic sounds clearly present an insurmountable problem, or at least one which no one attempted to surmount! By way of example, the earliest list of Slavonic letter names, dating from the twelfth century (known as the 'Paris Alphabet' or 'Abecenarium Bulgaricum'), has names like: 'ife, giuete, naf, faraue' (for iže, živěte, našь, čьrvь – see table 2.2), indicating no attempt, or at any rate complete failure, to come to grips with the Slavonic sounds [š], [ž], [č]. Another problem is that even in the writing of Latin, there were already local variations in the use of letters, but only one 'new' letter – the 'long' **f** – used at some stages in the early Middle Ages as a contextual variant of **s** like the σ/ς of Greek. By the late Middle Ages it would appear that this usage was confused, and in most cases the two forms **s** and **f** seem to be free variants. Certainly, in the early Slavonic examples there is no evidence of this symbol being used for a particular sound. In the Paris Alphabet, in addition to the above examples, where **f** represents all three palatals, it occurs also in 'af' (= azъ) and 'hieft' (= estь). The only innovation that we find at this stage (thirteenth century) is the (inconsistent) use of digraphs, like **ss**, **zz**, to indicate non-Latin sounds, specifically [š], [ž]. The different local varieties of Latin are reflected in the Latinica forms which became used for Slavonic languages when this usage was established in each area; thus, for example, the German (Gothic) tradition operated in Bohemia and Moravia, the Italian or Hungarian in Slovenia and Croatia; in Poland (as in Hungary) the use of digraphs remained particularly active.

4.2 Later history

The above situation continued up to the fifteenth century. There was still not very much written in Slavonic vernaculars till then, except in Bohemia, where the fourteenth century, especially under Charles IV, saw a considerable amount of local writing (for example, Chronicle of Dalimil, Legend of St Catherine). The early Polish hymn known as 'Bogurodzica' (thirteenth century) is isolated, but does indicate the use of digraphs. Let us now follow what happens in the various areas.

4.2.1 Czech and Slovak (Bohemia, Moravia, Slovakia)

In the fifteenth century we find the first serious attempt to go beyond the digraph system and make Latinica fit Slavonic in a more direct way: this is the system reputedly devised by Jan Hus and described in his *De Orthographia Bohemica* (of 1406?; see Schröpfer 1968), and it has been

improved on only in minor details. Its simple power is derived from the use of diacritics: Hus himself used dots over consonants for the palatals (e.g. **ṡ**, **ż**), and the acute for long vowels (for example **é**, **í**), while the modern Czech hook (the *háček* – the 'inverted circumflex', as in **š**, **ž**) appeared in the late sixteenth century. This system is so elegant and efficient that it is, in addition to being the standard orthography of Czech, Slovak, Slovene and Croatian, also the transcription system used by Slavonic linguists. (See below on transcription.)

Other general problems included the writing of /j/: the early tradition was to use **g** before front vowels and **y** before back, and this survived until the mid-nineteenth century, when **j** was established in this role (by Pavel Šafařík, a Slovak). At the same time the form **v** was fixed for /v/ in place of **w**. Phonological changes like the fusion of former variants [i] and [y] produced the usual problem of how far to retain etymological information as opposed to phonetic. Czech made the etymological decision early (eighteenth century) and has not retracted from it, retaining the letters **i** and **y** in spite of their phonetic identity (and it was followed in this by Slovak), while all other languages with the same fusion (Bulgarian, Ukrainian, Serbo-Croat, Slovene) have opted for phonetic spelling of this /i/. A parallel case of etymological spelling is the distinction between the two sources of [ū]: original /ū/ is rendered by **ú**, original /ō/ by **ů**. Thus the final Czech system was in place by the second half of the nineteenth century.

Slovak has not deviated far from the Czech system, having been able basically simply to drop unnecessary Czech letters like **ě**, **ř**, **ů**; its only innovation has been the vowel symbol **ô**, used to represent the diphthong [uo] (etymological [ō]). The etymological principle was accepted here too, by 1852, though only after proposals for phonetic spelling by notable names like Anton Bernolák and L'udovít Štúr (de Bray 1980, II: 132–3).

4.2.2 Polish and Sorbian (Poland, Lusatia)

The only Latinica Slavonic area not to adopt this diacritic system is the Polish one, apparently partly from initial resistance on religious grounds to borrowing the Hus system, but mainly perhaps because of its phonological system: while it has the usual collection of (old) palatals (/š/, /ž/, /č/), all now phonetically hard, it also has a new set of soft ones (from the soft dentals), and it may be that it could not come to terms with two (or worse, three) sets of diacritics. Thus Polish uses the acute to indicate the new soft palatals (**ś**, **ź**, **ć**), a usage established in the early sixteenth century, but digraphs for most of the old, namely **sz**, **cz**, **rz** (the latter an etymological spelling for phonetic [ž] from /r'/ and /rj/); for old /ž/ it retains the dot as in Hus's system – **ż**; and for both old and new /nj/ (now identical) it uses the acute – **ń** (see chapter 12 for details on the current situation). The

sixteenth-century digraph system still included **tz** or **cz** for /c/, and /š/ also appeared as **ss** (de Bray 1980, II: 231). While there were objections to the digraph system already in the fifteenth century, no doubt through the influence of the Hus system, Polish has not replaced its digraphs with the Czech *háček* symbols, and at this stage the problem is that of any language with an established literature behind it. English users are hardly in a position to be critical of the Polish spelling system in this respect, and the Polish system is as efficient as any other Slavonic language in terms of the correlation between sound and symbol.

As for the vowels, the nasals are the most typical feature of the Polish system, and their spelling by the symbols **ą** and **ę** is (morpho-)phonological, the surface realization being accessible by fairly consistent rules. These symbols date from the same period (early sixteenth century). So too does the one case of etymological spelling: the letter **ó**, which represents a former long /ō/, with the acute used as in Czech, but a modern short [u]. While long vowels were still around, in the fifteenth century, writing them double was a usage predating that of the acute, and perhaps matching the digraph consonant system.

For Sorbian, as for Slovak, the nineteenth century was the time for sorting out the orthography, though like Serbian and Croatian, there were major problems with the competing literary variants, not only Upper and Lower Sorbian, but also the Protestant and Catholic versions of Upper Sorbian. These major questions found their modern resolutions only in this century. The modern spelling system of both variants is diacritic, using both the Czech hook (for the old palatals) and the Polish acute (for the new palatals), and was developed by the mid-nineteenth century, being referred to then as the 'analogical' system (de Bray 1980, II: 342).

4.2.3 Croatian (and Serbian), Slovene (Croatia, Dalmatia, Serbia, Slovenia)

Apart from a few details of the application of certain Latin letters, these areas share the same tradition, originally that of Italian Latin. Thus, for example, /j/ was early represented by **g** or **i**; /č/ by **c** or **ch**, /š/ by **sc**, /c/ by **z**, etc. Local variations developed over the late Middle Ages in the representation particularly of the palatals. For /č/ northern Croatian (influenced by Hungarian) used **cs** or **ch**, southern Croatian used simply **c** and Slovene used **zh**; for /š/ northern Croatian used **f/s**, southern Croatian **sc** and Slovene used **fh/sh**; for /ž/, which did not occur in Italian, a new model was set up only in the Dubrovnik area: **sg** or **x**, while elsewhere there was the usual vague usage, northern Croatian again **f/s**, southern (Zadar) simply **s** or **z**, Slovene **sh** (see below on the use made of the 'long' s (f) as a visual differentiator in Slovene); for /c/ Croatian used **z** and (especially northern) **cz**, Slovene **c**. In so far as Latinica was used in

Serbia (mainly in the north), the only special usage was **x** for /ž/, though this was used, less consistently at first, in southern Croatia too, where it became common from the seventeenth century (however, Vuk Karadžić, in a comparative list of alphabets (*Srpski rječnik* 1818: lxix), lists **x** for /ž/ in the column headed 'Serb. Lat.', but for Croatian only **s** in this role, with **sz** for /s/, as in Hungarian; he also lists only **ch** for /č/ under 'Croatian').

The first area to start formalizing the language was Slovenia, where the late sixteenth century marked the appearance of a series of grammars, starting with that of Adam Bohorič in 1584. As far as spelling is concerned, these grammars fixed the tradition in use at that point. They did not embark on discussions of problems or suggest any changes. Some earlier non-grammatical works are of interest too in listing the letters before their text – usually biblical, for example Sebastian Krelj's *Otrozhia Biblia* of 1566. In these the items of interest are:

digraphs using **h**: **sh** for [ž] and [š], **zh** for [č], **ch** for [šč];
sh for [ž], but **fh** for [š];
s (low) is in fact commonly used for voiced [z] as opposed to **f** for [s], while **z** is used for [c], but so also is **c**.

Thus, overall, there was still a great amount of confusion in these about the use of Latin letters. This situation did not change much in published grammars for a long time: Marko Pohlin in 1768 and 1783 was still using the same system (known as 'bohoričica'), and so were Jernej Kopitar in 1808 and Pavao Solarič in 1814. None of these seems concerned about the orthography, and it was only in the 1820s that the debate hotted up into a 'war' between new systems proposed by Peter Dajnko ('dajnčica') and Franc Metelko ('metelčica'), the former proposing **ŋ**, **ɣ** (!) and **ч** for /nj/, /š/, /č/ respectively, and the latter a whole series of Cyrillic letters, some directly, notably **ш**, **ч** for /š/, /č/, also **щ**, **ф**, others in adapted form, notably **L** and **N** with hooks for /lj/ and /nj/ – modelled on Vuk's Cyrillic (see Jurančič 1977: 143–5). But it was, in fact, in Croatia that the orthography of both languages was sorted out: following Vuk's reform of Cyrillic (see above) in the early nineteenth century, Ljudevit Gaj in the 1830s performed the same operation on Latinica, using the Czech system and producing a one-to-one symbol correlation between Cyrillic and Latinica as applied to the Serbian and Croatian parallel systems. In turn, Gaj's system ('gajica') was adopted in Slovene in the 1840s, especially through the efforts of the editor Janez Bleiweis (see de Bray 1980, I: 235, 312). The modern systems date from this time.

4.2.4 Other Slavonic – Belorussian

After the Polish–Lithuanian Union of Lublin (1569) and the religious Union of Brest (1596), Belorussian came under direct Polish influence, at

the expense of the freedom it had had under Lithuania alone. At the level of orthography this led to the extensive use of Latinica to write Belorussian. By definition, the graphic system used was that of Polish, and this usage survived until early this century, when the Czech *háček* letters began to be substituted for **cz**, **sz** and **ż**, and **v** for **w**. Many publications at this time were printed in both alphabets. After the Revolution, the use of Latinica was virtually eliminated, though it survived in the western areas until the Second World War (see Mayo 1977: 29–31).

5 Greek

The consistent use of the Greek alphabet to write a Slavonic language is limited to the Macedonian area. It starts from the Turkish conquest in the late fourteenth century and continues through the later Greek domination in the nineteenth century, and, to the extent that any Macedonian is still written in Aegean Macedonia, which would be in private correspondence only, continues until now. In fact, however, since the First World War Greece has effectively not recognized any Slavonic Macedonian minority, and so there is no recognized need for the writing of their language. During the nineteenth century especially there was a large body of Macedonian literature written in Greek script; indeed the first book printed in Macedonian, in 1794, used the Greek alphabet (see Koneski 1983: 112–15).

The problem is the same as with Latinica: the Greek alphabet does not contain sufficient letters to cope with the extra Slavonic sounds, and so there is great vagueness in the application of Greek letters to these. This problem may be seen very early in the alphabet names reported by Banduri (in 1711), and supposed to date from the thirteenth century, though possibly Banduri's source was not in Greek. Here we find, for example, /š/ written as σ/ς ('σαα', 'vaς' for 'ša', 'našь'), /ž/ as ζ ('ηζε', 'ζηβητ' for 'iže', 'živite'), and /č/ as τζ ('τζερβη' for 'čьrvь'). The same problem is still around in the eighteenth century, for example the *Lexicon Tetraglosson* (see Kristophson 1974) also has τζ for /č/ (e.g. 'ρετζε' for 'reče'). This document does use the iota subscript on **a** to represent [ă] (e.g. 'τᾳρβα' for 'd(ă)rva'). Otherwise there appears to be no evidence of any innovation in the form or application of Greek letters to Macedonian.

6 Arabic

The use of Arabic script for a Slavonic language applies to Belorussia, starting from the arrival of the Tatars in the thirteenth-fourteenth centuries, and to Bosnia from the arrival of the Turks in the fourteenth century. For Belorussia we have many examples of such material from the mid-sixteenth to the twentieth centuries (see the collection by Антонович/Antonovič 1968). For Bosnia there exists an Islamic literature for a similar period,

including some forty printed books from the nineteenth and twentieth centuries (Janković 1989: 36). Ивић/Ivić (1986: 155) dates such a Bosnian literature from the early seventeenth to the nineteenth centuries. In theory, Arabic – at least in its classical form – would have presented similar problems to Latinica or Greek, since Arabic is also short on symbols for palatals and affricates; further, it would certainly have had problems also in the representation of the vowels; however, it did have extra symbols for 'emphatic' and fricative consonants, as well as the possibility of varying the number of dots on a given symbol. These resources had already been utilized in the application of Arabic script to Turkish, and it was this adapted script which was used by the Slavs, with some innovations of their own; thus, for example – according to the list given by Супрун, Калюта/Suprun, Kaljuta for Belorussian usage (1981: 11–14), the letter *ẓa* was used for /dz/, *tha* for /s'/, *ṭa* for /t/ and *ta* for /t'/; /c/ is written with *ṣad* but with three dots added, and the same three dots replace the one of *jim* for /č/, the one of *za* for /ž/ and the one of *ba* for /p/. For the vowels: /o/ is represented by the superscript **a** (') written over *waw*, that is, the classical Arabic diphthong [au] collapsed; however, so also is /u/ written thus (classical Arabic superscript **u** (') followed by *waw*), that is the superscript difference appears not to have been retained; /e/ is written by what appears to be the superscript **a** (') enlarged to letter size.

A similar description of late Bosnian Arabic script is given by Janković (1989: 32–3), indicating the additional influence of the Vuk/Gaj system of diacritics. Thus, the same three-dot versions as above are used for /č/, /ž/ and /p/; however, a two-dot version of *jim* is used for /c/ and a version of the latter with a circumflex instead of the dots for /č'/; /dž'/ and /dž/ are both simply the original (one-dot) *jim*; a circumflex is also used over *nun* for /n'/ and a *háček* over *lam* for /l'/(!); similarly /o/ is represented by *waw* with a circumflex and /u/ by *waw* with a *háček*; /e/ is simply the classical letter *ha*.

7 Hebrew

Some scant evidence exists of the use of Hebrew script by Belorussian Jews, but it was apparently never used for more than personal writing or place names (Wexler 1973: 47).

8 Orthographic principles (Cyrillic and Latinica)

Apart from odd proposals of a phonetic principle, the clear track in most areas has always been to work on the 'morphological' principle in orthography, that is to retain visible morphological relations in spite of surface phonetic facts. While this would seem to be natural, or at least a 'good idea', for languages like Russian (and English), where surface phonetic

changes can be major, theorists do not always see it that way: thus Belorussian has gone for a phonetic representation of the vowels, though, inconsistently, not of the consonants. Thus, while the common lexeme /solod/ 'malt' is realized (in the nominative singular) in both languages with a final [t], Russian spells it солод/solod, Belorussian солад/solad, both on the principle that in all the other forms of this word the /d/ is realized as [d], and so the root morpheme is {solod}. Languages with simpler surface realizations of the morphology do not have the same problems; thus, for example, Serbo-Croat, and to a lesser extent Ukrainian, can boast of having a 'phonetic spelling system' – they simply do not have the linguistic problems of languages like Russian; hence the above word in Ukrainian is realized as [solod], and thus may be spelt 'phonetically' солод/solod; the cognate Serbo-Croat /slad/ is realized as [slad] and spelt slad/слад.

Apart from Serbo-Croat and Ukrainian, all Slavonic languages devoice final obstruents, but none indicate this in the spelling. Ukrainian is odd in respect also of the (regressive) assimilation of obstruents in groups: it does not allow devoicing to occur, but does allow voicing, while all the other languages, including Serbo-Croat, allow assimilation of both sorts. Most languages are inconsistent in their attitude towards the spelling of this assimilation; however, Serbo-Croat maintains its 'phonetic' spelling by writing the surface value of the obstruent, for example sladak/сладак masculine 'sweet', feminine slatka/слатка.

As for the spelling of the vowels, only Belorussian attempts to spell them phonetically, as indicated above (солад/solad = [sȯlat]); thus those other languages which show surface changes related to stress position do not represent these in the spelling, for example Russian (солод/solod = [sȯlət]); in many cases the vowel changes occur only in certain variants of the standard languages, for example the raising of Bulgarian pre-tonic /o/ to [u], so that no single spelling would be phonetic for the whole standard, not to mention the non-standard variations.

All of the above applies equally to the Cyrillic and Latinica users; in fact, no Latinica user follows the Belorussian model for vowels, that is all follow the morphological principle, but the Latinica version of Serbo-Croat, of course, follows the Cyrillic one in spelling the voice assimilation (*sladak*, but *slatka*).

Finally, in many systems extra diacritics may be used in particular circumstances, for example to indicate suprasegmental features, or to aid disambiguation; examples of the first are:

1 stress position may be indicated by an acute accent in East Slav and Bulgarian dictionaries and textbooks for pedagogical purposes;
2 length and tone may be indicated in Serbo-Croat (with four accent marks) and Slovene's conservative variant (with three accent marks);

3 length (and also quality) of some vowels (*e, o*) may be indicated in the modern variant of Slovene (with the same three accent marks);
4 length may be indicated in post-tonic syllables in Serbo-Croat (by a macron).

Examples of the disambiguation function are:

1 Russian usually marks the word *čto* with an acute to indicate the object pronoun ('what') as opposed to the conjunction ('that');
2 Bulgarian marks the pronoun *i* ('to her') with a grave accent as opposed to the conjunction *i* ('and');
3 Russian has available the letter ë – phonetically representing stressed [o] after a soft consonant – which is used both in pedagogical functions and for disambiguation, as for example to distinguish всё/vsë (/vs'o/ 'all', neuter singular) from все/vse (/vs'e/ 'all' plural).

The regular orthographic use of diacritics to indicate suprasegmental features is limited to Czech and Slovak, which use the acute to indicate vowel length.

9 Summary

The modern situation is thus:

Glagolitic is no longer used anywhere (but is still recognized in Croatian Church usage, even if not decipherable by many – for example, Zagreb Cathedral bears a prominent Glagolitic inscription of recent provenance on its interior back wall);
Cyrillic is used throughout the East Slavonic area (Russian, Ukrainian, Belorussian); in the south in Bulgaria, Macedonia and Serbia, and also in Bosnia and Montenegro, thus in the whole of the east and south of the Balkan Slavonic region;
Latinica is used throughout the West Slavonic area (Czech, Slovak, Polish, Sorbian); in the south in Slovenia, Croatia, Bosnia and Dalmatia; also in Serbia, where it exists alongside Cyrillic, though the latter is dominant in most parts.

For details on the modern systems, see the relevant chapters of this book.

TRANSCRIPTION AND TRANSLITERATION

1 Definitions

Most of what is included under both of the terms *transcription* and *transliteration* can be generalized as 'conversion of scripts' (see Wellisch 1978), in the sense that almost all this activity centres around the business of making one alphabet (or, more generally, a 'script', or writing system) accessible to users of a different system. The one area which is not subsumable under this general description is the representation of *phonetic* elements (sounds) in a written form for use not only by users of different phonetic systems, but also by learners of native systems.

Of the two terms, the second – transliteration – is the easier to relate to its application: as its name suggests, it involves the transference (conversion) of *letters*. It is concerned with the conversion of one writing system – and specifically an alphabetic one – to another, and is not necessarily concerned with sounds at all. Transcription, on the other hand, in spite of the root 'script', is applied to the representation of either a writing system *or* a sound system in a written form which will allow users of other systems in particular to appreciate the *sounds* of the source system. While it may use the symbols of a particular target language's orthography, this is not essential, and any symbolic system may be used, depending only on the requirements of the target audience.

2 Target audiences

Who needs conversion of a writing system? It is the huge variety of answers to this question that produces the likewise huge variety of conversion systems. A few of the major customers are: librarians, who want to provide access via their catalogues to material written in scripts other than their own; editors of journals or newspapers, which have to refer at least to names of people, places and other journals in their discussion of foreign sources and events; and linguists, who want to describe languages at various levels of detail to other linguists who may not need (or desire) to access the source script.

While it is probably true to say that linguists are the easiest target audience, in that they are usually by definition well-informed in the business of scripts and sounds, even they may be subdivided into a variety of groupings related to the reasons why they want a conversion. There may, for example, be linguists reading this book who are unfamiliar with any Slavonic language, let alone the Cyrillic script, and whose needs in terms of conversion are limited to the minimum which will give them access to information about higher levels of the languages, say the morphology or syntax. For these – assuming they are users of a Latinica script – a basic

transliteration will suffice, and they may not need any comment at all on those languages which already use a Latinica script. Others will be interested in the sound systems themselves, and will want to know much more than can be got from simple transliteration: at the 'top' end, they will want a highly sophisticated transcription system which will give them minute details about the sounds of the source language.

Other customers for conversion are usually much more diverse in their needs, as also in their degree of sophistication in the general use of scripts. For example, librarians need not be concerned at all about the sounds underlying the foreign symbols, as their primary concern is that material can be accessed in the alphabetic order of their native system, and then its issue and return controlled by staff who do not know the foreign system. If staff or users need to say aloud names or titles, the roughest of approximations is quite satisfactory. On the other hand, radio commentators have to say aloud such foreign names all the time, and so are forced to make decisions about how best to approach this, that is how far they should go in imitating the source pronunciation, if indeed they have any idea of this. The practical situation which clearly occurs typically is that the foreign forms are first written, that is transliterated, by journalists or editors, using whatever degree of approximation suits their minimal written needs, and then the radio journalist, say a news reader, is required to read the form aloud, guided at best by formal recommendation (based, it is hoped, on informed sources), and at worst by uninformed common journalistic usage.

This last case, of the radio announcer's problem, is, of course, not limited to the situation of transliterated source scripts: that of other systems based on the same script as the target is just as great a problem, indeed potentially greater, as the expectation that the system is different is diminished by the apparent familiarity of the script. Two aspects of this are: (1) the different values of the same symbol in different systems, for example the sound value of **ch** in English, French and German; and (2) the use of diacritic signs which tend not to be transferred, and whose function cannot thus be carried over, for example the French vowel accents, whose omission does not produce serious problems, or the Czech consonant hooks, whose omission does produce potentially serious ones.

Ultimately, so long as we are dealing with proper names, as is usually the case in journalism, any approximation will do, so long as it is said often enough to be identifiable in a given form. However, much of this approximation is unsatisfactory even to non-linguists, if only in that it causes much puzzlement and confusion. English-speaking non-Russianists are frequently puzzled by the fact that a name spelt in transliteration with an **e** can be pronounced with an [o], as in 'Gorbach*ev*'; the problem is, of course, that transliteration and imitation have been mixed, in that the Russian letter is indeed the one transliterated as **e**, while the sound is indeed [o]; writing **e** is fine until the name has to be said, just as saying [o] is fine until

the written form interferes. This particular problem is reconcilable only by an editorial decision to temper the transliteration with phonetic information – in other words, to go rather for (or towards) a transcription; in the case in question the name would be better written 'Gorbachov' (as in the usual German version 'Gorbatschow').

Just as much confusion is caused by Latinica source names like the Czech names 'Dvorak' (for Dvořák) or 'Mecir' (for Mečíř) – in which the Czech letter ř represents a vibrant palatal fricative sounding somewhat like the sequence 'rzh', and moreover devoiced ('rsh') in final position, as in the second name. As they stand, without their diacritic marks, they can be used as satisfactory visual references to the people in question, but as soon as someone tries to say them (and in the case of a top tennis player, like the latter, this will happen rather often), problems arise: the uninformed, that is the vast majority, are highly puzzled by the addition of a spurious [ž] after r in the first name, and by the pronunciation of c as [č] and r as [š] in the second – always assuming that this is what is indeed said. A Polish name like 'Walesa' (for 'Wałęsa', where w represents [v], ł [w] and ę a front nasal vowel) presents similar problems.

Since the idea of transliteration is perceived as impossible within versions of the same script, we are left with the paradoxical situation that transliterated names, say from Cyrillic, are more likely to be pronounced accurately than names in other Latinica alphabets. It would be nice if everyone moved to a transcription system for all names, whether for print or sound-media use. However, against this stands the inertial force of tradition: once a particular form of a proper name has been used often or long enough, it becomes 'the' form of that name, and in all approximative uses will resist any attempt to make it 'more accurate'. Thus Russian names ending in [-skij] are happily written '-sky', this not conforming to any formal transliteration scheme (though often used in the style of individual journals), but providing a reasonable approximation; 'Dvorak' is generally pronounced (more or less) correctly with the 'extra' [ž], in spite of no clue to this being offered in its form; by now 'Gorbachev' is such an accepted form. A parallel situation is the use of established foreign versions of place names which are not transliterations of the modern native names, but usually represent old variants: for example, Moscow, Vienna, Copenhagen, China, or French 'Londres'. Only occasionally does tradition change, as in the recent conversion of Peking to Beijing in English usage.

Before passing to looking at the major transcription and transliteration systems which are (or have been) in use, we must first sort out a bit of terminological confusion: Wellisch, amongst others, refers to the script as 'Roman', and the alphabet as 'Latin'; it is probably useful to use 'alphabet' in the language-specific sense and 'script' in the general, and I adhere to this usage in general, but I do not regard it as crucial, since the context always makes the meaning clear; however, to many, 'Roman' – or at least

'roman' – refers to a type-style, opposed to 'bold', 'italic' and so on; furthermore, the term 'romanization', as applied to situations like Chinese and Japanese, is used in the sense of an alternative usable script for natives, rather than a transliteration for specific purposes or for foreigners. I therefore use only 'Latinica' for both purposes, and never the term 'Roman' in any form.

The assumption in this book is that we are concerned with conversion *into Latinica.* All other scripts have exactly the same problems, simply viewed from a different perspective; they all perform conversion into their own systems. Earlier in this chapter we considered examples like the representation of Slavonic sounds in the Greek or Arabic alphabets. The common problem is that languages have scripts which, at least by tradition, if not by consistency, represent their own sound system, which is different from others' sound systems; and since even native scripts are only rarely accurate representers of the sounds, it is hardly surprising that trying to apply a different script to a given sound system should create problems.

3 Transcription systems

As transcription is concerned with the transmission of the *sounds* of the source language, the two basic variants needed are: one for specialist linguists, and one for 'approximations' usable for the like of radio journalists.

In the first case, there are two subdivisions: one for the phonologist and one for the phonetician. The phonologist is interested primarily in the phonemes of the source language, and not in the fine details of pronunciation; for this purpose we need what is called a 'phonemic' or 'broad' transcription; thus, for example, 'Gorbachev' could be transcribed as /gorbačóv/. The phonetician is in addition interested in the details, including the effect of stress, and requires a 'phonetic' or 'narrow' transcription; the same name could be transcribed as [gərbʌʧ̍ɔ̍f]. (In each case I say 'could' be transcribed, because the given versions are only two of many possible transcription systems or degrees of detail, some of which we will take up below.)

The ideal 'narrow' transcription will allow the (ideal) phonetician to produce a native-like version of any sound or sequence of sounds from any language. The only condition is that the describer and the interpreter are using exactly the same system, that is the interpreter must know the exact (intended) value of every symbol used; for this reason it is necessary to establish widely recognized and accepted systems, indeed preferably only one such system, as transcription would then know no boundaries. This last situation is unfortunately not quite the case, although at least amongst professional phoneticians there is now one such system, known as the International Phonetic Alphabet, whose symbols are used in the above

phonetic transcription. Amongst local language pedagogues, especially those using scripts other than Latinica, there is much less acceptance of this alphabet, and more use of variants based on the native alphabet; thus descriptions of Russian intended for internal use prefer a Cyrillic-based system. It is notable, however, that recourse to IPA symbols is frequent in the transcription of details.

For most linguistic purposes, a broad transcription is sufficient, and this presents fewer problems, as the number of symbols required is much smaller. Again, most systems are based on the native alphabet, which means that Latinica is certainly the most popular, but not the only one; again, Russian sources use Cyrillic exclusively for phonemic descriptions. Further, there are variants in the Latinica usage also, depending on the typical sorts of phonemes in particular language groups; it has been common for descriptions of western European languages to use IPA symbols also for broad transcription, thus 'Gorbachev' could be transcribed phonemically as /gorbaʧov/ (the symbol /ʧ/ being the IPA one for the voiceless palatal affricate). Amongst Slavists, however, the tradition has for some time been to use, rather, the symbols existing in the Czech alphabet for this purpose. Thus, for example, palatal consonants are transcribed by the Czech 'hook' letters (š, ž, č); softness of consonants is indicated by an acute (s′, z′). This system has proved very efficient in describing all the Slavonic languages, including their older stages, and also the reconstructed forms of Proto-Slavonic.

One further advantage of this system is that it is easily used also for general transliteration purposes, as it does not use the 'odd' shapes of many IPA symbols, but only regular Latinica letters with diacritics. This is a major advantage, as it reduces the overall number of systems which any reader has to deal with. We will treat this usage below.

Finally, a word about the 'approximative' transcription: while the users of such transcriptions are normally not linguists, and need have no real interest in a 'good' pronunciation of a name, they must still produce something recognizable, and if they consider themselves professionals, they owe it to their public to make a serious stab at correctness. In this case, it is true that using any symbols which do not occur in the native alphabet is a waste of time: no non-linguist can be expected to know the significance of **č** or **ʧ**. Provided the sound concerned exists in the target language, there is little problem: here, English would use (its normal) **ch**; but with a foreign sound, there must inevitably be problems: some of these are insurmountable in the sense that given sounds will simply not be imitated, for example Czech /ř/; others are amenable to analogical transcription, like the writing of /ž/ in English as **zh** or of /x/ as **kh**, etc.

At this point the border between transcription and transliteration becomes blurred: such users are both transliterators and transcribers, and the importance of the sound media has made the latter function much more

important than previously. It is no longer good enough for a name to be simply transliterated: someone will soon have to say it aloud, and will get it badly wrong if transcription has not been considered, as in the case of the last vowel of 'Gorbachev', or the initial sound of 'Evtushenko' – an effective transliteration/transcription will allow for the pronunciation and produce rather 'Gorbachov' and 'Yevtushenko'. The marking of stress position is also highly desirable.

4 Transliteration systems

I have argued that some of the above instances of transliteration are negative in that what is required in such cases is at least some consideration of sound values, and thus of transcription. For 'pure' transliteration to be justified, the sound must be completely unimportant and irrelevant. Only then can one be 'scientific', that is consistent, in the activity. Consistency here means that a given letter (syllable, ideogram, etc.) of one language is *always* represented by the same *distinct* letter etc. of the target language, without any regard to the behaviour of the underlying sound. In this way 'reversibility' is assured. Potential users of such systems are such as librarians and cartographers; amongst the users of libraries and maps are those who are familiar with the source script, and they want to be able to reconstruct precisely the source form; in any case it must still be assumed that the transliterated forms are for *reading only*; as soon as the question of *speaking* them arises, as in, say, an oral request for a foreign title, or in the teaching of geography, some guidance on the source sounds is desirable to say the least. This is, of course, a linguist's view of the world, and while I would expect to find plenty of support among the readers of this book, it is certainly not a view held by the non-linguistically minded majority.

And so to the systems in use. Not surprisingly, it is the librarians who have done the most work in this area, with the major libraries of the world devising such scientific systems as mentioned above, in the first place for their own direct users, but indirectly also for smaller libraries and many others who adopt their systems. The two major English-language systems are those of the Library of Congress and the British (Museum) Library; both of these are based on the use of the Latinica letters without diacritics, though the Library of Congress system does use the ligature (‿) and breve (˘). Both of these systems date from early this century (1905 and 1917 respectively), and have thus built up a tradition as difficult to replace as a standard orthography.

For the great majority of letters there is no problem, since the basic sounds are the same in all the languages involved and both systems use the 'simple' letters 'a', 'b', 'd', etc. From those that do cause problems, examples of these two systems applied to Russian Cyrillic are as follows:

Russian	*Library of Congress*	*British (Museum) Library*
ж	zh	zh
х	kh	kh
ц	ts	ts
щ	shch	shch
й	i	i
ы	y	ui
ъ	″	(*omit*)
ь	′	′
я	ia	ya

Clearly, the two are very close, but they are nevertheless different enough to cause some confusion, and certainly some aggravation amongst library staff when the wrong one is used! A practical problem with the Library of Congress system is that the diacritics must be done by hand, undesirable in the mechanical age. However, this has been seen as a problem with any foreign diacritics, and in this age of computerised typography need no longer be insurmountable.

Alongside these two English-based systems there have existed others in various European countries, the most important and 'scientific' being the German *Preussische Instruktionen*; the system used in the French *Bibliothèque Nationale* catalogue is somewhat less 'scientific' (Wellisch 1978: 250). Most other countries likewise have used more or less local versions. The problems inherent in such diversity led to the desire for some uniformity, and to the production by the International Organization for Standardization (ISO) of its first transliteration standard (ISO/R9, published in 1954), which was, in fact, for transliteration of Cyrillic. It opted for the Czech/Croat-style use of diacritics, and the above sample letters were transcribed thus:

Russian	*ISO*
ж	ž
х	h
ц	c
щ	šč
й	j
ы	y
ъ	″
ь	′
я	ja

Adherents of the two English-based systems were not impressed, and stuck to their own systems (compare the 1958 British Standards Institution (BSI) system, virtually the same as the British Library, except 'ȳ' for **ы** and 'ŭ' for **ъ**; and the 1976 American Standards Association – now American National Standards Institute (ANSI) – system). The second edition of ISO/R9, in 1968, conceded defeat in acknowledging the British/American system as an acceptable alternative, but the 'double standard' involved

clearly defeated the whole purpose, and was reported by Wellisch (1978: 258) as about to be rejected in the third edition ('probably in 1977'), which was to revert to the system of the first edition. (In fact, there was apparently considerable disputation over this edition, and it was finally published only in 1986, as the first 'proper' edition (called ISO 9), as opposed to the previous 'recommendations'.)

Wellisch gives a useful comparative table of the many systems (1978: 260–2), as part of the history of Cyrillic transliteration (pp. 256–64). This includes the draft changes of the ISO third edition.

All of these systems were set up in principle by and for cataloguers. But others were using them too, and in our context most importantly the editors of scholarly journals and books. Inevitably, the confusion of the cataloguers has continued here, and the same variety of systems is apparent throughout the Slavist academic world. Moreover, not only are there different systems in different countries, but even internally in different disciplines. For English-language journals, the BSI or ANSI systems are normal, but now only for non-linguistic material: Slavist linguists have, not surprisingly, opted for the ISO system, based as it is on the Czech model, which thus conveniently serves both purposes of phonemic transcription and transliteration.

One final point must be made about the non-ISO systems: they are more than simply transliteration systems since they, somewhat surprisingly, take into consideration the phonemic system of the particular language involved; for example, in the BL and LC catalogues, the letter **х** is transliterated as 'kh' for all but Serbo-Croat and Macedonian, when it is 'h'; similarly for these two languages the Croat (= Czech) Latinica letters are used for **ж**, **ц**, **ч**, **ш** (see the *British Library Reader Guide no. 3 – Transliteration of Cyrillic*). This usage is based, presumably, on the parallel use of Latinica in Serbo-Croat, expanded to include Macedonian as another Yugoslav alphabet. Likewise, **щ** is transliterated as 'shch' for Russian, but 'sht' for Bulgarian. The example of **г** transliterated as 'g' for most, but as 'h' for Ukrainian and Belorussian, may be defended by the previous existence of the second letters **ґ** (Ukrainian) and **г'** (Belorussian), inconsistently, and no longer, used for [g] as opposed to [h] (Ukrainian) or [ɣ] (Belorussian).

Overall, one would think that a strict transliteration system should be 'language-neutral'. However, as linguists, we can have no objection to the inclusion of language-specific information of this sort, especially as it draws transcription and transliteration closer together. It is this principle which lies behind the system used in this book.

5 The system used in this book

This then being a book on Slavonic linguistics, for the joint purposes of phonemic transcription and graphic transliteration we use the system which

has become standard in Slavonic linguistics. This is based on the ISO system, with some relevant phonemic information being allowed as in the above cataloguing practices. These cases are few, and are as follows (the justification given in parentheses):

Letter	*Language*	*Transcription/transliteration*
г	Bg. Mac. OCS Ru. SCr.	g
	Bel. Ukr.	h (phonemic/phonetic)
и	Bg. Mac. OCS Ru. SCr.	i
	Ukr.	y (with i for **i**)
х	Bel. Bg. OCS Ru. Ukr.	x
	Mac. SCr.	h (to match SCr. usage in Latinica)
щ	Ru. Ukr.	šč
	Bg. (OCS Ш)	št (phonemic/phonetic)
ъ	Ru.	″
	Bg.	ă (phonemic/phonetic)
	OCS	ъ (phonemic)
ь	Bel. Ru. Ukr.	′
	OCS	ь (phonemic)

Further, certain local situations will call for some variation between the transcription and transliteration details (for example, in the rendering of soft consonants in Russian). These will be treated in the relevant chapters. Note too that we use the acute rather than the apostrophe for **ь** to avoid confusion.

Finally, where *phonetic* transcription is necessary, the IPA symbols are used, written in square brackets; phonemic transcription is written in oblique brackets. The full transliteration table is given on pp. xii–xiii.

References

Charanis, P. (1961) 'The Armenians in the Byzantine Empire', *Byzantinoslavica* 22: 196–240.

Cubberley, P. (1982) 'Glagolitic's Armenian connection', *Wiener Slawistischer Almanach* 9: 291–304.

—— (1984) 'The formation of Cyrillic and problems of the early Slav alphabets', *Wiener Slawistischer Almanach* 14: 283–302.

—— (1988) 'On the origin and development of the Slavonic letter-names', *Australian Slavonic and East European Studies* 2, 1: 29–54.

de Bray, R.G.A. (1980) *Guide to the Slavonic Languages*, 3 vols, 3rd edn, Columbus, Oh.: Slavica.

Gilyarevsky, R. and Grivnin, V. (1970) *Languages Identification Guide*, Moscow: Nauka.

Hamm, J. (1974) *Staroslavenska gramatika*, Zagreb: Školska knjiga.

Hoffer Edle von Sulmthal, A., and Margaritoff, M. (1989) *За Кирил и Методий*, София: Изд. на Отечествения фронт (Text in Bulgarian, German, Russian and English; English pp. 255–324, 332–3).

Istrin, V. (1988) *1100 Years of the Slavonic Alphabet*, Moscow: Nauka (translation of Истрин 1963).

Janković, S. (1989) 'Ortografsko usavršavanje naše arabice u štampanim

tekstovima. Uticaj ideja Vuka Karadžića', *Prilozi za orijentalnu filologiju* 38: 31–8.
Jurančič, J. (1977) 'Problem cirilice na severozahodu južnoslovanskega etničnega ozemlja', in F. Jakopin (ed.) *Slovansko jezikoslovje. Nahtigalov zbornik*, Ljubljana: Univerza, 135–52.
Kolarič, R. (1970) 'Slovenci in glagolica', *Зборник за филологију и лингвистику* 13, 1: 77–83.
Koneski, B. (1983) *Historical Phonology of Macedonian*, Heidelberg: Carl Winter.
Kristophson, J. (ed.) (1974) 'Das Lexikon Tetraglosson des Daniil Moschopolitis', *Zeitschrift für Balkanologie* 10, 1.
Kuna, H. (1977) 'Neke grafijske osobine bosanskih srjednjovekovnih kodeksa u odnosu prema staroslavenskoj glagoljskoj grafijskoj tradiciji', in F. Jakopin (ed.) *Slovansko jezikoslovje. Nahtigalov zbornik*, Ljubljana: Univerza, 153–66.
Mareš, F. (1971) 'Hlaholice na Moravě a v Čechach', *Slovo* 21: 133–99.
Mayo, P. (1977) 'The alphabet and orthography of Byelorussian in the twentieth century', *The Journal of Byelorussian Studies*, 4, 1: 28–47.
Schenker, A., Stankiewicz, E. (eds.) (1980) *The Slavic Literary Languages: Formation and Development*, New Haven, Conn.: Yale Concilium on International and Area Studies.
Schröpfer, J. (1968) *Hussens Traktat 'Orthographia Bohemica'*, Wiesbaden: Harrassowitz.
Stone, G., Worth, D. (eds.) (1985) *The Formation of the Slavonic Literary Languages*, Columbus, Oh.: Slavica.
Štefanić, V. (1963) *Tisuću i sto godina od moravske misije Ćirila i Metodija* (*Slovo*, Posebni otisak iz br. 13), Zagreb: Staroslavenski institut.
Vajs, J. (1937) 'Hlaholica na Rusi – Novgorodské sgrafity', *Byzantinoslavica* 7: 184–8.
Wellisch, H. (1978) *The Conversion of Scripts – its Nature, History, and Utilization*, New York: Wiley.
Wexler, P. (1973) 'Jewish, Tatar and Karaite communal dialects and their importance for Byelorussian historical linguistics', *The Journal of Byelorussian Studies* 3, 1: 41–54.
Zor, J. (1977) 'Glagolica na Slovenskem', in *Slovansko jezikoslovje. Nahtigalov zbornik*, Ljubljana: Univerza, 483–95.

Антонович, А.К. (1968) *Белорусские тексты, писанные арабским письмом, и их графико-орфографическая система*, Вильнюс.
Велчева, Б.А. (1989) *Глаголица на Балканите до началото на XIII век*, София: БАН (Автореферат).
Ивић, П. (1986) *Српски народ и његов језик*, Београд: Српска књижевна задруга.
Истрин, В.А. (1963) *1100 лет славянской азбуки*, Москва: ИАН.
Карский, Е.Ф, (1928) *Славянская кирилловская палеография*, Ленинград (Reprint 1979, Москва: Наука).
Медынцева, А.А. (1969) 'Глаголические надписи из Софии новгородской', *Советская археология*, 1.
Мирчев, К. (1963) *Историческа граматика на българския език*, София: Наука и изкуство.
Младеновић, А. (1989) 'Око Лукијановог (Вуковог) слова «Ђ»', *Српска Академија Наука и Уметности. Балканолошки институт.* Посебна издања 38: 155–8.
Супрун, А.Е., Калюта, А.М. (1981) *Введение в славянскую филологию*, Минск: Выш. школа.