

3 Concepts, Theories, Methods

3.0 Introduction: Preliminaries on Language Contact

The discussion in the first two chapters has already introduced a number of key concepts pertaining to language contact and language comparison. In this chapter, we develop our ideas further in order to facilitate the presentation in the subsequent chapters where the substance of Balkan language contact is discussed. Moreover, some consideration of general methodology is necessary, to elucidate the bases upon which our analysis of the Balkan sprachbund data rest.

Any model of the language contact situation in the Balkans has to reckon with a highly multilingual area, as the material in Chapters 1 and 2 makes abundantly clear. This truism is evident not just from the many languages co-existing in the region and the obvious signs of contact (loanwords, etc.), but also from first-hand accounts by travelers in earlier periods. One telling description is that given by Brailsford 1906: 85–86 of Skopje (Trk *Üsküb*) in the Ottoman period:

With luck, a traveller in Macedonia may hear six distinct languages and four allied dialects spoken in the same market-place. If it is a northern centre, for example Uskub, the peasant women who handle raw wool and hawk their own homespun, may use two Slavonic dialects, which vary slightly but still appreciably. . . . The merchants in the booths belong to three races. There are some Greeks, but these probably are immigrants from Salonica and the South, who came with the railway. Jostling with them are the Jews, who also come from Salonica and speak their Spanish jargon. More numerous are the Vlachs or Wallachians . . . and they speak indifferently their own Latin (Roumanian) dialect, or the Slav of the peasants with whom they have done business for centuries, or the culture-language of the Greeks which has come to them through their schools and their Churches. . . .

In short, one may distinguish in the Babel two Slav and two Albanian dialects, Vlach, Greek, Turkish, Hebrew-Spanish, and Romany. . . . There is no *lingua franca*. Hard necessity imposes some knowledge of elementary Turkish upon the majority of the male population, but they rarely speak it with ease; the peasants seldom understand more than a few words of it, and it has never become a medium of intercourse between different Christian races. Greek is more serviceable as a polite or commercial language, but its day is over, and save in the south, it is only the older generation of Slavs which can speak it. French is now taking its place. Slav (and particularly the Bulgarian dialect [in modern terms, Macedonian – VAF/BDJ]) is the one language with which no native of the northern and central districts can dispense.¹

1 While these quotations from Brailsford render a sense of the impressions of an early twentieth-century British traveler, it must be remembered that Brailsford's account suffers from his background and limitations. Aside from his use of *race* to refer to what today we would call *ethnicity*, he was,

Moreover, such a mix of languages is consistent with what Capidan 1940 describes for Aromanian villages and what Mirčev 1952: 124 describes for eastern Bulgaria; multilingualism has been the norm rather than the exception as far as the Balkans have been concerned. Still, it is important to keep in mind that there was at least a gender-based differential to bilingualism in the Balkans. Récatas 1934 discusses the effects of gender-differentiated roles in Balkan and especially Aromanian bilingualism, in that men were the ones who moved around outside the sphere of the home and who thus interacted with speakers of other languages; women generally stayed close to home and thus were less likely, he found, to be particularly conversant in other languages (cf. Daniel et al. 2021: 152 with regard to similar gender-related patterns in Daghestan). Women, for Récatas, therefore, were linguistically conservative in this regard. A glimpse into their attitudes about bilingualism can perhaps be discerned in Récatas's observation (p. 30) that some Aromanian mothers, presumably with some degree of passive/receptive bilingualism (see §3.2.1.1), would make fun of children for using *ohi* (his transcription for Greek *ὄχι*, the word for 'no'), and would say back to the child *psohi*, with a mocking "prefix," or *ohiksufu*, with a mocking "suffix."² Stojanov 1952: 218 reports similar conditions about bilingualism for eastern Bulgarian; Grannes 1996: 4 (= 1988: 225) sums it up as follows: "Often bilingualism would be more common among men than women: '[. . .] when it was necessary to conceal something from their wives, the men spoke Turkish' (Stojanov 1952, 218)." And, Bunis 1999: 91 notes a similar situation involving Judezmo-speaking men, as opposed to the Jewish women, of Ottoman Thessaloniki with regard to the learning of Turkish: "Females, much more confined to the domestic scene and the narrower Jewish world of friends and relatives, were much less exposed to Turkish and other Gentile languages"; by contrast, "Jewish males spent much of their day outside the home conducting business, which brought them into frequent contact with spoken Turkish."

It is against a backdrop of this sort that the most fruitful discussions can take place of the methods needed to analyze the situation and of the constructs one must assume in order to make those methods work.

3.1 Languages versus Speakers

It is customary in discussions of the diffusion of linguistic features from one language to another to talk in terms of "language contact," as seen in numerous

apparently, unaware (or unwilling to inform his audience) that Turkish was a major medium of communication among all urban dwellers in Ottoman Europe. See Chapter 4, footnote 364.

- 2 Koufougiorgou 2004 offers an updated view of gender-based use of Aromanian in Aminciu (Grk Métsovo, in Epirus) and interestingly, in her broad-based survey, women report less use of Aromanian than men, perhaps because women are at the forefront of the shift to Greek as the prestigious and socially and economically advantageous language. Nonetheless, she reports something analogous to the mocking noted by Récatas, with regard to more traditionally oriented women making fun of women who dress "European," suggesting that women play a role in perpetuating traditional aspects of Aromanian life, even if the language is receding. The linguistic mocking of Récatas's era seems to be at odds with the Balkan proverbs valuing bilingualism given in our Introduction (p. 1), but may reflect a more local dynamic.

linguistic textbooks, from a classic work such as Weinreich 1953 (widely cited also, as throughout here, in its sixth printing by Mouton from 1968) to more recent efforts such as Thomason 2001 or Winford 2003 or Matras 2009 (cf. also Aikhenvald 2007a). While we recognize this usage as well-established in the literature, we agree with Weinreich 1968: 71 that it is merely a convenient shorthand for what is really involved in the influence of one language on another, namely *speaker* contact. That is, just as “language” can be said to be constituted in and through acts of communication,³ so, too, “language contact” takes place in the interaction between or among speakers of different languages; the languages themselves, the abstract entities, do not engage in contact per se but do so only through the medium of actual speakers engaging in contact via face-to-face communicative interactions.⁴ The qualifier “virtually” is needed here for two reasons.

First, it can be argued that in the case of a bilingual speaker,⁵ one, though not necessarily the only, locus of contact must really be, as Weinreich 1968: 71 has observed, within the mind of that individual, what has been referred to (e.g., by Pavlenko 2014) as the “bilingual mind.”⁶ In such a mind, interference and transfer may take place between the different internalized systems, a first (native) language (L1) and a second (usually – but not always – not native or not as native) language (L2), represented therein.⁷ Second, exception must also be made for borrowings that occur through the medium of written sources, especially so-called learned borrowings, where no contact between actual speakers is needed; thus a speaker of English who uses a Latin phrase was never conversing with a living native speaker of Latin, but rather is “interacting” with a writer through the printed page either directly or through the mediation of some other nonnative speaker. The same would hold for “contact” through more recent types of noninteractional (and perhaps less learned) media such as audio- and video-recordings, movies, tapes, the Internet, etc.⁸ Nonetheless, most “language contact” is “speaker contact.”

3 Indeed, all definitions of “language” we have ever seen start with some reference to “communication”; the OED (s.v.), for instance, in its definition, states that language is “the system of spoken or written communication”

4 This point is made also in Joseph 1992f, where a speaker-oriented approach to the study of language change in general is advocated. Cf. also Matras 2012, who makes a similar point in his study of Domari.

5 We refer here and elsewhere to “bilingualism” and “bilingual speech” as a matter of convenience, fully recognizing that in many instances more than two languages are involved. Thus our “bilingual” is to be understood as “bi-, tri-, quadri-, or multilingual.”

6 This point is made by several scholars from time to time, and by citing it here, we are giving it our endorsement; see also Lehiste 1988: 28, Masica 1992: 10, and Giannini & Scaglione 2002: 152.

7 We are deliberately being noncommittal here about the nature of the “different internalized systems.” This notion could refer to two different, separate, and distinct grammars, or to a set of options for the realization of a given feature or construction, and no doubt to other models as well. However interesting the question might be in general, for our purposes, nothing crucial hinges on this, as long as there is *some* representation in the speaker’s mind, in *some* form, of the parts of the two discrete languages (this issue is addressed from different perspectives in many of the chapters in Bhatia & Ritchie 2004, especially Ijalba et al. 2004, Meisel 2004, and Costa 2004; see also Mufwene 2008). In fact, data from the Balkans (and elsewhere, e.g., Auer & Muhamedova 2005 on Kazakh-Russian contact) support the argument that no single model can account for the variety of observed phenomena. See also the discussion below in §3.2.1.6 on codeswitching and Friedman 2008a.

8 See footnote 14 below for some concrete examples.

Moreover, while it is customary to talk in terms of *language contact*, especially in the Balkans, to a certain extent it would be more accurate to talk about *dialect contact*. Both inter-dialectal contact and inter-language contact are inter-speaker contact. The only difference is that in dialect contact the relevant systems are more similar to one another (typically) than in language contact.⁹ Thus the distinction between language contact and dialect contact is arguably more a matter of degree than of kind.¹⁰ As far as the Balkans are concerned, no matter how this distinction (or lack thereof) might be judged in absolute terms, keeping it in mind is important. The main focus in Balkan linguistics has always been at the language level (albeit originally with dialectal data), but, as noted in §2.5.2 and discussed in §3.2.2.9, over-reliance on standard languages can obscure rather than enlighten. Moreover, inter-language contact will always necessarily involve dialects of each language, since speakers are always speakers of (at least) one dialect, and even a standard language is a dialect – a privileged one – of the language viewed as an abstract totality.¹¹

The need to focus on speakers as opposed to languages when talking about contact becomes evident also when one considers the nature of borrowings in the Balkans. That is, on the one hand, there are many loanwords involving aspects of contact between different cultures that conceivably could be transmitted via non-interactive media such as the written language. On the other hand, there are also numerous loans of a more “intimate” nature, involving expressive items and common conversational elements. These are all discussed in detail in §4.3.

What is telling about this latter group is that, given their conversational nature and thus the fact that they are precisely the sorts of words that do not occur readily in higher registers of usage, including written language, their diffusion around the Balkans constitutes proof that the different peoples of the Balkans were in fact talking to one another and were in direct contact. Such markers are the “glue” of conversation and can be considered grammatical elements at the discourse level; they are both colloquial and abstract. We would expect, therefore, that they can

9 This means that the task confronting a speaker of one linguistic system (“A”) needing to make any sense of a different system (“B”) will typically be harder in the case of cross-language contact as opposed to cross-dialectal contact. Moreover, we recognize that there can be quite different cultural associations and other semiotic processes (Gal & Irvine 2019) with the respective systems in both cross-language and cross-dialectal situations.

10 Emeneau 1980: 127 argues that “the same process produces the same result whether between dialects of a language, languages of a family, or languages of several families.” This is similar to the argument made by Mufwene 2000 concerning creole genesis (cf. also Enfield 2005 on transmission versus diffusion and Gal & Irvine 2019 on the semiosis of both language and dialect usage). Some linguists see dialect borrowing as a different process altogether, however; see Nakhleh et al. 2005, for instance.

11 See Alexander 1983, 1984–1985 on inter-dialectal contact in the Balkans, and Trudgill 1986 for general discussion. Mæhlum 1992 offers an interesting case of dialect mixing and contact involving Spitsbergen Norwegian. Le Page (e.g., Le Page 1992, continuing themes from his 1985 book with Andrée Tabouret-Keller) challenges absolute notions of concepts such as language and contact. We can note that the competition between language and religion as sources of identity in the Balkans raises similar issues. Max Weinreich’s famous quotation and its background can be found at <https://en.wikipedia.org/wiki/A_language_is_a_dialect_with_an_army_and_navy>.

spread only through actual use in (conversational) encounters between speakers. As Civ'jan 1965: 9 puts it, "the stimuli for the formation of the [Balkan sprachbund] are to be sought . . . in the daily necessity of the use of language as a means of communication," and these are precisely the sorts of words that find a home in the articulation of "daily necessities."¹² If language contact took place in the abstract, away from the everyday colloquially based interactions that real speakers engage in,¹³ and in particular if contact in the Balkans in the relevant historical period were more like the casual "contact" that has led in the contemporary period to the spread of technical terminology from English (e.g., involving computers, drugs, and sports) into hundreds of languages, it would be hard to understand how these highly discourse-bound elements, along with so many others like them, could have diffused so widely and so readily in the Balkans.¹⁴

Moreover, as the later chapters make clear, it is essential to utilize notions like referentiality and argument structure in accounting for various pan-Balkan phenomena. Object reduplication (see §7.5.1) is a case in point. Such notions are necessarily tied to interaction between speakers, who do not trade morphological categories like one trades valued objects such as fur pelts and beeswax. Rather, there is generally communicative intent behind such contact-related phenomena. We might say such interaction and the resulting loans reflect "human-oriented" contact rather than "object-oriented" contact.

Nonetheless, throughout this work, as a practical matter, we refer to *language contact*, as it is accepted phraseology in the field. In keeping with the importance we place on the social setting and the language-external factors that were vital the forming of the Balkan sprachbund, our reference to language contact is always to be understood in terms of the speakers involved and the dialects they use(d) on a regular basis.

12 The translation is ours. Civ'jan's original formulation in Russian is: "стимулы для образования БЯС следует искать . . . в повседневной необходимости использования языка как средства коммуникации." Grannes 1996: 4 (= 1988: 225) takes a similar position with regard specifically to Turkish influence on Bulgarian, saying the influence "is the result of direct contact between speakers of the two languages in everyday life."

13 It may not be the case that anyone would want to claim that no speaker contact was involved in language contact, but the cautionary note is important to sound, since it is easy to be misled by metaphors in discussing linguistic phenomena, and in practice, this distinction is not always observed (as discussed in Joseph 1992f).

14 Admittedly, American colloquial usage has spread through the medium of popular songs and films, and especially the Internet, and also when speakers who have been abroad bring expressions back with them, as with Macedonian *okej* 'OK,' *orajt* 'all right' used in a play set in the early twentieth century (Krlje 1967). Still, such borrowings can be located in speaker-to-speaker contact, in the sense that a speaker returning from abroad or using the Internet is generally the medium through which such foreign forms enter, and indeed, in a case like this, English is really only indirectly the source of the usage. Nonetheless, some borrowing of elements that are conversational and intimate can take place without direct contact through nonspeaker-centered interactions (e.g., through songs). The Internet has provided an entirely new source of language change and contact, but for the most part it is a medium for the spread of English. However, it is also a medium for dialect contact. Such research vis-à-vis the Balkans must be the topic of a separate study.

3.2 Basics on Methodology

Examining languages in contact with one another is a form of comparative and historical linguistics, in that the elements of one language typically are compared with those in another and judgments are made as to the historical source of comparable features in order to determine which of several possible causes is responsible. In the sections that follow, various aspects of comparative linguistics and related notions pertaining to causation, to dialects, and to methodology in general, all of which are highly relevant to subsequent chapters, are introduced and discussed, with two main foci: the recognized mechanisms of contact-induced change and the methods utilized for assessing contact situations.

3.2.1 Mechanisms Relevant in Contact-Induced Change

On the one hand, the social conditions under which contact among speakers takes place are critical, i.e., whether the contact is a result of (1) conquest and imposition of a new language on a population, (2) a forced dislocation of speakers of different languages suddenly thrust together and needing to communicate with one another, (3) a relatively peaceful coexistence involving considerable interaction for economic or other reasons, (4) a voluntary or involuntary shift in language use in the face of social or political pressures such as potential economic benefits, avoiding prison, and so on. All of these are social settings for language contact that are found in various venues around the world, and in microcosm in the Balkans, where scenario (1) describes the arrival of Greek in the second millennium BCE and Latin in the early Hellenistic period (c. second century BCE), (2) describes the arrival of Slavic in the sixth and seventh centuries CE and Judezmo at the end of the fifteenth century, (3) describes the later situation of Balkan Romance, Albanian, Greek, and Balkan Slavic as well as the position of Romani (which probably arrived at the beginning of the second millennium CE), and (4) describes shifts to Turkish, especially among urban populations, during the Ottoman period, but also possibly shifts from Balkan Romance to Balkan Slavic in parts of Macedonia and from Balkan Slavic to Balkan Romance in parts of Romania in pre-Ottoman times (cf. Gołąb 1976) as well as shifts that have occurred in certain twentieth-century nation-states.¹⁵ We can speak of *substratum* and *superstratum* to describe the participants in and the direction of shift when language shift occurs, and *adstratum* to describe influence without shift. On the other hand, as emphasized above, language contact as discussed here requires speaker contact, and thus involves individual speakers using and/or hearing some language other than their first (native) language or acquiring more than one language natively from childhood (cf. Friedman 2000d). Thus, language contact can be considered a matter of how individual speakers experience and perform language.

¹⁵ This list is merely illustrative, not exhaustive.

3.2.1.1 Bilingualism

The most basic notion, touched on above with reference to the so-called bilingual mind, is that language contact implies some degree of bilingualism on the part of speakers, that is, some degree of access to, if not actual command of, more than one language. Definitions of what it means to be bilingual and what bilingualism entails vary considerably.¹⁶ Haugen 1953: 7 defines it in a fairly broad way as the ability to “produce *complete meaningful utterances* in the other language” (italics in the original), and Lehiste 1988: 1 has a similarly broad definition (referring to the ability “to produce grammatical sentences in more than one language”). Roman Jakobson, on the other hand, is said to have held, informally at least, to a stricter criterion for bilingualism involving the ability to lecture spontaneously in more than one language on a topic with which one is familiar, and Bloomfield 1933, as Edwards 2004: 8 notes, “observed that bilingualism resulted from the addition of a perfectly learned foreign language to one’s own, undiminished native language . . . though admitting that the definition of ‘perfection’ was a relative one.” The more recent trend is in the direction of broad views of how to define bilinguals, with the characterization of Butler & Hakuta 2004: 115 being representative and quite useful for present purposes: “individuals or groups of people who obtain communicative skills, with various degrees of proficiency, in oral and/or written forms, in order to interact with speakers of one or more languages in a given society.” This broad view can be explicitly broadened even further; some definitions above refer just to production of another language, i.e., *productive* bilingualism, but there is also *passive*, or *receptive*, bilingualism (Edwards 1994; Romaine 1995; Basham & Fathman 2008) whereby a speaker of one language can understand another language to varying degrees but cannot (readily) or does not produce utterances in that other language. Such situations occur not only in everyday public interactions but also within families involving linguistically mixed marriages (Friedman, field notes). Unless otherwise specified, “bilingualism” as used herein includes both productive and passive/receptive bilingualism.

Diebold 1961 takes what may be thought of as a realistic, and similarly useful, view and refers to “incipient bilingualism,” a situation where speakers of one language have fragmentary knowledge of another language, but presumably know enough to recognize elements in the other language. Since language-learning takes place over time and one necessarily has to start with learning a little and then adding to that, presumably “incipient bilingualism” is the stepping-stone every user of multiple languages starts with on the way to greater knowledge

16 Edwards 2004 has very useful discussion of different approaches to defining what bilingualism is; see also Grosjean 2004. Note too terms used by Monica Heller to refer to certain results of governmental “policies practiced in the name of bilingualism,” as Moschonas 2019: 206 puts it, namely “double monolingualism” (Heller 1999: 97), “two separate monolingualisms stuck together” (Heller 2000: 14), or two “parallel monolingualisms” (Heller 1999: 139). The situations Heller refers to do not find realization historically in the Balkans but their existence elsewhere makes it clear that it is important to recognize that the Balkans enjoyed a particular kind of bilingualism and that not all bilingual situations are the same.

of and skill in a second language. At the same time, though, it need not be the case that all incipient bilinguals develop their second-language skills to the same degree, so that some may end up with so-called fossilized L2 abilities (see Han & Odlin 2006). Once one recognizes different degrees of ability developmentally in a second language for a given individual, one has to recognize as well that on a societal level, in a given bilingual community, there can be speakers of differing abilities in the other language, some fluent or nearly so, some conversant and functional, some fossilized in a rudimentary stage of learning, and so on. Such differential ability may result from differential access to the other language and differential opportunity to hear, learn, and use it,¹⁷ or it may be generationally based, if the language of the home is different from the language of the streets or of schooling. Factors involving individual motivation for and individual abilities in learning a second language cannot be discounted either.¹⁸ Differential abilities in L2 across a community can figure in the spread of some nonnative features into the native language; some concrete examples are given in §5.3.

Here the notion of so-called *imperfect learning* is also relevant. At issue is the question of the extent to which nonfluent L2 speakers, as opposed to fluent L2 speakers, were influential. Added to this is the fact that well into the twentieth century, the standard versions of the Balkan languages were contested and in flux. Greek was in a state of diglossia until 1976, Albanian had two standards until 1968–1972, Macedonian was not recognized as a standard until 1944, Romani and Aromanian have country-specific standards, but these are still contested at various levels, Turkish has been through a series of reforms since the introduction of the Latin alphabet in 1928. Bulgarian, Romanian, and the former Serbo-Croatian had what passed for standard languages by the end of the nineteenth century, but all have been subject to various vicissitudes in the twentieth. Our point here is, in part, to question the very notion of imperfection in language learning as it relates to language contact. To be sure, different speakers had different abilities, but in a world where most people were illiterate, and there were no modern nation-state standard languages, a variety of mechanisms can be invoked, and written evidence, while highly useful, can itself be an imperfect guide. Thus, for example, Pulevski 1873, 1875 are extremely valuable indications of colloquial structures in the Balkan languages (Dombrowski 2015; Hamiti 2005; Sonnenhauser 2020). At the same time, however, the fact that Pulevski was striving to write, rather than simply speak, means that some of his language did not actually represent spoken language but rather his idea of what was appropriate for writing (e.g., final *-d* for etymological *-t* in Macedonian). Another

17 See §3.0 above, for instance, on a gender-based difference in bilingualism in various Balkan contexts, where access is a crucial factor.

18 The ontogeny-phylogeny model (OPM) of Major 2001 recognizes differences of this sort, in part through its recognition of a “monitoring” function on the part of the language learner and of the possibility of “considerable individual differences in the amount of monitoring” (p. 117).

point to keep in mind is that, as Labov 1964, 2001 has shown, adolescence is the real locus of language change, which raises the possibilities of inter-linguistic modifications based on youth identity, as is found elsewhere in the world, rather than imperfect acquisition (and see also footnote 42).

Another dimension to differential abilities for a given language in a community in which there is more than one language spoken is that bilingualism need not always be mutual. That is, there can be a different directionality in who learns which language in a multilingual community, and it can be the case that some or all speakers of language A will need to learn language B but only a few speakers of B, if any, will need to learn language A. This is typically the case when speakers of language A are in a distinct minority and are socially isolated or socio-economically disadvantaged. As pointed out in Friedman 2000d, in the Balkans, for instance, very few non-Romani speakers learned Romani and very few non-Judezmo speakers learned Judezmo; rather, Roms and Jews learned the dominant language(s) of their locales and communication with others in their communities took place via the medium of that dominant language. Some consequences of such social segregation can be seen in aspects of Romani phonology, for instance, as discussed in §5.3.

These considerations also raise the important question of who counts as a speaker of a language in the first place. Especially when generational considerations are taken into account, one has to wonder what level of command of a given language X qualifies one as a “speaker of X.” Take the case of an Arvanitika speaker living in Greece who, due to the pervasiveness of Greek in all domains of daily life now (and the history of repression of Arvanitika), is a few generations removed from the period in which the language was used exclusively on a day-to-day basis and is more familiar with Greek than with Arvanitika; does such a speaker “count” as an Arvanitika speaker affected by Greek? Is the Arvanitika of such a speaker still Arvanitika in some sense, or is the speaker really a speaker of Greek with just a smattering of imperfectly learned and controlled Arvanitika available to him/her? Tsitsipis 1981, 1998 uses the somewhat chilling¹⁹ designation “terminal speaker” for such individuals, and terms like “semi-speaker” are also to be found in the literature.²⁰

Bilingualism, like language in general, is thus to be viewed both on an individual, psychological, basis and on a group, societal, and community, basis. For the most part, in discussing the effects in speakers of knowledge of one language on another language, most researchers take a speaker’s first (native) language as the baseline and measure effects of learning other languages on it. In some instances, however, it is necessary to think of the first language in a societal rather than individual sense, so that a third-generation terminal speaker of Arvanitika, with a limited command of Arvanitika even if

19 This is the (very apt) descriptor used by Mackridge 1992: 117 in his account of Tsitsipis’s terminology.

20 See Dorian 1977, 1980 on the notion of “semi-speaker,” with specific reference to Scots Gaelic but with broader relevance.

exposed to it early in life in the home, can shed light on the extent to which a heritage language for an ethnic group can be affected by another language. Instances of just this type come up again and again in the various chapters of this book as we survey the varying results of Balkan bilingualism.

3.2.1.2 Agentivity in Language Contact Situations and Bilingualism: A First Look

In bilingual contact situations in which the contact results in innovative elements entering one or more of the languages, an important parameter to take into consideration is agency, that is, who the agent of change is, who brings the innovation into the affected language. There are several dimensions to agentivity, and these are explored here and in some of the following sections. An approach to contact-induced change that has focused attention on agentivity is that of van Coetsem, as seen in van Coetsem 1988, 2000, and it has been particularly influential in its insistence on recognizing different paths of agency. In a sense, focusing on agentivity is a way of recognizing the role of the individual in contact-induced change. Thus we outline the relevant aspects of individual agent-driven contact effects here, drawing some terminological comparisons with other notions often employed in the language contact literature, but also using this discussion to recognize other kinds of agentivity to be discussed in later sections.

The most important distinction in van Coetsem's approach is between recipient language agentivity and source language agentivity. In the former case, a speaker of one language, in receiving material from another language, is the agent of change, taking a form from outside his/her linguistic system and using it. This corresponds to the traditional notion of borrowing, and the borrowing speaker is indeed the agent, actively taking the foreign element and adopting it (perhaps in an altered form) as his/her own, thereby adding material to the native system. Hock 1991: chapter 14, *passim*, calls this "adoption," with reference to lexical borrowing, and Johanson 1992, 2002: 3 uses the same term (in German: *Entlehnung*), and for him, adoption involves a process in which speakers, agentively, "copy" material from another language. In the latter case, also referred to as "imposition" by van Coetsem – note that Johanson employs the same term (in German: *Unterschiebung*) – a speaker of one language, in using a different language (the "target" language, the "second" – or secondarily acquired – language) imposes structures from his/her linguistic system onto another. This corresponds to the traditional notion of substratum influence,²¹ and thus to "transfer" in the literature on Second Language Acquisition, and "interference through shift" in the (also) highly influential model of Thomason & Kaufman 1988, in that a speaker filters the use of the other language through his/her own system, thereby effecting change in that other language and producing an altered

21 And somewhat more, in Johanson's case, as is discussed in §3.2.1.3.

form of it. In such a case, the second/other language is the recipient, and the agent is the language (actually, speakers of the language) that is (are) the source of the altering force.

Thus borrowing of material by speakers of language A from language B is recipient language (A) agentivity, and the imposition of patterns from language A onto language B still has A as the agent, but the agentivity is from A as source language. In the first instance, B is unaffected and not even necessarily directly involved (cf. §3.1 on borrowing without face-to-face communicative interactions), while in the second instance, B is the recipient language, the one affected by the imposition through the creation of a new form of B, namely B as filtered through the substratum of an A speaker's system.

While one can question whether a single speaker's accented pronunciation of a target language constitutes an alteration to that language in the abstract,²² clearly if there are enough speakers engaging in the same sort of alteration in a concentrated geographic area, then a new variety of the target language can arise.²³ This process underlies, for instance, the development of Indian English as a new dialect of English, or, for that matter, most of the "new" Englishes in the world,²⁴ and it can be seen as operative in the phenomenon of Yugoslav Standard Turkish, which differed from the Standard Turkish of Turkey (see Friedman 1982c, 2002a).

By way of summarizing the different labels attached to the individual-based contact effects discussed here, we offer the following table (Table 3.1) ("—" signals the absence of an agreed-upon term):

Table 3.1 *Mechanisms/processes for contact-induced change – some comparisons*

<i>Traditional approach van Coetsem</i>	borrowing recipient language agentivity	substratum influence source language
<i>Thomason & Kaufman</i>	borrowing	agentivity (imposition) interference through shift
<i>Second language acquisition literature</i>	—	transfer
<i>Johanson</i>	adoption	imposition

22 Major's OPM (see footnote 18) focuses on the individual and thus offers a way of recognizing the substrate effect, i.e., source language agentivity, through the stages of language learning and monitoring the learner goes through.

23 It is important to recognize that "engaging in the same sort of alteration" need not involve diffusion from speaker to speaker but rather can be independent on the part of each speaker; presumably with the same systemic pressures acting as a filter for each speaker, the same filtering (substratum) effects are likely to arise in individual speakers independently of one another.

24 See E. Schneider 2003 on the various new Englishes in the world, with ample references.

3.2.1.3 More on Agentivity: Reverse Interference

The notion of agentivity is indeed useful in characterizing certain kinds of contact-induced change, but van Coetsem's model is not the whole story. There are further agentivity-related effects not covered by van Coetsem's approach that appear nonetheless to play an important role in contact-induced change.

One that comes up repeatedly in Chapter 5 regarding Balkan phonological contact is what we here call *reverse interference*, whereby a second language has an effect on a speaker's first (native) language.²⁵ An example is the use of Greek-like interdental fricatives in some Macedonian dialects of northern Greece, Macedonian elsewhere being a language without such fricatives (see §5.4.4.3). Similarly, the Albanian of Mandrica has lost interdental fricatives under the influence of Bulgarian (Sokolova 1983). This effect was envisioned by Seliščev 1925: 43 when he wrote "*Tel ou tel fait de la langue acquise a exercé une influence sur la langue maternelle et réciproquement*" ('This or that fact of the acquired language exerted an influence on the mother tongue and vice versa'). In his classic work, Weinreich 1968: 109 also recognizes such a possibility: "language shift does not exclude linguistic influence in the opposite direction."²⁶ Moreover, this effect has been demonstrated experimentally, under controlled laboratory conditions, in numerous studies by Flege and others, e.g., Phillips 1982; Flege 1987; Major 1992; Hussein 1994; Yeni-Komshian et al. 2000; Flege 1998; Guion 2003; and Bond et al. 2006; cf. also Laeufer 1990, 1996. Recapitulating his work, Flege 2007 states that while "interference was usually regarded as the influence of prior learning on subsequent learning, that is, as an effect of the L1 on the L2 ... Grosjean 1982 suggested ... that interference is bi-directional, and that the dominant language will influence the nondominant language to a greater extent than the reverse." This work collectively has confirmed that knowing and using a second language, especially in an intense and sustained way, can have a subtle but real effect on the production of native-language sounds. For instance, native speakers of English living in Brazil showed Voice Onset Time (VOT) in their English stops that approximated but did not necessarily match exactly Portuguese VOT (Major 1992) and native Arabic speakers immersed for many years in an English-dominant environment had English-like vowel lengthening before voiced stops in their Arabic (Hussein 1994).²⁷

25 The term *reverse interference* was used in linguistics at least as far back as Phillips 1982, who defines it as "interference from a second language (L2) to a first language (L1)" experienced by bilinguals. It was also used in Felton 1990, Laeufer 1990, and Hussein 1994. The term is not uncommon in works – mostly on phonology – in the late 1990s and into the 2000s.

26 Liu et al. 1992 refer to this phenomenon as "backward transfer," which they define (p. 455) as "a process whereby strategies that are appropriate for L2 feed back on L1"; Chen 2006 also uses this terminology, as does Winford 2013, while Matras 2009: 224 writes of "phonemes diffusing 'backwards' to substitute inherited phonemes in selected words."

27 Such "reverse" effects also show up in speech perception; perhaps the first experimental study on reverse interference, Caramazza et al. 1973, found that French–English bilinguals perceived stops in both languages in terms of the VOT of their second language, English. Interestingly, these speakers did not show reverse production influence from English into their French, suggesting that this overall phenomenon is perhaps more complex than might first be supposed.

Effects such as these are found in more naturalistic contact conditions. A. Schmidt 1985: 229 found for Dyirbal that there were phonological effects from secondarily acquired English onto the first language, Dyirbal, of various speakers in the younger generation (ages 15–39), several of whom were quite fluent speakers.²⁸ She notes that the “English fricative /f/ occurs in Y[oung people’s]D[yirbal] pidgin-type pronouns; wefela ‘we plural’, youfela ‘you plural’” and comments further that “there are signs of phonetic insecurity as the rhotic contrast weakens in non-minimal pairs; YD speakers waver in the realization of the T[raditional]D[yirbal] phonemes /rɾ/ and /r/. ”²⁹ She also documents this effect in syntax (p. 230), with English word order in YD speakers (“the exceptionally-free TD word order is rigidified in YD as an A-V-O [agent-verb-object] pattern as in English”) and morphosyntax, with English prepositional marking for peripheral cases among less proficient YD speakers, instead of TD case-suffixes.

This sort of effect could in principle just be considered, in Thomason & Kaufman’s (1988) terms, a type of “borrowing,” since the altered VOTs of Major’s speakers and the YD use of [f] and of English word order all represent features that a speaker has gotten from another language. However, especially when a compromise outcome is involved rather than the incorporation of a foreign feature per se, it seems to be a different sort of effect from others seen so far. Although something enters a language from outside that language system, the effect is (typically) far subtler than in cases of borrowing in the more usual sense. Whereas borrowing can be quite casual in nature (a lexeme here, a structure there), this effect of reverse interference involves the filtering of one language through the system of another. In that way, it is just like a substratum/interference-through-shift/imposition/transfer except that the languages have altered, actually reversed, their role as to source and filter in that the original second language becomes the filter through which the original native language “flows” in production and perception.

Still, despite the considerable number of studies showing this effect, there is no agreed upon term for it in the literature.³⁰ Johanson 2002: 3 describes his “imposition” in simplest terms by saying that for a group of speakers of language A, “[language] B influences their A,” suggesting that the reverse effect we have in mind here is possible, and some of the situations he describes apply here. Nonetheless, “imposition” as a term has been put to other uses in the literature, as §3.4.1.2 makes clear.

28 The youngest group of speakers (ages 5–15) knew virtually no Dyirbal, and were able to say some words but not to construct any sentences; to the extent that they can say anything in the language, “their pronunciation of Dyirbal words is distorted by the English sound system” (p. 194). Such speakers are L1 speakers of English, so their minimal Dyirbal just shows regular interference (that is, substratum/imposition) effects from their L1 onto their attempts with an L2.

29 Here, /rɾ/ is an alveolar trill whereas /r/ is an alveolar retroflex.

30 Numerous colleagues have confirmed for us that *reverse interference* is not a widespread standard term, although, as seen in footnote 25, it has been around for decades. See also footnote 26 on terminology.

We thus endorse the designation *reverse interference*.³¹ One language “interferes” with another, but in the reverse of the widely recognized substrate effect of L1 influencing L2. Factors such as the age at which the L2 is learned, the extent of use of the L2, and so on, play a role, but the effect is real. Examples from the Balkans are presented in Chapter 5 (see, e.g., §§5.4.3.6, 5.4.4.2, 5.4.4.9).

As for agentivity, reverse interference is not agent-driven in the sense of borrowing or imposition, or even copying, but rather simply happens to speakers without their actively doing anything other than attempting to speak a second/foreign/other language with increasingly greater fluency. There is agentivity, in that some activity takes place, but the effect, it would seem, is not a conscious one.

Somewhat more consciously or overtly agentive, but with a similar outcome, is what Matras 2009: 225, 228 calls “authentication,” whereby bilingual speakers who are borrowing words from another language, with sounds that are not found in their speech otherwise, “are able to identify and produce those sounds.” This control that he, quite properly, envisions speakers having over the way they render a foreign word in their usage means that they can essentially choose to let the donor language affect their other language’s phonology or not as they borrow a word. An example Matras gives from the Balkans is “Macedonian Turkish /ts/ with Macedonian loans,” /ts/ being alien to Turkish but borrowed as such nonetheless, thus affecting the local Turkish of North Macedonia and other WRT areas.

3.2.1.4 More on Agentivity: Accommodation and Selection of Variants

A second agentivity-related effect is accommodation. When a speaker of one language alters his/her speech patterns and structures in the direction of what another speaker is doing, or in the direction of what s/he believes the other speaker should be doing, we can say that one speaker is accommodating linguistically to the other. Work by Giles 1973 alone and in collaboration with others (e.g., Giles & Johnson 1987) developing a “Speech Accommodation Theory,” since renamed “Communication Accommodation Theory” (see Sachdev & Giles 2004), offers a useful account of the dynamics of shifts in language use by an individual (as agent) in response to the social setting of interacting with a particular interlocutor and that interlocutor’s own speech patterns. Trudgill 1986 (cf. especially chapter 1) develops accommodation into a key driving force in determining resolutions to contact among dialects.

The reasons for accommodation may be social in nature, having to do with gaining approval from the interlocutor or others or establishing a particular social identity, or they may be more functional in nature, motivated by a desire for – or a belief in the ability to create – greater communicative effectiveness. Major 2001:

31 Flege himself (p.c., September 5, 2007) has said that “‘reverse interference’ seems like a reasonable term to use,” though he wonders if just plain “interference” would then be suitable for the opposite (and more usual) effect.

78 summarizes the motivation behind accommodation thus: “One’s desire to accommodate ... can depend on the prestige (or perceived prestige) of a particular accent or sound, one’s attitude toward the target language and culture, and the perceived power gained or not by acquiring the language.” We can add that other factors such as the desire to create social solidarity can also influence accommodation. This holds for interactions among speakers of different dialects of the same language as well as among speakers of different languages. But whatever the reason, the result can be convergence of one speaker’s linguistic patterns to those of another.³²

In some instances, the accommodation might consist of adopting phraseology or pronunciations of another speaker wholesale, while in other instances, the result of accommodation may be a reduction in the percentage of use of a particular variable or even compromise usage of one type or another; when a particular form is involved, we have what Trudgill 1986: 62, writing about cases of dialect contact, has called an “interdialect form.”³³ To some extent, accommodation at the dialect level is the first step in koineization (for a useful summary on which see Kerswill 2002), the development of a new speech form based on elements from mutually intelligible but nonidentical dialects through selection and simplification leading to leveling of differences. At the language level, accommodation can lead to convergence.³⁴

Two other approaches worthy of mention in this context as involving accommodation to some extent are feature selection, as in Mufwene 2001a, 2008, and the epidemiological model of Enfield 2003, both of which take approaches from the natural sciences and adapt them to linguistics. In the case of feature selection, communicative efficacy can be said to function like evolutionary pressure in such a way that when a language has a number of features at its disposal for achieving a communicative goal, that feature which proves most effective in a particular environment is more likely to be selected and transmitted. Thus, in an environment of language contact, if each language has several features in competition for expressing a grammatical concept and both happen to share a particular feature, that feature is more likely to be selected and developed. A concrete example is the expression of Balkan futurity by means of a particle derived from etymological ‘want’ (see §6.2.4.1.1). Such a device is attested as an independent development in each of the Balkan languages for which we have adequate documentary evidence, but the impetus for choosing precisely this grammatical means among those that

32 Sachdev & Giles 2004: 356 notes that the processes of interaction can also lead to a divergence, to “an accentuation of language (and cultural) differences” (so that there really is no “accommodation” in the strict sense, only a reaction to the potential for accommodation, cf. also Labov 2007). We are concerned here more with processes leading to convergence, but note as a possible case of divergence, or at least differentiation, the situation with Romani phonology described in §5.3.

33 See §5.2 and especially footnote 14, for a striking case of accommodation bringing interlocutors down to the same level of proficiency, based on Young People’s Dyrbal as studied by A. Schmidt 1985 (and see §3.2.1.3 above for more on Dyrbal).

34 Note that we are using *convergence* here to refer to a result of processes and mechanisms activated in contact situations, and not, *pace*, e.g., Myers-Scotton 2002: 164, as a process or mechanism in and of itself.

were competing, e.g., in Old Church Slavonic, where etymological ‘have,’ ‘be,’ ‘begin,’ and plain perfective presents were all also available features, can be seen as feature selection triggered by adaptation to the surrounding linguistic environment.

Enfield’s epidemiological model looks at feature spreads in mainland Southeast Asian languages (see also Enfield 2017). This approach is related to approaches that treat language as a kind of virus passed from person to person and capable of an altered shape. While this approach takes a different view from that of feature selection, nonetheless the two models are not incompatible. An example from the Balkans of compromise is the form taken by the negation marker of Tsakonian Greek in the first half of the twentieth century (Pernot 1934).³⁵ The inherited negator in Tsakonian is [o] (from Ancient Greek οὐ) but due to influence from Standard Modern Greek with its negator δεν, Pernot reports that some Tsakonian speakers developed a compromise form [ðon], adopting the standard form but blending it with the native Tsakonian vowel; in a sense, then, such a form accommodates to standard Greek usage but maintains a native aspect as well (and see §3.2.1.8), having the consonantal shell of the standard form with the vocalism of Tsakonian. While this particular example involves only a single lexical item, it is one with an important grammatical function, and it illustrates the nature of linguistic compromise. Clearly other areas of grammar and lexical form can be the targets of such a process of accommodation (see §3.2.1.5 for examples from Russenorsk and from Haitian and St. Lucian creoles).

An example from the Balkans of this process at work across different languages comes from the Greek of Southern Albania. As reported on by C. Brown & Joseph 2017, one finds words there which for the most part have the expected Greek segments but which substitute a sound from an Albanian word with a similar meaning and similar phonological shape.³⁶ For instance, for ‘mechanic, engineer’, one can hear [mexanikós] with -e- in the first syllable, even though Greek more generally, and properly from an etymological standpoint, should have -i- (cf. Standard Greek μηχανικός). The parallel Albanian word *mekanik*, with its [e], seems to have impinged on the Greek form and led to the accommodated compromise form μεχανικός ([mexanikós]), with features from both Albanian (the -e-) and Greek (the suffix -os).³⁷

35 Tsakonian has been described as a “Greek dialect” (e.g., Browning 1983: 124), but it is divergent enough from standard Modern Greek to warrant being called a separate language, and is implicitly recognized as such in Greece and generally by linguists. All Modern Greek dialects (except Tsakonian) are descended from the Attic Koine, while Tsakonian is the only clear direct descendant of the ancient Doric dialect. We use the qualification *clear* because the Greek dialects of southern Italy preserve some Doricisms and have some independent developments, although overall they participated in the Byzantine world and thus reflect the same Attic Koine that is the basis of Modern Greek. On the significant participation of Tsakonian in that world, see Kisilier 2021.

36 C. Brown & Joseph refer to such forms as “hybrids” and offer examples of not just phonological hybrids but also morphological and semantic, and even paralinguistic, hybrids; further fieldwork by Joseph in 2017 has revealed a number of additional examples, e.g., [dáfni] (with [d]) ‘laurel,’ versus expected δάφνη with [ð], based on Albanian *dafinë*. See §4.3.10.2.3 for more discussion of semantic hybrids and §§6.1.4.1 and 6.2.1.1.1 on instances of what can be treated as morphological hybridization. Ndoci & Joseph 2024 explore the ideological implications and foundations of such hybrid forms.

37 It is difficult to claim that [mexanikós] is a borrowing from Albanian *mekanik*, as the stress in the Albanian is on the -ni- syllable, whereas the Greek stress is on the -ko- syllable. As it happens, *mekanik* might be an older borrowing from Greek, from a time when the vowel in the initial syllable

At the language level, too, the case of the spread of the loss of the infinitive in the Balkan languages and its replacement by finite complements, as discussed by Joseph 1983a: chapter 7, can be cited. In that work, the author outlines a scenario for this development (for more details about which see §7.7.2.1) in terms of accommodation in a multilingual contact situation. He suggests that the use of finite forms represented in part an attempt by speakers to produce utterances in a form that would aid communication and make their utterances more readily parsable, and thus intelligible, to their listeners. Other factors can be identified, to be sure, including ambiguities of form that enhanced the possibility of interpreting an infinitive as a finite form, but in this scenario, accommodation was the key, and it was carried out through the selection of particular variants already available in the language that served the communicative function at hand.

Moreover, accommodation can occur on both sides of a communicative exchange. Ferguson 1971 identified what he called “foreigner talk” as a special case of what can happen between speakers of different languages with a pressing need to communicate. Alteration of the language occurs, with speakers of both codes accommodating one to the other, and this in turn affects the target language in two ways. It creates both an L2 variety used by nonnative speakers and an altered variety of L1 used for talking to those nonnative speakers. Implicit in such a construct is a notion of what the “bare bones” are that are needed for communication, so this then leads to a consideration of “simplification” in contact situations.

3.2.1.5 Simplification and Pidginization

Simplification is viewed by some observers as the primary mechanism of contact-induced change, although empirical data show that complexification can also occur. When new material enters a language, it can be at odds with – and thus add a complication to – the already-existing system of the language.³⁸ The introduction of finite complementation in Turkish from Persian, in the form of subordinate clauses introduced by *ki*, increased complexity in the grammar of Turkish by extending the range of possible ways of expressing subordination.³⁹ However, it is hard to see complexification as a usual strategy or mechanism that speakers might employ when faced with an outside linguistic influence; one does not

(AGrk < η >) had not yet raised from the ancient pronunciation [ē] to the later [i]; alternatively, the vocalism of *mekanik* may have been mediated through written sources or other nonancient sources, or perhaps most straightforwardly, it simply reflects a borrowing from Italian *meccanico*. See footnote 36, and in §3.2.1.6 footnote 50, for other examples of such hybrids involving the blending in of phonological elements. On the other hand, given that Albanian stress is, as is normal in the language, on the last syllable of the stem, here [nik], it is possible that Albanian *mekanik* was simply Hellenized with the suffix *-os* and the Albanian stress rule applied to stress [kos]. A similar explanation is not available for other hybrids, such as [dáfni] in footnote 36.

38 As Nichols 1992: 193 says: “It can be concluded that contact among languages fosters complexity, or, put differently, diversity among neighbouring languages fosters complexity in each of the languages.” See Thomason 1997 for various empirical examples.

39 See §3.2.1.7 for some discussion of the theoretical import of Turkish *ki*.

ordinarily set out to make a system more complicated and look for material to accomplish that end, particularly if communication is a goal.⁴⁰

On the other hand, the act of simplifying structures encountered in another language as they are brought into one's own language can quite plausibly be seen as a mechanism that speakers employ in a contact situation, and simplification can occur in contact-induced change. The most extreme cases of linguistic simplification known involve multi-language input in a contact situation. We have in mind here the creation of so-called pidgins, structurally impoverished contact "codes" whose emergence in a contact situation is necessitated by circumstances of communicative exigencies. But pidgins are functionally quite restricted, and Balkan language contact does not seem to have led to anything like pidgins emerging anywhere in the region. Thus, more germane for the Balkan situation is simplification as a means by which a learner of a L2 deals with the challenges that material from that language poses.⁴¹ That is, similar processes of simplification may be at work in pidginization and in other types of language contact but the result does not have to be a restricted pidgin. The role of simplification would be especially pronounced in the kind of second-language learning that is most relevant to the Balkans, namely untutored, informal learning, what is sometimes referred to as "natural second language acquisition" (Winford 2003: 209), with learners learning spontaneously, and developing communicative abilities adequate for their circumstances, without overt instruction guiding their learning.

As for resulting simplification, it should be clear that someone with an imperfect, even if functional, command of an L2 is not going to be able to control the native speaker's full range of structures, forms, and even sounds in the target language (TL). The difference between the native's TL and the nonnative's L2 can generally be characterized as a reduction or simplification of aspects of the TL when it becomes an L2. Indeed, as Winford 2003: 217 observes, following Meisel 1977, "learner versions of the TL are typically reduced in lexicon and structure, and consequently in communicative power." Moreover, it is commonly held that such reduction reflects, as Winford 2003: 16 puts it, a "universal tendency toward simplification of target structures, at least in the early stages of learning." Some of this simplification is the result of transfer effects from the L2 learner's first language, i.e., substrate effects, or imposition in van Coetsem's schema (see §3.1.2.2). For example, Winford (p. 218) points to the fact that in the learning of Dutch by English speakers, "the three-way phonemic distinction between /i/, /ü/, and /u/ in Dutch is . . . reduced to a two-way distinction (/i/ versus /u/)" ; inasmuch as /ü/ is not part of the English system, speakers simplify by reducing the contrast. In some instances, though, especially if native speakers of the TL accommodate in

40 We realize that contact between speakers of different languages could in principle lead one group to consciously or unconsciously alter their language so as to differentiate themselves from the others, rather than assimilate in any way to them. But while a possibility, we see this as a less usual circumstance (see §3.2.2.10 for a possible case of this sort, and Kulick 1992 for concrete examples).

41 Kerswill 2002: 696, drawing on work by Siegel (especially Siegel 2001), differentiates pidgin outcomes in language contact from koines that arise via accommodation (see §3.2.1.4).

the direction of simplifications made by L2 speakers of the TL, there can be a concomitant effect on the TL, with the emergence of what is referred to above in §3.2.1.4 as a compromise structure. For instance, in Russenorsk, a Russian-Norwegian pidgin, sounds and sound combinations that occur in only one of the languages contributing to the pidgin are generally replaced (Broch & Jahr 1984), and a similar set of outcomes is described by Broussard 2007 for Haitian and St. Lucian creoles (from French and Fon-Gbe input). Balkan examples where this kind of effect may be evident are discussed in §§5.2 and 5.4.1.6.

The role of simplification in language contact outcomes is complicated by the fact that there is no clear metric for determining what is simpler and what is more complex. It is generally held that more of something, and in greater variety, is more complex and that less of something, and less variation, is simpler. Thus Winford, above, labels a shift from a three-way to a two-way contrast as simplification, and Trudgill 2002: 711–712 characterizes as products of simplification the innovative two-form verbal system and three-form adjectival system of Norwegian Bokmål as opposed to the conservative Scandinavian five-form verbal system and two-stem, eleven-form adjectival system as seen in Faroese. On the other hand, the replacement of infinitives with subjunctive clauses in the Balkan languages resulted in increased complexity in agreement requirements in that more rather than fewer agreement contexts occurred in all multi-clause sentences. Thus, simplification in one domain may be counterbalanced by an increase in complexity in another. It does seem, however, that unmotivated, arbitrary differences within a system could place a greater burden on human memory and learning capabilities and thus be more difficult. This is especially so if adult learners are involved, as in some cases of language contact, and if adults simply are not as good at learning new material and routines as younger speakers.⁴²

One claim that has been made in this regard is the suggestion by Trudgill 2006 that certain types of contact situations can promote the elimination of morphophonemic alternations via leveling. Since leveling involves a regularization of patterns, it can be viewed as a type of simplification. Trudgill conjectures that if there is widespread second-language learning of the language in question by adults, then, to the extent that simplification is a characteristic of second-language varieties of a target language, one might expect to find fewer unmotivated morphophonemic alternations in speech communities where there are, or at least were, large numbers of nonnative speakers secondarily acquiring the language, especially as adults. As

42 We do not want to go so far as to say that adults lack the language-learning capacities of children, despite widespread belief in a “critical period” for native-like language learning. It is clear that language learning and thus language change go on throughout a person’s lifetime, as studies such as Sankoff & Blondeau 2007 demonstrate (see also Labov 2007: 380). Moreover, adults learn in different ways from children and with different motivations (e.g., to communicate rather than to fit in) and under different circumstances (e.g., they come at the language-learning task with a full language system, their native language, already in place – children simply have *more* acquiring to do), and these factors may go a long way to explaining differences between child language-learning abilities and those of adults. Moreover, there are adults who acquire near-native abilities in a second language so any seeming learning gaps or deficits on the part of adults cannot just be a matter of age.

a case in point, he suggests that the greater degree of regularization of strong verbs in North American English than in British English, giving e.g., *burned*, *spelled* instead of *burnt*, *spelt*, etc., may be due to the fact that historically there have been many more second-language learners of English in America than in Britain.⁴³ A Balkan parallel to this idea is the fact that Macedonian, especially the western dialects in the heart of the Balkan convergence area, has leveled out more morpho-phonemic alternations than Bulgarian (Markovikj 2007; cf. Elson 1995; see also §5.6).

3.2.1.6 Codeswitching

One type of apparent complexity that arises in contact situations, and in multilingual societies with bilingual speakers, is codeswitching, the apparent mixing of two different languages (“codes”) in one and the same utterance. Weinreich 1968: 73–74, in a passage reminiscent of Whitney’s 1868 or even Müller’s 1861 claim that mixed grammars or mixed languages could not occur (see §3.4.1.2) states: “The ideal bilingual switches from one language to the other according to appropriate changes in the speech situation (interlocutors, topics, etc.), but not in an unchanged speech situation, and certainly not within a single sentence.” At the very least, he treats intrasentential codeswitching as aberrant, but, as he admits (Weinreich 1968: 74): “The whole problem has hardly been explored.” It was arguably only in the 1980s that codeswitching studies became numerous enough to constitute a sub-discipline in linguistics, and we can cite here especially Poplack 1980, whose very title (“Sometimes I’ll start a Sentence in Spanish y termino en Español: Toward a typology of Code-Switching”) defies Weinreich’s formulation (as does, we can note in passing, the Hebrew Bible, Daniel II:4).⁴⁴ Nor is it a coincidence that her choice of codeswitching is between English and Spanish, since it was precisely Spanish in elementary education in the United States that was the primary focus of the Bilingual Education Act. The great majority of codeswitching studies, when not

43 Inasmuch as leveling is simply a type of analogical change, it is important to keep in mind that analogy can also introduce complication into the grammar, for instance by spreading, instead of eliminating, a minority pattern. A good example of this is found in the Balkans with the spread of the 1sg present athematic ending *-m* from the IE **mi*-conjugation throughout Balkan Slavic (and beyond, to varying degrees, in the rest of South and West Slavic). At the time of the earliest documents (OCS) only five verbs still had remnants of this conjugation, most characteristically the 1sg present in *-mъ* rather than back nasal *-o* found in all other verbs. By the modern period, however, 1sg *-m* had become regular for the most productive conjugational class in Bulgarian, it had become regular for all but two verbs in BCMS, and it was completely generalized as the ending in Macedonian. (For full details for all of Slavic, see L. Janda 1996.) Similarly, although the general trend in English is to level out strong verb preterites in favor of past-tense forms built on the present stem with the more productive inflectional material (e.g., *helped* replacing older *holp*), older *dived* has given way in American English to innovative *dove* with strong verb inflection.

44 It may well be that the social upheaval of the 1960s had an effect on the disciplinary position of the study of certain types of language contact (see also §3.4.1.2). Thus, for example, as Bokamba 1988 points out, interest in codeswitching grew significantly in the United States after it was given legislative impetus by the passage of the Bilingual Education Act in 1968 (cf. also Myers-Scotton 1993a as well as Mufwene 2004: 202 on the effects of political facts on the study of language endangerment).

concerned with the situation in the United States, have generally examined encounters between colonizing Europe and colonized Africa (e.g., Heath 1989), Asia, and the Pacific or between migrant laborers from poorer countries to the wealthier European employing countries (e.g., Backus 1996).⁴⁵ As observed in Friedman 1995b, it seems that the only pre-1989 study of codeswitching as such in the geopolitical Balkans is McClure & McClure 1988, a work on Vingar, Romania.⁴⁶ With this exception, the languages of Eastern Europe in general figured as topics of pre-1989 studies on codeswitching only when historical circumstance placed the speakers inside the borders of Western Europe (e.g., Gal 1979, 1988) or when Eastern Europeans emigrated to the West (e.g., Ewing 1984). Since 1989, Balkan codeswitching has generally received only the most passing of mentions in the general literature (e.g., Myers-Scotton 1993b: 219). Friedman 1995b is the first study to examine codeswitching in a Balkan context, looking at the phenomenon involving Macedonian and various other languages, especially Turkish, Albanian, and Romani, while Joseph et al. 2019 and Sobolev 2021a treat codeswitching specifically between Greek and Albanian in primarily Greek-speaking villages in southern Albania.⁴⁷ Treffers-Daller's 1992 observations on intergenerational differences in French–Dutch codeswitching in Belgium in the construction of monolingual (and destruction of bilingual) identities find a reflection in the reduction of discrepancies between declared nationality and declared mother tongue in what is now North Macedonia between 1953 and 1994 (Friedman 2003b). More recently the boundaries between codeswitching and second language acquisition have, in some cases, been blurred (e.g., Schmid 2005).

As Auer & Muhamedova 2005 point out, Stolt's 1964 study of Luther's German–Latin codeswitching *Tischreden* 'table speeches,' albeit of necessity based on written sources, is arguably the earliest modern codeswitching study. We can note, however, that Haugen 1953 also described codeswitching, in an immigrant context (see pp. 53–73, 298–300). With regard to the Balkans, Kappler's 1998abc studies of macaronic Balkan poetry demonstrate the same type of hybrid/cosmopolitan identity creation/expression as the French–Dutch codeswitching of the older generation of Belgian bilinguals. It can also be argued that Ottoman Turkish, with its integration of Persian and Arabic grammatical components, represents precisely the type of codeswitching phenomena that have received increasing attention in recent decades.

45 Menovščikov's 1969 article on Eskimo–Chukchi contact in Siberia deals with a then-communist country, but within the context of the former USSR, this latter represents a context of colonialism or ex-colonialism analogous to that found in many Western codeswitching studies. Labor migration in and of itself is a widespread phenomenon in both time and space, but its linguistic consequences did not receive much attention except in some studies of cryptoglossia (see §4.4.3).

46 The two languages, however, are Hungarian and Romanian, and so this is not a study of codeswitching within the Balkan sprachbund. We can also note here Weinreich's 1968: 74 passing mention of codeswitching between Romanian and Hungarian and between Romanian and French.

47 Also, there are codeswitching studies of languages from the Balkans in the diaspora; Hlavac 2003, for instance, examines Croatian in contact with English in Melbourne, Australia.

A prime focus of the codeswitching debate centers around whether the grammars of the two (or more) languages in question constitute a single unit in the mind of the individual (e.g., MacSwan 2009, 2021, and minimalist approaches in general) or whether there are two grammars, and, as a correlate to that, whether one of them must dominate at any given moment (e.g., Myers-Scotton 2002, 2006); see also footnote 7. Both these approaches beg two important questions: (1) social factors in determining codeswitching, and (2) the fact that in multi-generational stable bilingual situations it is possible to have contexts in which neither language is dominant, while in other situations it may be possible that a single mixed grammar is operative with sociolinguistic constraints on the relative dominance of one or the other component. Attempts to theorize codeswitching that rely on purely formal constraints, as in the minimalist approach, end up making false or unfalsifiable claims, since they must posit the impossibility of data that actually do occur or they must rely on so-called grammaticality judgments in situations where the interactions of the grammars involved are as fluid as the social situations in which they occur; moreover, the creation of artificial situations (elicitations) gives artificial data.⁴⁸ Thus, for example, while it may be the case that English articles do not occur in Spanish-dominant (matrix) codeswitching situations, this has nothing to do with the properties of the definite articles in the respective languages, since English articles will occur with Spanish nouns in an English-dominant code-switch.⁴⁹ On the other hand, in a fully bilingual Spanish–English household, where one parent has Spanish as the first language and the other parent has English and the children are raised bilingually, the situation is much like that identified by Auer & Muhamedova 2005 for Kazakh–Russian codeswitching or Friedman 2009a for Romani–Turkish in some Romani dialects, namely, at any given moment, any given language may be employed. However, both the Russian–Kazakh and Turkish–Romani data demonstrate that code-switches can also result in the creation of utterances in which neither language can be said to be dominant because a new, third, interlingual, situation is created without, however, creating a third language. It can be argued that precisely such transferences were involved in Balkan contact

48 That is, merely presenting a sentence or phrase to a speaker and asking for a judgment as to its status may well summon up prescriptivist and puristic pressures for the speaker which are especially problematic for judging codeswitching examples, given the often-negative attitudes towards language “mixing.” It can be argued that much of the work on constraints on codeswitching in recent years especially focuses on what educated speakers do in controlled laboratory settings, whereas the Balkan situation has involved untutored second language acquisition so that any observed codeswitching involves knowledge of other languages, to be sure, but not guided or consciously arrived at knowledge.

49 We note though that although it has been claimed (MacSwan 2007; similarly Costa 2004: 209) that one will not find codeswitched mixes of Spanish nouns with English adjectives, and vice versa, with inappropriate syntax for the lexical items involved as well, i.e., *the casa white / *the house blanca, just such modified noun phrases have entered English through contact with other languages: *galore*, for instance, from an Irish adverbial modifier (*go + leor* ‘enough’), is now a somewhat anomalous obligatorily postposed adjective in English (*He has friends galore* / *galore friends); cf. also the anomalous Noun–Adjective structure of *attorneys general*, due to its French source (though some speakers Anglicize this to a Noun–Noun compound, *attorney generals*, reanalyzing the etymological adjective as a head noun).

phenomena, where phonology,⁵⁰ morphology, morphosyntax, syntax, and lexicon were all exchanged, but not in a uniform manner and under varying social and geographic circumstances.

To summarize, then, we can say that for the Balkans, codeswitching of both inter- and intra-sentential and clausal varieties all occur and have occurred, although historical data are not always clear or available.⁵¹ The triggers are as varied as the situations in which the phenomenon occurs. Issues of prestige, solidarity, and identity formation are all obviously present. At the same time, however, the fact that in a centuries-long, stable, multilingual situation the languages that we see today are all known to have retained their distinctiveness – the obscure factors being which language survived as Albanian, on the one hand, and how it came to pass that those languages that died out, e.g., Germanic, Celtic, some Turkic, and various Paleo-Balkan languages, shifted, on the other – points to a combination of widespread bilingualism and equally widespread linguistic identity maintenance.⁵² Thus, while codeswitching in the Balkans is both attested historically and occurs today (see also Friedman 2006b; Stoica 2003–2004), its role in contact-induced change is difficult to assess in the absence of longitudinal studies of significant time depth. It would seem that in each of the Balkan languages, the maintenance and shifting of boundaries at each linguistic level from the phonological to the morphosyntactic and lexical would indicate that mechanisms other than codeswitching were at work.⁵³ The one area where, in the Balkans, as elsewhere, codeswitching is arguably a path to language change is the transition from so-called nonce-borrowings – single-occurrence, one-word switches – to integrated lexical borrowings (see Heath 1989). And with this we come to the issue of borrowing as a mechanism.

50 An oft-cited putative phonological constraint in codeswitching is that it cannot occur word-internally (cf. MacSwan 2005, 2007; see also MacSwan 2021), yet there are cases of the phonological “intrusion” of one language into just part of a word, as the case of Lorain, Ohio Puerto Rican Spanish /r/ discussed by Ramos-Pellicia 2004 (and see §5.4.1.6) shows; see also Rezaeian et al. 2006 on examples of Persian codeswitching with Persian-inflected English words with initial clusters that are otherwise impossible in Persian (e.g., *problem-i* ‘problem-INDF’). We note too that Holden 1976 demonstrated variable application of different Russian phonological processes in English loanwords into Russian of the mid-twentieth century, again suggesting that the phonetic implementation of a single word need not have input just from one language. And, the example of the phonological compromise hybrid form [mexanikós] discussed in §3.2.1.4, from the Modern Greek of southern Albania, offers a case from a present-day Balkan situation, as does [dáfni] cited in footnote 36 for δάφνη (with initial [ð]) elsewhere in Greek, due to the influence of Albanian *dafinë*; see also footnote 37.

51 Friedman 1995b has argued that ethnic anecdotes (jokes) collected in the nineteenth century can give us a window into colloquial codeswitching practices at a temporal distance no longer accessible to us by other means. Moreover, the Balkan literary examples alluded to above also have heuristic value for speculating on the speech of educated individuals.

52 See Friedman 1995b on emblematic codeswitching in the nineteenth-century Balkans. Cf. also codeswitching as a politeness convention in the Skopje *čaršija* ‘bazaar’ (Friedman 2019c).

53 This might well be the case in similar situations such as that described by Gumperz & Wilson 1971 for Kupwar. Muysken 2000: 274 provides a useful typology for codeswitching (or code mixing in his terms) vis-à-vis contact-induced language change, with the following parameters: borrowing, relexification, calquing, genesis, shift, convergence measured against insertion, congruence, and congruent lexicalization.

3.2.1.7 Borrowing

It has been impossible to discuss the Balkan languages up to this point without some mention of ‘borrowing,’ understood as “the incorporation of foreign features into a group’s native language by speakers of that language” (Thomason & Kaufman 1988: 37; see now also Poplack 2017 and Adamou & Matras 2021). Not only has such incorporation of foreign words into the lexicon – the *locus classicus* (to borrow a Latin phrase) of borrowing – been exemplified for several languages, whether ancient and now dead or contemporary and still actively used, but the process itself has also been used earlier in this chapter to make a point about speaker-to-speaker interactions and about speaker agentivity.⁵⁴ Still, there are important issues about borrowing as a process of inter-speaker interaction in situations of language contact that are appropriately summarized and signaled here in the context of a consideration of mechanisms relevant in contact-induced change.

The key questions concern just what is “borrowed” when borrowing takes place: what is it that speakers do in the “incorporation” process and what sorts of material can they incorporate?

To address the first question, a key observation is that what is borrowed is not abstract aspects of languages but rather concrete and surface-oriented material. There are differences of opinion as to what constitutes “abstractness” in these cases of cross-language interaction. King 2000: 82, for instance, says that while “many linguists conceive of lexical borrowing as the borrowing of ‘merely’ words,” she feels that “following recent work in generative grammar . . . words are borrowed in a fairly abstract form, and the transfer of bundles of syntactic and semantic features along with phonological information is involved.” Labov 2007: 349, on the other hand, observes that in diffusion (i.e., in this case, borrowing), the copying (i.e., the spread) of material “is limited to the most superficial aspects of language: words and sounds.”⁵⁵ He continues (p. 349, footnote 6), “More precisely, adults borrow observable elements of language, the same elements that can be socially evaluated.” And, Joseph 2001a: 22 points out that convergent phenomena typically involve superficial linguistic material, focusing on common words, common uses, common combinations, common strings of elements, and so on. (Cf. Silverstein’s 1971 discussion of Chinook Jargon and Gal & Irvine 2019.) Moreover, since lexical borrowing – a quintessentially surface-oriented phenomenon so widespread in the Balkans – can shade off into construction borrowing and thus syntax, it becomes, Joseph argues, “problematic to view [. . . Balkan structural]

54 See §1.2.1.1 on Ancient Macedonian possibly showing non-Greek loanwords, §1.2.1.5 on a pre-Greek Indo-European stratum in the Greek lexicon, §1.2.1.9 on old Germanic loans in Albanian and Romanian, §1.2.3.1 on other loans in the Albanian lexicon, §1.2.3.5 on Romani borrowings, including some from Turkish, and so on. Note too the discussion in §2.3.2.2 of the predominance of a concern for borrowings and lexical parallels more generally in Sandfeld 1930, and Chapter 4 in general.

55 Cf. Wertheim’s 2003 characterization of folk approaches to language as consisting of “nouns and sounds.” Also, see further in this section on the matter of the borrowing of grammatical material and structure.

similarities in terms of deep syntactic features such as parameter settings” if some “Sprachbund significance for these features” is to be claimed. Attributing sprachbund relevance for such features, he continues, “would be inconsistent with their deep nature, since the ‘action’ in language contact, so to speak, is at the surface, not at a deep level, yet contact is crucial for the development of a Sprachbund.”

We note here in this respect that this means that the quest for “universal grammar,” generated and dominated for several decades by the Massachusetts Institute of Technology school of linguistics, does not pose a fruitful avenue of investigation for Balkan linguistics, i.e., for examining sprachbund, contact-induced, phenomena as opposed to linguistics of the Balkans. One can wonder, for instance, how insights obtained from the quest for universals, oriented as it is to levels below that of surface structure are, as Joseph puts it, “revealing beyond what might be found if one were to compare any arbitrary set of typologically related languages chosen on a basis other than geography.” Thus, studies of abstract structures in language convergence in the Balkans “are interesting from the perspective of the ‘Linguistics of the Balkans’/Comparative syntax of the Balkans, but not from the perspective of ‘Balkan Linguistics’/comparative Balkan syntax.” In other words, studies of the Balkan languages that explain their commonalities in terms of universal linguistic features tell us nothing about the contact phenomena that have actually led to the observed commonalities (see also §7.1).

To return to the first issue, it has long been noted (see, e.g., Hock & Joseph 2019: 223; Myers-Scotton 2006) that the term “borrowing,” though in widespread use and solidly entrenched in the literature on language contact, is something of a misnomer. That is, the donor language is not deprived of anything in the process and what results in the receiving language is not an exact duplicate, but generally an imperfect replica that does not mirror the original source exactly. For instance, as discussed in §5.3, the sounds of a borrowed word are often, albeit not always, adjusted to suit the phonological system of the borrowing language.⁵⁶ Similarly, the morphology, meaning, and even word-class of the original need not be preserved as it is incorporated into a different language.⁵⁷ Such considerations have led some researchers to develop alternative terminology. As seen above in §3.2.1.2, Johanson 1992, 2002, 2023, for instance, uses the term *copying* (as does Gołąb 1976) for what is usually called “borrowing” (or calquing), partly in recognition of the replicative nature of the incorporation process.⁵⁸ We can also note here Matras’ 2000 term *fusion*, which he uses to describe a type of borrowing that involves the integration of material from L2 into L1 as a means of reducing cognitive load, not because of but despite the L2 form belonging to a different system. He uses the

56 For instance, English borrowings from languages with word-initial unaspirated stops typically are pronounced with English-style aspiration; thus *taco*, from Spanish [tako], is usually pronounced [tʰakow] by English speakers (with an English diphthong as well).

57 Consider, for instance, *affidavit* and *alibi*, a verb and adverb respectively in the source language, Latin, but nouns in the recipient language, English. For several other examples, especially regarding semantic shifts in borrowing, see Winford 2003: 33, 43–46 and Matras 2009: 174–175.

58 In Haugen’s 1950: 212 classic definition of borrowing, the term “reproduction” is used (“the . . . reproduction in one language of patterns previously found in another”).

example of discourse particles, a category that is important in the discussion in Chapter 4 (see §4.3.4). Ross's *metatypy* (cf. Ross 2006) can also be noted here. While the term's basis involves massive grammatical calquing, the calquing can also have effects at the lexical level, e.g., *bor-goun* 'animal,' lit., 'pig-dog' in the Austronesian language Takia, is a calque on the model of *buruk-kasik* 'idem' in Waskia, which is in the Trans-New-Guinea family. While we generally use *borrowing* here,⁵⁹ we recognize the value of new terminology for the new perspectives that it affords onto a familiar phenomenon.⁶⁰

Once one recognizes that it is actually copying of material, rather than borrowing per se, that is involved in bringing elements of a foreign language into one's own, then it is just a short conceptual step to the notion of "calquing" or "loan translation."⁶¹ In such a contact-induced change, native morphemes are substituted for the pieces of an external model, as in Greek ουρανοξύστης 'skyscraper' based on an external model, possibly English *skyscraper* directly (given the parallel meaning and order of each element of the compound, ουρανο- 'sky' / ξυσ- 'scrape'), though German *Wolkenkratzer* (lit., 'cloud-scraper') or possibly even French *gratteciel* (lit., 'scrape-sky') conceivably could have played a role in the Greek form, despite the difference in the meaning of the compounding elements in the German and in their order in the French. We see such calques as just one instantiation of a more general process of blending or fusion of elements properly belonging to different languages in contact, with the elements in such cases being the semantics and word-formation "template" from one language (e.g., SKY + SCRAPE + AGENTIVE) and the actual forms from another (e.g., ουρανο- + -ξύσ- + -τη-, from Greek itself).⁶²

Calquing therefore means that material other than the word, and possibly greater than the word (in the sense that 'SKY + SCRAPE-' is a multi-word compound), can be "copied." While on the one hand, such a process indicates a cross-language awareness and at least minimal bilingual knowledge on the part of calquing

59 Matras 2004 and Sakel 2007 have made a distinction between "mat-borrowing" and "pat-borrowing", referring to the borrowing of phonological/morphological material (matter) on the one hand and the borrowing of a pattern on the other. We see this distinction as valuable but in essence no different from simply talking about borrowing and the more traditional notion of calquing (on which see further below).

60 Müller 1854: 32 provides an interesting insight into the ideological value of such terminology when he refers to borrowed vocabulary in Romani as "stolen" from other languages. In a form of tit for tat, Bryer 1983 referred to putative Romani words in English as "stolen" from Roms by English speakers.

61 Indeed, Comrie 2002: viii in his introduction to Johanson 2002, explicitly draws the connection between (some) copying and calquing: "Johanson employs the useful term 'copying' to subsume both kinds of contact [his adoption and imposition], with a distinction between global copying, where the whole of the form and function of a structure is copied, and selective copying, which is similar to the traditional 'calquing', for instance copying of a structural pattern but using indigenous morphemes." See also footnotes 58, 59, 60.

62 Note that other aspects of this word are the result of Greek-internal processes, e.g., the formation of compounds with the stem-form of the noun, with the perfective aspect stem of the verb, etc. We do not consider such calques to involve abstract elements in contact, even with notions such as 'agentive' or semantic templates with particular combinatorics involved, since the surface realizations are crucial, giving the basis upon which speakers can recognize elements of the template.

speakers,⁶³ on the other hand, it can be a way for new patterns to enter a language. That is, when calquing leads to the deployment of recipient-language morphemes and words in novel ways, then new patterns are created in the recipient language. For instance, double determination in Balkan Slavic as illustrated by Macedonian use of the definite article together with a demonstrative, as in *ovie deca-va* ‘these here kids here’ (lit., ‘these.PRX kids-DEF.PRX’ and related structures in Bulgarian and Torlak BCMS; see §6.1.2.3), is an innovation compared to the rest of Slavic; it can be taken to be based on a Greek, Albanian or Balkan Romance pattern, as seen in Grk αὐτός ο ἄνθρωπος ‘this man’ (lit ‘this the man’), a pattern that has been a part of Greek since the development of a definite article in the language in the seventh to fifth centuries BCE.⁶⁴ Balkan Slavic speakers innovatively expanded the distribution of their definite article to match the distribution of the corresponding element in the Greek (and Albanian and, to some extent, Balkan Romance), thereby creating a new pattern for the language of demonstrative–noun–article where previously only demonstrative–noun or noun–article had been possible. This example also offers insight into the second key issue indicated above, in that it shows that function words can be the focus of cross-language spread. Such an example therefore leads to the question of whether there are any restrictions at all on what can be borrowed across languages, a matter of considerable controversy in the contact linguistics literature and touched on above as the second key issue, stated there in terms of what sorts of material speakers can incorporate into their grammar as a result of borrowing. The controversy is summarized well in Thomason & Kaufman 1988: chapter 2, to which the reader is referred for details and relevant bibliography, but this key issue basically boils down to whether borrowing is only lexical in nature or instead can target structure as well. Thomason & Kaufman argue for a very broad view, claiming – after a careful survey of the relevant evidence – that there are no purely linguistic constraints on borrowing: given the right social circumstances, any element of one language can be borrowed into another language (cf. also Friedman 2007a: 204).

A very strong contrary stance is taken by Nakhleh et al. 2005, echoing some of the earlier positions reviewed by Thomason and Kaufman. They dispute the cogency of the examples of structural borrowing that Thomason and Kaufman raise, and they rely heavily on the findings of Ruth King 2000. King examines the entry of preposition stranding from English into Prince Edward Island French (in Canada), as in (3.1):

- (3.1) a. Où ce-qu’ elle vient de
 where that she comes from
 ‘Where does she come from?’
 b. Quoi ce-qu’ ils parlent about
 what that they talk
 ‘What are they talking about?’

63 The value of calques for identifying and assessing sprachbunds is explored in §3.4.2.2.

64 It is a widespread pattern in Albanian, Balkan Romance, and Romani, but owing to the lack of textual evidence in the case of Albanian, and the later arrival in the region of Balkan Romance and Romani, the only attested likely source is Greek, although Albanian may also have played a role.

She argues that this phenomenon is not a case of structural diffusion but rather represents a type of lexical borrowing, in that the properties of particular borrowed prepositions of English origin, like *about* in (3.1b), have been carried over from English into French, with the properties of such originally English forms extended “to the whole set of Prince Edward Island prepositions” (p. 147).⁶⁵ Nakhleh et al. 2005 generalize from King’s demonstration regarding Prince Edward Island French and categorically rule out structural borrowing altogether. Winford 2003: 79–80 is less categorical, but takes a very cautious position on the possibility of structural borrowing, saying (p. 64) that “the case for direct borrowing of structure . . . has yet to be proved.”

We consider it unfortunate that so much of this recent debate on structural diffusion has taken place in the absence of virtually any consideration of the most intensely studied instance of structural diffusion, namely what is to be found in the Balkan sprachbund. The example noted above of the demonstrative + definite article pattern spreading from Greek (and/or Albanian or Balkan Romance) into Balkan Slavic would seem to be a clear case of structure diffusing. Indeed, this controversy would seem to be a nonissue if one were to take at face value the evidence from the Balkans, where so many structures appear to have passed between languages. That is, besides the demonstrative example, one can also note the enclitic definite article itself, found in Albanian, Balkan Romance, and Balkan Slavic,⁶⁶ which, while likely to be part of a Balkan substratum language (so Hamp 1982), nonetheless diffused into these languages, and in the case of Balkan Slavic clearly did so after the creation of a definite article in these languages within the historical period. One might well think here in terms of transfer in second-language acquisition as the substrate speakers shifted to a new language (whatever gave rise, e.g., to Albanian and Romanian) but that is surely a type of adult-to-adult diffusion; moreover, the spread to Balkan Slavic would seem to qualify as structural diffusion. Further, in some western and southern Macedonian one finds the marking of direct objects with the preposition *na* (otherwise meaning ‘to’ or ‘of’), a structural innovation that can be traced to the influence of the co-territorial Aromanian language, in which the preposition *pe*, from Latin *per*, is used to mark animate direct objects (as in Romanian), just like the structure found in

65 We discuss a case like this from the Balkans, involving the positioning properties of the Greek borrowing *vte*, quite likely of Turkish origin, in §4.3.4.1.2 and §5.5.3, though these properties did not spread to other elements in any of the languages; so also with Turkish postpositions borrowed as such into generally prepositional Balkan languages, discussed in §4.3.3.2. We note further that if syntax – as is increasingly the case in recent lexically oriented theoretical frameworks, e.g., Head-driven Phrase Structure Grammar or the Minimalist Program (and even Construction Grammar, though the unit of interest there is larger than the lexical item) – is a matter of the projection of lexical properties onto syntactic structures, then the distinction between lexical borrowing and structural borrowing, as far as syntax is concerned, becomes rather tenuous at best. To some extent, this is at the core of King’s approach to grammatical borrowing in Prince Edward Island French, though her demonstration of the lexical character of that situation is not tied to her theoretical assumptions.

66 See also §§1.2.1, 3.2.2.5, and 6.1.2.2.1 for more discussion of this feature, from different perspectives.

these Macedonian dialects. See also Matras 2007 on structural diffusion and §6.1.1.1.2 on differential object marking.

In each of these cases, it is hard to say that the properties of some individual lexical item were the point of entry for the new structure, which is what a “structural-borrowing-as-lexical-borrowing” account in the manner of King or Nakhleh et al. would presumably say: in the case of the demonstrative and definite article, there are two lexical items involved (demonstrative and article) so it is not clear which would be said to be the basis for the “lexical” borrowing, and at any rate, the material itself is not borrowed but rather the structural properties of one language’s elements are transferred onto a recipient language’s corresponding element; in the case of the enclitic article, we can say that the article would seem to be a rather unlikely candidate for being the focus of lexical borrowing, given its prosodic weakness (and again, it is not the item so much as the properties that are transferred); and finally, there is no single lexical item involved in the object-marking example, but rather a structural position/entity, namely direct objects.

It is of course possible that these patterns entered the “borrowing” language in a single borrowed construction or phrase and spread from there. There is no way to argue against that scenario, and it can and does happen. A telling example is the spread of the VERB-‘not’-VERB construction in the Balkans with the meaning ‘whether one VERBs or not,’ which varies in productivity across Balkan languages though they all share the construction with the verb ‘want’; presumably, that lexicalized instance was the model upon which even the productive pattern in some of the languages are based (see §4.1 and Joseph 2000c). But if that is the type of explanation one resorts to in each potential case of structural borrowing, even when no clear lexical items are involved, then the issue of structural diffusion ceases to be an empirical matter. It should be clear by now where we stand on this issue: we take the Balkan evidence to offer strong support for the Thomason and Kaufman (and our own) position that under the right social circumstances, anything – including structure – can be borrowed from one language to another.⁶⁷ Thomason and Kaufman were perhaps not as explicit as they might have been about the social conditions they had in mind for structural borrowing and surely some instances of apparent structural borrowing involve generalization from a lexical borrowing source, as King proposes in her detailed (and quite convincing for what it examines) case-study, but ruling out this type of borrowing altogether is not justified by the larger body of evidence.

By way of concluding this discussion, we consider two further points about the nature of borrowing affected by Balkan evidence. First, functional (grammatical)

67 Labov 2007: 349 recognizes this, though he seems to do so only reluctantly. He cites approvingly Sankoff 2002 and her claim that “linguistic structure overwhelmingly conditions the linguistic outcomes [of language contact]” (p. 658), though her use of “overwhelmingly” carries an implicit recognition that in some instances there can be no constraints on what is borrowed. And, Labov himself cites a few instances of structural borrowing in the Balkans – the Verb-‘not’-Verb construction, which he (perhaps) rightly dismisses, and the replacement of the infinitive by finite complementation, which he accepts – but nonetheless concludes his brief review of the matter by saying “contributors to this debate agree – with the exception of Thomason and Kaufman – that there are

material (function words and morphology) is typically said to be harder to borrow than content lexemes, and it has been claimed that bound morphemes, specifically affixes, i.e., material smaller than the word, are not borrowed per se but come into a language attached to lexical material and are then extracted from that material and extended to other forms.⁶⁸ There is no linguistic reason, however, to restrict by grammatical or semantic class the types of elements that can be borrowed: Friedman 2003a: 6–21, for instance, documents the range of Turkisms in Macedonian and notes (p. 14) that “Turkish lexical borrowings belong to all levels of vocabulary and almost all parts of speech – noun, verb, adjective, adverb, conjunction, preposition, pronoun, exclamation, particle,” and though no numbers per se have been borrowed in any of the Balkan standard languages, “there are Turkisms in numerical expressions.”⁶⁹ And, one can see that “Turkish vocabulary has penetrated every facet of Macedonian life: urban and rural,” including material culture, abstractions, and even kinship terms, color names, and body parts. We would thus, with Thomason and Kaufman, attribute any apparent limitations on the nature of borrowings to the social conditions under which the borrowing occurs rather than the grammatical status of potential candidates for diffusion across languages.

Moreover, Turkish derivational suffixes occur in all the Indo-European Balkan languages. Admittedly, these might have entered in particular words, but a few have taken on productive lives of their own and have been extended well beyond the bounds of their original distributional locus, making it hard to see what the original locus of diffusion would have been. With regard to inflectional morphology, the occurrence of the Turkish plural suffix *-lar* in Albanian (and some dialects of Macedonian) would seem to be a case of the affixes entering attached to specific lexical items, since it is generally found only with Turkish lexemes (e.g., *baba-llar(ë)* ‘fathers,’ *bej-ler(ë)* ‘Turkish notables’),⁷⁰ but it has spread beyond that to other nouns for males with prestige (e.g., *kalogjerler(ë)* ‘monks,’ ultimately from Greek), and in any case, it is not clear that one can generalize from such instances to all cases of the apparent borrowing of bound morphemes. For instance, some Romani dialects of eastern Bulgaria in contact with Turkish borrowed the Turkish second person plural marker *-niz* as *-[n]əs* in the preterite and then generalized it to the native Romani first plural preterite, e.g., Turkish *gel-di-k/gel-di-niz* ‘we/y’all came,’ Romani *gelj-am-əs, gelj-an-əs* ‘we/y’all went’ (see Elšík &

structural limitations on what types of linguistic patterns can be transmitted across languages.” Add our names (including Friedman 2007a) to his list of “exceptions”!

68 Lefebvre 1984, for instance, makes this claim, citing, among others, the example of the entry of the Spanish plural marker *-s* into productive use in Quechua through the vehicle of borrowed Spanish nouns.

69 Moreover, Turkish numerals were borrowed into a number of Bulgarian and Pomak dialects of Thrace, and Common Romani borrowed the numerals for ‘seven,’ ‘eight,’ and ‘nine’ from Greek. For a fuller presentation of Friedman’s Macedonian material, see §4.3 (in which there is also relevant discussion of “ERIC” loans).

70 See §5.6 for more discussion of Turkish plurals in Albanian from the phonological perspective, and §6.1.4.1 for more from the morphosyntactic perspective.

Matras 2006: 135–136).⁷¹ Moreover, on a conceptual basis, it seems that speakers who come to be highly familiar with a second language, even if that familiarity is acquired in a naturalistic rather than tutored setting, can recognize repeated instances of a piece of a word, i.e., an affix, work with it, and even give it a life of its own in their first language; the *Bela di Suprã* (Mac Gorna Belica) Aromanian admirative using the Albanian third person admirative marker *-ka* (Friedman 1994b) shows that affixes can be salient to second-language users.⁷²

A second relevant claim is that structural congruence between systems is a necessary condition for structural or grammatical borrowing. This view has been espoused by such influential thinkers as Meillet 1921: 87, with reference to languages with very similar systems, and Jakobson 1938/1962: 241, suggesting that “a language accepts foreign structural elements only when they correspond to its own tendencies of development.” Still, as Thomason & Kaufman 1988: 14ff. say in their discussion of this claim, “it is particularly hard to disprove, because it has at least the potential of being circular,” in that the ability of a language to borrow feature X could be taken to demonstrate a predisposition to borrowing it. Nonetheless, we note here one striking case involving a language of relevance to the Balkans that would seem to be a *prima facie* case of grammatical borrowing across noncongruent systems. As mentioned previously (see §3.2.1.5), Ottoman Turkish borrowed from Persian a subordinating conjunction (complementizer), *ki*, that had the property of selecting for a finite verb in the clause it headed. This finite complementation strategy, however, did not correspond to any sort of structural possibility already present in Turkish at the time of its introduction into the language; in fact, Turkish had only nonfinite complementation prior to the borrowing (involving verbal nouns, infinitives, and participles) and the *ki*-complements now still constitute the only finite subordination that occurs in the standard form of the language.⁷³ It thus presents a convincing counterexample to claims that borrowing a structure requires congruence between the systems involved.

Therefore, we conclude, siding with Thomason & Kaufman 1988, that grammatical borrowing is unconstrained by purely linguistic structural conditions. The challenge is to identify the social conditions that are conducive to the borrowing of structure; clearly, the cross-language interactions in the Balkans constitute the right sort of social conditions and attitudes for that to happen.

71 The resemblance between the roots for the motion verbs in these two languages is coincidental and not a case of borrowing. The native Romani forms would be *geljam*, *geljan*. See §6.2.1.1.1 for a fuller discussion of the borrowing of Turkish inflection into a wide range of Romani dialects.

72 See §6.2.5.6 for details on this example of grammatical borrowing, and §6.2.1.1.3 for a debunking of a much-cited putative case involving Meglenoromanian and Macedonian.

73 See §7.7.2.1.1.2.6 on the introduction of finite subjunctive complements in Balkan Turkish under influence from other Balkan languages. Slobin 1986 demonstrates that this pattern is cyclical in Turkic languages in general: subordinating conjunctions are borrowed from Indo-European languages, eventually turned into agglutinative suffixes, and then new subordinators borrowed.

3.2.1.8 Language Ideology

Underlying any interaction among speakers, including speakers of different languages or different dialects is some form of language ideology, defined by Silverstein 1979: 193 as “any sets of beliefs about language articulated by the users as a rationalization or justification for perceived language structure and use.” That is, language ideologies represent a guiding system of beliefs and attitudes that speakers hold about their language or dialect and about other languages or dialects, and about language and dialect in general. These beliefs can be very strong, and need not have any overt verbal manifestation, but they are very real for the speakers who hold them, and thus they can be powerful forces in shaping the direction of change, especially in contact situations. Such beliefs include ideas about correctness in language use, about the values attached to certain registers or varieties of speech, even about what constitutes part of “our” language: all such attitudes can be subsumed under the general rubric of language ideology, and, all of these can come into play in language contact.⁷⁴

The ways in which language ideology influences the effects of language contact are as varied as the social situations of contact themselves. Thus, Kroskrity 2000a documents the strict exclusion of borrowed lexicon and grammar from Arizona Tewa, a Kiowa-Tanoan language in contact with Hopi, a Uto-Aztecan language, on a Hopi reservation surrounded by Navajo, a Na-Dene language, and all of them part of a formerly Hispanophone and now Anglophone larger society. (See also Dozier 1956 on other Tewa communities.) Arizona Tewa-speakers are generally trilingual in Tewa, Hopi, and English and some also know Navajo and Spanish, but speakers of these other languages do not learn Tewa. Kroskrity suggests that the centering of Tewa norms on ceremonial language provides an ideological model that results in a relatively high degree of linguistic purism that affects both lexicon and grammar. A different kind of purism is described by Aikhenvald 2007a, 2007b for the Vaupés linguistic area of northwestern Amazonia (a region spanning Brazil and Colombia).⁷⁵ Here, for example, Tariana, an Arawak language, is in contact with, among others, East Tucanoan languages through an exogamous marriage network that requires multilingualism of all speakers. The kind of structural isomorphism that we associate with the Balkan sprachbund is present, but an ideology of linguistic separation results in a situation where, contrary to the usual expectation, the borrowing of free morphemes does not occur, and the few morphemic borrowings are bound (which, Aikhenvald 2007a: 40 argues, is due precisely to their being less overtly detectable as units). In the case of Romani, to come back to the Balkans, the segregation of native and pre-Byzantine borrowed vocabulary from post-Byzantine borrowed vocabulary by means of the presence versus the absence,

74 See §2.5.2 on some additional examples of language ideology in Balkan linguistics, and §5.2 and §5.3 for specific cases involving ideology and its effects on phonology. For general considerations see Friedrich 1989; Silverstein 1979, 1998; Schieffelin et al. 1998; Woolard 1992; Woolard & Schieffelin 1994; Kroskrity 2000b; and Gal & Irvine 2019.

75 See now Epps & Michael 2017 on areal issues in Amazonia more broadly.

respectively, of a thematic vowel (e.g., Friedman 2009b) is yet another example of how attitudes toward maintaining boundaries are realized via grammar.

For the Balkans, the study of language ideology is an area in need of further exploration. Friedman 1997a addresses general issues of language ideology in the Balkan languages, particularly in connection with the creation of standard languages and modern nation-state identities; Gal & Irvine 1995 and Irvine & Gal 2000 use the case of Macedonian, among others, to illustrate broader theoretical generalizations on how ideologies construct differences and boundaries of languages and disciplines (see also Gal & Irvine 2019); and Tsitsipis 1998 gives a specific case study, namely Arvanitika (Albanian) in Greece, but much more remains to be done. Among the larger political effects of language ideology was Bulgaria's blocking accession of North Macedonia to the EU in 2020 claiming, among other things, that Macedonian is a dialect of Bulgarian (Friedman 2020a). An interesting point regarding ideology can be made using the epigram at the beginning of this book. The proverb suggests an ideology of openness to multilingualism, at some period in the past at least, that is not necessarily the norm in all contact situations.⁷⁶ Moreover, in the case of Judezmo, the fact that the proverb is used but cited in Turkish is, in a sense, an instantiation of ideology behind the proverb itself as well as of the position of Turkish for Judezmo-speakers in Ottoman Europe (cf. Friedman 1995b).

What might be considered the basic act in language contact, namely that of delimiting one's own language or dialect and thus of recognizing what falls within the scope of "my language" and what falls outside, is also basically ideological in nature. Given that there are generally no sharp boundaries between co-territorial languages that share certain features or are used by many speakers, or between dialects of a language, speakers are more or less free to construct boundaries where they see them or where they perceive them, even if the analytical eye of the linguist might see the situation differently.⁷⁷ And, speakers are free to (and sometimes do) believe that languages can be subject to ownership, that such ownership can – or should – be marked in some overt way, and that their actions linguistically can work towards those ends. Thus it is possible to view the creation of a mixed form like the Tsakonian negative marker *don* – described above in §3.2.1.4 in purely structural linguistic terms as a blend of or compromise between two competing forms (inherited Tsakonian *o* and standard *δεν*) – as being an ideologically driven act.

⁷⁶ As examples of the effects of language ideology on the historical record, see Fine 1983: 220 citing Mošin 1963: 54–69 on the destruction of Slavic manuscripts by the Byzantine Greeks, and Lindstedt 2000 on the connection between ideology and resistance to borrowing.

⁷⁷ Note that positing the existence of an "objective" manner of gauging language or dialect differences, i.e., the assumption that there is a "truth" to be uncovered about language divisions, is itself an ideological belief. It is occasionally useful for heuristic purposes, but it is a matter of ideology nonetheless. Just as Stolz 2006: 45 writes of sprachbunds that they "are not objects in the physical world but rather projections from the minds of linguists" (though see now the various pieces in Hickey 2017 for relevant discussion of variously real linguistic areas) and that therefore "truth conditions cannot be applied to the various competing alternative notions," so, too, divisions of speech communities into languages and dialects and gauges of difference among them are basically subjective phenomena constituted by speakers and/or observers through acts of instantiation.

Under such an interpretation, Tsakonian speakers recognized or imputed a value to the standard language form and decided to act on that evaluation, making that form theirs first by incorporating it into their usage and then “owning” it through the substitution of the Tsakonian vocalism. So too with choices that speakers make about accepting (copying) loanwords (copies) and about how to incorporate and make part of their system any outside element they accept: all such choices have an ideological dimension.

Another contact phenomenon with a distinct ideological dimension is code-switching, the general principles of which are discussed in §3.2.1.6. Here we can note that speakers of more than one language at times constitute their identities precisely by alternating between (see Trudgill 2018) the languages they know, signaling various social factors by the choice of language in which they express any given thought or part of a thought. In this sense, codeswitching functions as a kind of register (Paz 2018). Kappler 1998abc thus discusses Balkan macaronic literate codeswitching involving Greek, Turkish, and Balkan Slavic and Friedman 1995b treats codeswitching in nineteenth-century Macedonian ethnic anecdotes as indices of various types of language and gender statuses (see also Friedman 2006b, 2018d); see also Le Page & Tabouret-Keller 1985 on codeswitching in general as acts of identity. We can note here that to the extent that codeswitching is socially determined, attempts to formulate strictly formal constraints on the phenomenon will founder on the fluidity of actually occurring situations as well as speakers’ abilities to manipulate their own resources. Woolard 2004 offers important discussion of indexicality, footing, voicing, and contextualization as theoretical constructs that can advance codeswitching research, and Muysken 2000 offers a useful typology that also critiques the shortcomings of formalist approaches to codeswitching.

Finally, we can point to the concept of creoles and creolization as a socio-historical phenomenon (Mufwene 2008) that has a language ideological dimension. The idea that abrupt or traumatic language shift results in linguistic phenomena that do not occur in other situations of language change or contact, an idea labeled *creole exceptionalism* (see §3.4.1.2), has engendered a passionate debate over whether *creole* refers to a linguistic type or to a group of languages defined by socio-historical circumstances exhibiting contact-induced changes that do not differ qualitatively from other types of language change (DeGraff 2003, 2004; Bickerton 2004).⁷⁸ The language ideological dimension is manifested in precisely the question of how history relates to language structure. At issue are questions not unlike those surrounding the debate over defining a sprachbund or distinguishing a language from a dialect (cf. Irvine & Gal 2000 on boundaries and also Gal & Irvine 2019). To the extent that the results of creolization are no different from those observed in other contact situations, creole studies are not of any greater

78 Bickerton’s 1984 attention-grabbing assertion that creoles recapitulate a human linguistic “bioprogram” can also be noted here in passing as the ultimate example of creole exceptionalism (see also McWhorter 1998). Data from collections such as Thomason 1997 argue strongly against the notion of creole exceptionalism in its various manifestations, in part via examples where no Indo-European languages – the so-called lexifiers for most creoles – are involved.

significance to the Balkans than other subfields in contact linguistics. Even if, on the other hand, creoles are exceptional, the socio-historical situation of contact in the Balkans was quite different from that surrounding the emergence of creole languages, and we would therefore expect different results. We return to these questions in §3.4.1.2.

At base, then, language ideology in terms of Balkan linguistics is concerned with the establishment or erasure of boundaries, the assignment or creation of social meaning through linguistic activity, and the processes by which these phenomena are constituted. In the chapters that follow, we are sensitive to language ideological issues as they relate to Balkan linguistics, and we address them when appropriate, but the primary focus of our treatment remains the history and present state of the grammars and lexicons themselves.

3.2.2 Assessment Methods for Contact-Induced Change

Discussing the factors that play a role in contact-induced change presupposes knowing that language contact is the source of change in a given case. Thus general concerns about how to assess a particular instance of change must be examined. Here we discuss how to determine whether language contact is involved in a given instance, through a consideration of where innovations come from, and we offer examples of such assessment.

3.2.2.1 Genetic/Genealogical Relationships

A key concern that drives comparative linguistics is what may be called “genetic linguistics” (a term found, for instance, in the title of Thomason & Kaufman 1988), i.e., a focus on understanding the origins of a linguistic phenomenon, whether it involves recognizing lineal descent (‘transmission’, in the sense of Labov 2007) or borrowing/contact (Labov’s ‘diffusion’).⁷⁹ In this enterprise, one first determines which elements can be compared between (Trudgill 2018) various languages, in that they show correspondences (often but not necessarily, similarities) in meanings, in their phonetic forms, in their structural positions in their respective systems, and the like.⁸⁰ The source of the correspondence can then be sought, and one can determine which of four basic causes is responsible for any particular case of the occurrence of comparable elements across languages.

One such cause is mere chance, as in the case of the Nahuatl *teo(tl)* and Latin *deus* both meaning ‘god’ and both sharing, purely accidentally, a phonological structure of

79 Note that the Greek term γενετικός, from which *genetic* is derived, means “pertaining to origins” (i.e., to γένεσις ‘origin’). Some linguists now prefer the term *genealogical*, given the biological (and, in some contexts, racist) implications that *genetic* can summon up.

80 For a strict application of what has come to be called the “Comparative Method,” this statement is an oversimplification, but hopefully a benign one. Comparisons can be made for reasons other than determining the genesis (origin) of a feature; for instance, comparisons for purely typological purposes are possible, by way of classifying languages according to the types of elements or systems they manifest. These two goals – genetic and typological – can be quite different in nature,

a dental stop followed by a front vowel followed by a back vowel.⁸¹ Or, the comparable elements may result from a common “reaction” to a common stimulus, a cause that may be termed “universality,” as with onomatopoeia (where the common stimulus is some natural sound) or other aspects of human existence common to all speakers by virtue of their humanity; for instance, the widespread occurrence of aspirated labial stops in words having to do with blowing, e.g., Armenian *p^huk^h* ‘breath’ and Siouan *p^hu-* ‘blow,’ presumably derives from a linguistic mimicking of the act of blowing. Alternatively, borrowing can be a source of readily comparable items across languages, as is common with names for food items, e.g., ModGrk μουσακάς from the Turkish *musakka* ‘eggplant casserole (usually with meat).’⁸² If these three causes can be ruled out, then one is left with a single final hypothesis concerning the reason for the corresponding elements, namely that they are due to what has come to be called a *genetic* or *genealogical* relationship between the languages in question. “Genetic” here, as noted in footnote 79, has nothing to do with biological genetics or DNA, and refers to descent from a common ancestor language, what is usually referred to as a proto-language, hence the alternative *genealogical*. Languages for which such a relationship can be demonstrated⁸³ are said to be genetically (or genealogically) related, and are presumed to have their origins in differentiation within a common speech community at some point in the past.⁸⁴ The relationship they show after such

leading to different sorts of questions (and thus answers). For instance, in a typological comparison, it might be sufficient to note that two languages, e.g., English and Greek, each have a voiceless labial stop (/p/) and a voiceless labiodental fricative (/f/) in their respective phonological inventories, whereas in a genetically oriented comparison, it is not the sounds alone that are to be compared but rather the occurrence of a given sound in a given morpheme. Thus the genetically interesting fact about /p/ and /f/ in English and Modern Greek is that in words of comparable meaning, English /f/ occurs where Greek /p/ occurs, as in the initial segment of comparanda like *father*/πατέρας or *five*/πέντε (and Greek /f/ occurs where English /b/ occurs, as in *φέρεις* / (you) bear, and so on), not simply that there is a /p/ and there is an /f/. We note in this regard the useful perspective of Hamp 1977a: 279, who says: the “comparative method establishes the descent and genetic (sub)classification of languages and families” and later on the same page uses ‘comparative method’ on a par with ‘areal linguistics’ and ‘typology’ as the “three great categories of linguistic study [...] that rely on the comparison of linguistic features and grammar.”

81 Other examples of this sort can be found in various standard introductory works and textbooks on historical linguistics, e.g., Aitchison 2001: 26; Campbell 1999: 316–317, 322; Hock 1991: 557–559; Hock & Joseph 2019: 399–403, 430–434; Ringe & Eska 2013: 270; and Sihler 2000: 138, 148, among others. We add here one from the Balkans: as noted in §3.2.2.7, the Istro-Romanian word for ‘hand’ is *mâră*, which, totally coincidentally, closely mirrors the word for ‘hand,’ *mara*, in several languages (e.g., Arabana, Doyari, Pitta-Pitta, Mithaka, and Ngamini) of the Karnic subgroup of the Australian Pama-Nyungan family. To this can be added the observations going back to Sapir 1921 of the fact that drift can produce coincidences that have nothing to do with regular sound change or language contact.

82 And the Turkish, in turn, is from the Arabic.

83 See any of the works mentioned in footnote 81, but cf. also Campbell & Poser 2008, on how one might prove genetic relationships for a set of languages.

84 Sometimes the differentiation can be the result of a split whose causation has a known history, as when a group of Gaelic (Irish) speakers moved into the Scottish highlands, and eventually came to be different enough in their language from homeland Irish speakers, thereby constituting a new language, Scots Gaelic. In other cases, it has to be assumed that such a split occurred, due to voluntary migration, forced relocation, geographic disruption, or the like. In yet other instances, differentiation can be the result of the tendency of dialects to grow further apart owing to lack of interaction among speakers, as happened, e.g., on much of Romance, Germanic, Slavic, or Indic territory.

splits can be modeled as a sort of “family tree,” often referred to by the German term *Stammbaum*, in which successive branchings represent successive intermediate proto-languages for later instantiations under a given node in the tree. But at every node, there are other languages interacting with it.

3.2.2.2 Areal Comparisons

Comparativists interested in the genetic relationships between/among languages recognize borrowings – and other contact-induced changes – in order to be able to dismiss them from consideration and thereby to focus on similarities that are a matter of common descent, i.e., inherited from the ancestral form common to the languages in question. For linguists interested in language contact, on the other hand, recognizing the inheritance factor is also critical, but in this case, these are the features that need to be identified in order to be dismissed. That is, one must control for inherited features (and other possibilities) in order to be able to discern possible borrowed ones, and control for borrowed features in order to highlight the inherited ones. The two pursuits – of genetic linguistic connections and of areal/contact-related connections – are not antithetical to one another but rather are complementary. As Campbell 2006: 21 puts it, historical linguists try “to answer the question, ‘what happened,’” and to do so, it is immaterial whether contact or inheritance is the cause for the presence of some feature in a language.

Areal concerns are crucial, since languages in contact with one other, in cases where actual speakers are involved (see §3.1), are co-territorial, occupying the same general geographic area and are thus areally related. They thus display similarities that are due to spatial contiguity rather than descent from a common source. Studying languages in contact thus typically involves making areally based comparisons, whence such terms as “areal linguistics” (for the subfield in general)⁸⁵ or “linguistic area” (for a convergence zone or *sprachbund*).

3.2.2.3 Sources of Innovation

Therefore, given the presence of some feature in a language, be it a sound, a word, a construction, etc., the investigator must entertain two basic possibilities for characterizing its historical source:

- a. the feature has been a part of the language for as far back in time as is relevant
- b. the feature has entered the language at some discernible point in its development.

⁸⁵ This term is given significant establishment in Hickey 2017, a handbook of “areal linguistics.” That volume is a significant resource of information about numerous linguistic areas concerning both well-recognized and newly posited *sprachbunds*; Campbell 2017 includes a compendium of such linguistic areas, seventy-five in all, that have been proposed in the literature; see also Aikhenvald & Dixon 2007.

It is the (a)-type features that are customarily called “inherited” or else referred to as “retentions” from earlier stages of the language, whereas the (b)-type features are “innovations.”⁸⁶

Further, three types of innovations must be distinguished. The first two roughly correspond to Andersen’s 1973 evolutive change vs. adaptive change, respectively (though he uses “adaptive” more for matters of dialect contact):⁸⁷

- (i) internal, i.e., based entirely on existing material in the language and motivated by system-internal pressures within the language itself or by a speaker’s language-using apparatus itself, i.e., elements of the language system, the language production/perception “organs,” and such; examples include the coining of a new word in English such as *chocaholic* from the existing words *chocolate* and *alcoholic*, or new realizations of consonant clusters with *s* in relatively recent English, especially [ʃtr] for canonical [str] (see Janda & Joseph 2003 for discussion and references)
- (ii) external, i.e., via an importation from another language system or through the influence of some other language system, e.g., borrowing a word such as *printer* ‘device for printing material from a computer’ into Bulgarian (as [pɾɪntɐr]) from English
- (iii) mapping of external onto internal, i.e., those cases where the impetus for an innovation comes externally but the material used to make that innovation manifest is already present in the system; this involves elements of both (i) and (ii), as illustrated by calques (loan translations) such as the *sky+scraper* example discussed in §3.2.1.7

Type (iii) requires a bit more discussion, since calques are not the only instantiation of (iii). In some instances, the effects are more overtly a blend, e.g., when

86 There are some problems with this division and terminology, though common sense allows us to continue to use them. For one thing, once an innovation enters a language, it becomes part of the material to be inherited. That is, to the extent that a language is re-created again and again as new speakers (typically children) acquire it, there is inheritance for all features that are not lost or altered, even if they entered the language not long before a given act of language acquisition takes place. Also, the term “inherited” is most typically used for features that can be traced back to an ancestor language separated at some temporal distance from the stage of the language that is the linguist’s object of study. Thus, for instance, one can talk of an element that was inherited from Proto-Germanic into English, such as the word *path* (cf. German *Pfad*) even if it was not inherited from Proto-Indo-European into Germanic (being instead a borrowing into Proto-Germanic, possibly from Iranian).

87 Unlike the authors of such works as Shapiro 1991 and Keller 1994, who take a deterministic/teleological view of language change – essentially looking to the outcome as being what it is because it is what it ought to be – our approach to language change is that (to borrow the *bons mots* used by Hans Henrich Hock) “shi(f)t happens” in language, and we end up with results, some (perhaps all) of which are really just accidents of history, not preordained outcomes. This is not to say, however, that there is no *telos*, no goal, in language change, especially contact-induced change, since convergence typically involves types of feature selection that can be assumed to be aimed at a goal, namely effective communication. Similarly, prestige-driven linguistic borrowing or adoption (whether of pronunciations or of structures or of words) is done with a goal in mind, namely contributing to the status of the borrower. In this way, we diverge from Stolz 2006: 35 who sees Keller’s “invisible hand processes” in “speakers of languages whose communicative practices unintentionally trigger processes of convergence.”

a native morpheme and a functionally parallel external morpheme come to coexist redundantly in a word – a lexical example is the diminutive form of the adverb *λίγα* ‘a little’ in the Greek of Megara (near Corinth), an area in which Greek and Albanian (Arvanitika) speakers coexisted and interacted.⁸⁸ Megaran Greek into the early twentieth century had a form *λιγάζα* that functioned as a diminutive⁸⁹ of *λίγα* ‘a little,’ with the Albanian diminutive suffix *-Vz-* “grafted” onto the Greek stem (cf. Alb *fole* ‘nest,’ *folezë* ‘little nest’), but also a further characterized form *λιγαζάτσι*, showing the addition of a Greek diminutive suffix *-(α)τσι* (cf. Standard Greek *-ακι*, thus with dialectal affricatization) that was isofunctional with the Albanian suffix.⁹⁰

In some cases, too, the blend can have grammatical consequences and be rather subtle, involving the emergence or enhancing of an existing tendency or potential within the affected language due to contact. There is no ready agreed-upon name for the general class of innovations of this sort, but such a situation has been referred to in the literature as an enhancement (e.g., by Porter 1989: 118; see also Aikhenvald 2007a: 22), so we propose calling this situation *contact enhancement*. Although the classification can be difficult to apply in a consistent way,⁹¹ it is a recognized phenomenon. The category of evidentiality, as discussed in §6.2.5, provides a particularly apt example from the Balkans in that the Turkish evidential system was already in place by the time the Ottomans came to the Balkans, whereas medieval Slavic documents provide only hints of usages resembling evidential strategies.⁹² Thus the makings of an evidential system seem to have been present in early East South Slavic (the dialects that would become Macedonian and Bulgarian), and were brought out in a more fully developed way through the influence of Turkish.

Each of these types of sources for features turns out to be relevant for the Balkan shared features that are discussed throughout this work.⁹³

88 Based on Fourikis 1918, discussed in Joseph 1985b.

89 Perhaps, given the meaning of *λίγα*, it functions as an intensifier, further decreasing the “size,” in a sense, of the referent of *λίγα*, already itself a “small” word (a decrease in size being an intensification of smallness).

90 Johanson 2001 develops an account based on social attitudes and ideology for the appearance in Bulgarian of some hybrid forms with a Turkish suffix and a native Bulgarian (though not isofunctional!) one, e.g., *avdžijstvo* ‘huntsmanship, hunting.’ He suggests (p. 179) that in the aftermath of the independence of Bulgaria from the Ottoman Empire, “it was considered particularly important to eradicate the abstract nouns of Turkish origin . . . [and] the suffix *-lâk* was replaced, mainly by the Slavic suffix *-stvo* . . . [leading] to ‘hybrid’ formations” (with a Turkish and a Bulgarian suffix). We can thus speculate that the addition of Greek *-ατσι* onto the “Albanized” form *λιγάζα* might have been ideologically motivated, a “reclaiming” of the Hellenic character of the word, overtly marking it as Greek.

91 The scenario envisioned in such a case is somewhat akin to the claims discussed in §3.2.1.7 concerning some degree of necessary congruence or the like between systems for grammatical borrowing to take place.

92 Van Wijk 1933: 243 cites as possible evidence contrasting uses of the perfect and aorist to evaluate reports in the aorist in the oldest Slavic *Paterikon*.

93 We are deliberately not addressing here the question of the value and meaning of distinguishing between innovation proper and the spread of an innovation, as it is a thorny matter that pertains most directly to what for us is an ancillary issue, namely what constitutes a real change (the initial

3.2.2.4 Yet Another Approach: The Role of Typology

Thus far, we have discussed the genetic (genealogical) and areal approaches to linguistic comparison. A third approach that involves dealing with systemic resemblances that are not mere coincidences is the typological approach. Whereas genetic similarities result from inheritance and areal similarities result from contact, typological similarities are those that do not have either of these sources of causation but rather arise owing to the way human language works.⁹⁴ In the case of such “typological causation,” a given innovation is judged so “natural” (in some often ill-defined but equally often intuitively clear sense), i.e., so much a part of the human language faculty, that it could happen in any language at any time, given the necessary structural environment. In such a case, the correspondence between two languages is not necessarily a matter of inheritance or contact, even if it is found in two genetically or geographically related languages. Comparable elements of this type are usually referred to as “independent” or “parallel” innovations.⁹⁵

3.2.2.5 Putting the Preceding Constructs and Concepts to Use

The taxonomizing of comparisons and innovations from the previous few sections means that the evaluation of putative contact-induced changes must always take place against the backdrop of what is known – or presumed – about inherited features and about linguistic innovations in general.

As a case in point, consider the postposed definite article, a feature in the Balkans which attracted Leake’s 1814 and Kopitar’s 1829 attention (see §3.2.1.7 and §2.2.1, as well as §1.2.1).⁹⁶ This feature, found in Albanian, Balkan Slavic, and Balkan Romance, has a good chance of being a contact-induced phenomenon in the Balkans (though there is controversy as to the exact nature of the contact that led to it) for the simple reason that there is absolutely no basis for positing a definite article in any position for Proto-Indo-European nor in any of the ancestral languages attested in interpretable textual evidence (but cf. Hamp 1982, discussed in §1.2.1). For one thing, there are many Indo-European languages that do not have a definite article, including some of the earliest attested branches, e.g., Anatolian (Hittite, Luvian, etc.) and Indo-Iranian (Sanskrit, etc.), as well as the Baltic languages and most of the modern Slavic languages, although remnants of Common Slavic pronominal definiteness on adjectives does survive a bit. More important, however, in all of the IE languages that have definite articles where early stages of the languages are known, this category appears rather late in the language’s development. Thus, for instance,

innovation or the generalization of an innovation). Our intent, though, is not to make the spread seem somehow irrelevant; see Janda & Joseph 2003 for a consideration of this issue.

94 Thus, this notion is reminiscent of the role of “universals” discussed above in §3.2.2.1 with regard to the application of the Comparative Method.

95 See also footnote 80 for a slightly different application of the notion of “typological comparison.”

96 By *postposed* we mean occurring at the end of the first inflectable element in the noun phrase, so that one might characterize its behavior as “enclitic within the noun phrase.” See §6.1.2.2.1 for more discussion and examples.

the element in Homeric Greek of the eighth century BCE that gave rise to the Classical Greek definite article of fifth century BCE Attic Greek is clearly a demonstrative pronoun in its Homeric use. Similarly, while Classical Latin had no definite article, one has developed in all the Romance languages, again out of a demonstrative, and this same path of development occurred in some Slavic languages. And so on. The absence of a definite article is thus a well-supported feature in the reconstruction of Proto-Indo-European, and the appearance of a definite article, postposed or not, in the Balkan sprachbund cannot be a shared inheritance, i.e., a feature of their common Indo-European heritage. From a methodological perspective, therefore, the postposed definite article stands a chance of being a contact-induced feature in the languages in question.⁹⁷

There are also complicating factors to consider that have a bearing on the ultimate assessment of the source of the postposed definite article and at the same time bring out a further methodological caveat. In particular, there are other languages, and specifically Indo-European languages, that are outside of the Balkans and have innovated a postposed definite article. This is most notably the case for the definite article found in North Germanic languages, beginning as early as Old Norse and continuing into the modern Scandinavian languages (Dahl 2004; Faarlund 2009), but holds as well for Armenian (Mann 1968: 142), and has been claimed for northern Russian dialects, which would be of particular interest since it involves a Slavic language (but see Seliščev 1968: 188–189). These facts complicate the picture since they show that a language need not be in the Balkans to develop a postposed definite article, and also that this development need not involve language contact and could thus emerge in any language at any time.⁹⁸ There is a possibility therefore that the postposed definite article is simply an independent but parallel development in Albanian, Balkan Romance, and Balkan Slavic. The plausibility of such a conclusion is enhanced somewhat by the observation (J. Greenberg 1991; Mladenova 2007; but cf. Sobolev 2009b) that the course of development in the Balkans, from demonstrative to article, is quite common cross-linguistically outside of the Balkans and outside of Indo-European.⁹⁹

97 Definiteness can certainly be marked in other ways in contact situations. Thus in Turkish and modern Indo-Iranian languages (e.g., Modern Persian and Hindi), definite direct objects are marked in a special way, and it is possible that language contact has figured in the emergence of this feature in at least a subset of these languages. In the case of Bulgarian, language planners intervened to bring two different dialectal shapes of the masculine definite article – both *-ăt* and *-ă* – into the standard language and to assign the meaning ‘oblique’ to the latter, although this distinction did not occur in any Bulgarian dialect. Here, the ideology that inflection for case was a desirable feature since it occurred in prestigious languages like Russian and Church Slavonic influenced the codifiers’ choice (see Friedman 2017b; Fielder 2018). Conversely, definiteness can be affected by contact in other ways; northeastern dialects of Romani, for instance, have lost an inherited definite article through contact with speakers of Balto-Slavic languages without such definiteness markers.

98 Suggestions that the postposed definite article in the Balkans is the result of contact with Armenian rest on assumptions that contact with Armenian in the Balkans involved a single, partial metatypy while all other language contact included both full calquing and borrowing.

99 On the other hand, the fact that only in the Balkans did Romance and South Slavic develop postposed definite articles, and moreover, must have done so at approximately the same time, seems more than a coincidence. To this can be added the fact that close contacts between northern

This would mean that there is no reason to suspect that language contact must necessarily be responsible for the appearance of these innovative articles in any of these languages. The occurrence of even a single parallel outside of the Balkans is enough to raise the spectre of independent but (accidentally) parallel development for a given feature as realized in two or more Balkan languages. While some linguists might even have their sense of skepticism about a claim of a common origin for the Balkan facts heightened if more instances of such parallels were to be found, playing a numbers game here does not seem fruitful, for naturalness, we would say, is defined by occurrence in natural languages, pure and simple, not by widespread occurrence. Each case must be examined for its own merits in its own context – after all, parallel contact-induced developments are still contact-induced.

Nonetheless, comparative evidence is relevant whatever the ultimate truth might be for this feature and its status as a significant Balkan feature. Indeed, whether it is substratal in origin or not,¹⁰⁰ a suitable assessment cannot be made without reference to the status of the definite article in Proto-Indo-European, the common ancestor to Albanian, Romance, and Slavic.

Still, though, the fact that the article in these Balkan languages developed as a postposed one, and one that is enclitic (or at least varies positionally) within the noun phrase (see footnote 96 and §6.1.2.2.1), precisely during the period when they came into contact with one another, and, moreover, that a parallel syntactic development did not occur in related languages that also developed definiteness as an overtly marked grammatical category (e.g., Slovene or Czech, or French or Italian) can be justifiably regarded as a degree of similarity too great to be attributable to mere coincidence.¹⁰¹ Such reasoning is similar to argumentation concerning genetic/genealogical relationships among languages in which justifying the positing of a common ancestor for various languages relies ultimately on similarities and especially correspondences being too numerous to be considered to be mere coincidence.

3.2.2.6 Geographical Distribution

One crucial fact remains about the postposed definite article as a putative “Balkanism” (see §3.5) that can strengthen its assessment as a Balkan feature, namely its geography. In particular, the geographical range of the appearance of the definite article in its postposed form in the Balkans is striking, taking in three contiguous representatives of the language groups of the Central Balkans, as is its absence in related languages outside the Balkans. Treating the feature as

East Slavic speakers and Nordic speakers during the Byzantine period might provide an explanation for the independent development of postposed articles on these two Slavic peripheries. The tendency of demonstrative pronouns in colloquial Czech and Slovene to function like articles in the context of intensive contact with German is also worth noting here.

100 See the discussion in §1.2.1 and §6.1.2.2.1.2 on such an account for this feature.

101 Note, too, that even in those languages that developed a postposed article, such as Armenian or Scandinavian (but not North Russian, cf. Zaxarova & Orlova 1970 vs. §6.1.2.2.1), those articles are not enclitic within the noun phrase in the way that the Balkan articles are.

having an independent origin in each of the Balkan languages that show it would be tantamount to saying that its occurrence across these three languages is merely coincidental. While geography by itself is not a compelling argument, it is at least suggestive, even if not probative, essentially allowing for what Campbell et al. 1986: 534 and Campbell 1997: 330–331 have called a “circumstantialist” argument for a contact-induced change.¹⁰²

Similarly, geography plays a key role in the discussion of the confirmative/nonconfirmative opposition in the Balkans, regarding, among other things, the use of the descendants of the inherited perfect in *l* in most of Balkan Slavic and the use of the descendants of the inverted perfect (the admirative) in Albanian. As seen more fully in §6.2.5.1, these uses are found only in the Balkan varieties of these languages (or a subset thereof) but not in the non-Balkan varieties, suggesting a Balkan locus for the feature. Such a conclusion becomes all the more compelling when one considers that there is a possible source language that fits geographically, namely Turkish, and moreover is a co-territorial language with a verbal paradigm (the past marked with the suffix *-mİş*) that displays developments and meanings parallel to what is seen in the Indo-European Balkan languages.

3.2.2.7 Considerations of Chronology

Sometimes not only the geography but also the chronology of a feature and especially of its appearance in the Balkans can be crucial for assessing its status. The case for a contact-induced innovation can therefore be circumstantial, based on geography and chronology, but reasonably compelling nonetheless.

For example, again with regard to confirmativity, it is known that the oldest attested representative of Slavic, namely Old Church Slavonic, does not show any nonconfirmative uses of its *l*-perfect (though see §3.2.2.3 and footnote 92), yet later forms of East South Slavic that came into intensive contact with Turkish do.

In some instances, too, the chronology of particular changes, when mapped onto the *Stammbaum* model for the representation of language relationships, allows for a firm conclusion regarding the status of a possible connection between innovations in different languages. A case in point involves rhotacism in Romanian and Albanian (discussed more fully and from a different perspective in §5.4.4.10.5).

Miklosich 1862: 6–8, in writing about parallel features in the Balkans, noted that both Albanian and Romanian show seemingly parallel instances in which an intervocalic *n* became *r* (hence the label “rhotacism,” e.g., PIE **woinos* > Tosk Albanian *verë* ‘wine,’ Latin *manu-* > I-R *mără* ‘hand’ or Latin *bene* > I-R *bire* ‘well’ (Rmn *bine*). This feature has occasionally been cited by others as a special commonality between Albanian and Romanian.¹⁰³ The suggestion that the changes

102 “Circumstantialist” approaches are contrasted with “historicist” ones, the latter insisting on historical documentation about the diffusion of traits (cf. (§3.2.1.7)). Fortunately, for many features in the Balkans, both approaches can be drawn on. See also §3.4.2.2.

103 See Sandfeld 1930: 125, 126–127 for discussion of this trait, with references to some early opinions on a special link here, both pro and con, to which may be added the positive views of

in these two languages are in some way linked, however, falls apart when one considers the chronology of the respective developments that emerges from Albanian and Romanian dialectology. That is, in Albanian, rhotacism is restricted just to the southern (Tosk) diasystem and is not found in the northern (Geg) diasystem. Since the comparative evidence shows rhotacism to be the innovation – other Indo-European languages have *n* in this position in comparable words (cf. Greek οἶνος, Latin *uinum*, etc. for ‘wine’) – the Stammbaum model would require that the innovation took place after Geg and Tosk split off from Common Albanian, that is, on the way from Common Albanian to Tosk. Similar considerations apply on the Romanian side, where the rhotacism is most regular in Istro-Romanian, but sporadically associated with various dialects/subareas within the rest of Romanian and is absent from South Danubian Balkan Romance (Sala 1970: 56). For Romanian, Rosetti 1924 (see also Rosetti 1985: 268–269) notes some rhotacism in Bucovina and Transylvania and Sala 1970: 56 points only to some scattered instances in Transylvania and northern Moldavia; however, the absence of rhotacism from Slavic loans into Romanian would suggest that to the extent it occurs at all it is an old feature within Romanian (as Sala 1970: 52 notes). Still, it is not pan-Romanian, much less pan-Balkan Romance in any sense, and, as in Albanian, would be assigned in a Stammbaum model to the period after the break-up from Proto-Balkan Romance. Putting those two results together, one is led inescapably to the conclusion that rhotacism in Albanian and rhotacism in Romanian have nothing to do with one another, even if at first glance they would appear to be good candidates for being connected;¹⁰⁴ chronologically, each one happened at a period well after any possible early linkage or contact between these two speech communities.¹⁰⁵

To build a scenario in which the two developments could be connected, one would have to assume that what became (Southern Albanian) Tosk and what we

Hasdeu and Pușcariu (as reported in Sala 1970: 47, footnote 17). More recent views are mixed as well: Sh. Demiraj 2004: 94 denies any connection (see also 1996: 244), as does Rosetti 1985: 268–269, while Stankiewicz 2002: 369 and Banfi 1985: 141 equivocate and Asenova 2002: 27–28 is cautious but somewhat more positive; most positive among recent scholars is Thomason 2001: 108, who gives the feature just a brief mention but includes it nonetheless in a list of “less widespread phonological Balkanisms.”

104 Another instance of a superficial similarity not panning out is discussed by Friedman 1978 in which the apparent similarity of Bulgarian perfects with *e* (3sg copula) and Turkish *mls*-perfects with the copula *-Dir* are shown to fit into their respective systems in very different ways: in Bulgarian, it is the omission of 3sg/PL copular *e/sa* that is the special case while in Turkish it is the addition of *-Dir* that is the special case. See §6.2.5.1 for more details on this.

105 What we have in mind here is the old-looking shared vocabulary between Balkan Romance and Albanian alluded to in §1.2.1.4 (see also §4.2.1.1) that is suggestive of some sort of prehistoric link between the two languages, the exact nature of which, however, is most controversial. If such words point towards a common substratum that fed into Albanian and Balkan Romance, as many believe, then still no solution to the rhotacism issue emerges; if rhotacism were somehow a substratum feature that surfaces in each of the languages (as Miklosich apparently believed), then its restricted distribution in each language (e.g., in Tosk but not Geg within Albanian, and in only part of Romanian and not at all in SDBR) would not be explainable – such a feature, if truly old, should be found throughout all of Albanian and all of Balkan Romance (exception made for secondary obscuring developments).

can for the moment call “Northwest” Romanian were contiguous and were the only parts of the Romano-Albanoid area affected by this change. However, for all other purposes, the connections one finds rather link Tosk and Geg on the one hand and “Northwest” Romanian and the rest of Romanian on the other. Thus the Romanian phenomenon would seem to be best treated as occurring after the break-up of the different varieties of Romanian (which in turn is after Balkan Romance) and thus as rendering it unconnected to the Albanian phenomenon.¹⁰⁶

Therefore, the dialectological distribution and the Stammbaum indicate clearly, decisively, and unambiguously here that these are independent developments, however striking the parallel may be in terms of the details of the change (same input, same output, and same general conditions).

3.2.2.8 Retention versus Innovation

Being able to identify with some degree of certainty the inheritances from a proto-language thus figures heavily in assessing the status of innovative features in a sprachbund, but only indirectly, by providing the baseline from which to judge

¹⁰⁶ Moreover, the evidence of Slavic loanwords vis-à-vis the Albanian/Balkan Romance rhotacism picture is highly relevant. In general, rhotacism in Albanian does not affect loanwords from Slavic (Sh. Demiraj 1996: 239–244) though Çabej 1964: 12, following Jokl, suggests a few such forms (see also Asenova 2002: 38–40), e.g., southern Albanian (from the town of Himarë along the southwestern coast) *tërësirë* ‘cable; rope’ (possibly) from South Slavic *tracina* (itself perhaps from the IE **terk^w* ‘twist’ root, as in Latin *torquere*). Moreover, no Slavic loans in Romanian are affected by rhotacism. Thus, to focus on Albanian, where, unlike Romanian, there is a large dialect area (Tosk) with systematic rhotacism, this phenomenon would most likely date to before the time of the Slavic invasions of the Balkans (sixth to seventh centuries) or somewhat later, when loanwords first began to work their way from Slavic into Albanian. Himarë *tërësirë* would therefore either be a very early borrowing from Slavic, at which point its restriction to the Himarë area would be hard to explain, or would show rather that rhotacism persisted actively somewhat longer in an isolated southwestern corner of Albanian territory. In fact, Janson 1986: 193 on somewhat similar grounds dates Albanian rhotacism roughly to the period between 800 CE and 1000 CE, and even Çabej views it as generally lasting no longer than up to 1000 CE (see Çabej 1979: 56; Sh. Demiraj 1996: 243, 277) for ordinary vocabulary (suggesting it persisted longer for toponyms, inasmuch as some place names now with *-r-* show up with *-n-* even into the fifteenth century, e.g., von Harff in 1496 has *Velona* for modern *Vlorë*; other explanations of the toponyms are possible, though, such as the (re)introduction of earlier forms or names taken from nonrhotacizing dialects or languages, e.g., the form *Avlona* is attested in English into the twentieth century, and Italian still has *Valona*). Hamp’s 1981–1982 explanation of the development of the name *Ohrid* out of the old Greek name for the city, *Λυχνιδός*, rests on assuming the word passed into Albanian in time for the Tosk denasalization, in order to account for *r* from earlier *n*; as to chronology, he believes that either “the borrowing by Albanian must have been early enough to allow the passage of *n* to *r* . . . [or it] must have occurred sufficiently soon thereafter to permit the Slavic *n* to be heard as significantly different phonetically from Common Albanian **nn* (> *n* in both Tosk and Geg), so that the denasalized *r* was provided as the nearest equivalent.” Returning to the matter of Albanian and Balkan Romance regarding rhotacism, while most of these dates for rhotacism might be early enough to make a linkage between Balkan Romance and Albanian possible on this feature, such a chronology runs afoul of the very limited distribution of the change in both Albanian and Balkan Romance respectively; if it were that early and if at that period there was not significant dialect differentiation within each language group, we would expect to see a wider instantiation of rhotacism in the two languages. Thus no matter how one views these developments, even from a substratum perspective (see footnote 105), one is led to the conclusion that each language’s rhotacism is independent of the other’s.

whether a feature is an innovation at all. It might well be asked, therefore, if there is any direct contribution that inheritances can make in and of themselves to the sprachbund. In other words, can shared retentions be significant in a contact situation? Some linguists have responded positively here, and the intriguing hypothesis has been put forward, e.g., by Evans 2003, that in cases of intense language contact based on multi-lateral multilingualism, the presence in the languages of the area in question of features inherited from the proto-language can reinforce the occurrence in each language of inherited material that might otherwise be replaced.¹⁰⁷ In this way, shared retentions can be a matter of contact-induced stability, and we may speak of “conservation areas” thereby constituted.¹⁰⁸ To some extent, such a scenario seems quite reasonable, since speakers, in the usual case, do not know if a particular feature of theirs that matches a feature in a neighboring language they are in contact with is a shared retention or a shared innovation; all they are aware of is the noticeable commonality between the two languages.¹⁰⁹ This is in keeping with a truism about language contact that figures in the discussion in later chapters: Language contact typically involves the transfer of elements of “surface” structure, i.e., the more “tangible” aspects, and not the “deep” and more abstract elements of the grammar.¹¹⁰

3.2.2.9 The Relevance of Dialects

As the discussion of Balkan Romance (especially Romanian) and Albanian dialectology shows, within a general discussion of methodology, it is necessary to return also to a theme noted in the Introduction, and mentioned again in Chapter 1, with its discussion of standards and norms, and above in §3.1 as well, namely the importance of considering varieties of the languages other than the standard languages.

Standard languages typically are constructs that do not necessarily reflect the everyday colloquial speech of any individual (cf. J. Joseph 1987), the speech style most relevant to and characteristic of intimate contact between speakers. It is important to keep in mind the words of Bailey et al. 1989: 299:¹¹¹

107 While Hamp 1977a argues that in areal linguistics, as in genealogical linguistics, only shared innovations are diagnostic, this does not eliminate the value of shared archaisms if they are retained precisely and only in the contact area.

108 See §5.2 for an example of a contact-based retention in phonology, and §6.2.2.1 for an example from morphosyntax.

109 We note, for instance, that speakers, when asked about dialects other than their own, will often venture observations about similarities (“our word for X is the same as their word for X”; “they say X the same way we do,” etc.), and will base folk judgments as to the relatedness of dialects on degrees of similarity of certain salient elements, which may or may not be shared innovations, whereas linguists typically look to significant and identifiable shared innovations as the mainstay of determining Stammbaum-like dialect relatedness.

110 In a sense, then, the situation in contact is WYHIWYG (“What you hear is what you get”), as the earlier discussion about the role of surface structure (§3.2.1.7) indicates. Mitchell 2015 is an experimental study of how what is heard can matter for contact-induced phonological change.

111 And to this we can add the classic statement of Henry Sweet 1900: 51: “The real life of language is better seen in dialects and colloquial forms of speech.”

... [T]he history of ... language is the history of vernaculars rather than standard languages. Present-day vernaculars evolved from earlier ones that differed remarkably from present-day textbook[-varieties] ... These earlier vernaculars, rather than the standard, clearly must be ... the focus of research into the history of ... [languages].

The value of this important injunction can be seen forcefully in the Balkans through a few examples.

For instance, in Standard Modern Greek the mid-vowels /o, e/ are stable when unstressed and remain as mid-vowels. A similar situation is found in Standard Bulgarian and Standard Macedonian, so that it might well seem that there is nothing interesting or worthy of examination in the vowel systems of these languages other than noting the occurrence of certain segments. However, in eastern Bulgarian dialects and also in the colloquial usage of speakers of the standard language, a somewhat different situation is found, with the unstressed mid-vowels /o, e/ being raised somewhat in the direction of [u, i] respectively (more robustly the farther east one goes and more for the rounded back vowel). The actual phonetics of this development may well involve, in the case of /o/, increased lip-rounding which then has the effect of increasing the gravity¹¹² of the vowel and thus making it perceptually more like an [u], and thus somewhat raised. Raising is also characteristic of southeastern Macedonian unstressed mid-vowels. Still, this development allows for an intriguing comparison with Greek to be made, since the raising of unstressed mid-vowels does in fact occur in Greek; just as in Bulgarian and some southern dialects of Macedonian, though in Greek, too, it is dialectally restricted, occurring generally only in the northern dialects of Greek, and not in the standard language, which historically is largely based on a southern dialect (that of the Peloponnesos). Moreover, we know that the mid-vowel raising is a relatively recent development in northern Greek, since it came later than another vowel change, the loss of original unstressed high vowels /i, u/, that sets the north off from the south within the Greek-speaking world; that is to say, the raised mid-vowels were generally not subject to the loss that original high vowels were.¹¹³ Since the northern dialects of Greek, of course, are geographically contiguous or co-territorial with Bulgarian, and southern Macedonian and have been for more than a millennium, the possibility of linking the Northern Greek raising phenomenon with that in Bulgarian and Macedonian dialects must be considered, and it is conceivable that the mid-vowel raising is a contact-related phenomenon in the eastern central Balkans (see §5.4.1.5 for further

112 “Gravity” in the sense of the acoustically based feature [±grave] (on which see Jakobson et al. 1967).

113 For instance, corresponding to Standard Greek *παιμῆναι*, a form which represents well the Proto-Modern-Greek form, northern dialects have [pirmén], with loss of the unstressed [i] of the second and fourth syllables, and raising (but with no subsequent loss) of the unstressed [e] of the initial syllable to [i]. The high vowel loss thus preceded the mid-vowel raising, since otherwise the raised vowels would be expected to have been lost too. There are occasional instances of raised mid-vowels being deleted – e.g., as Newton 1972: 190 reports, [bði] occurs for Standard Greek *παιδί* ([pæði]) ‘child’ and for the more usual northern form [piði], possibly due to relexification of the word as if it had an underlying high vowel which could then be subject to loss by the synchronically active rule of high-vowel deletion.

discussion). What is relevant here is that if one were to look solely at the standard languages, an interesting and potentially significant Balkan parallel in vowel developments would be missed entirely.¹¹⁴

A second such case in which comparisons involving only a standard language can blind the researcher to important aspects of a contact situation in the Balkans is the fate of certain Turkish suffixes in Bulgarian, as discussed by Johanson 2001. As becomes evident, considerations of chronology are also crucial here. Johanson's focus is the Turkish abstract noun-forming suffix *-IVK*,¹¹⁵ which normally shows up in Bulgarian on both native and copied stems in a uniform shape, *-lāk*, e.g., *terzilāk* 'tailoring' (from Trk *terzilik* 'idem'), *vojnīklāk* 'military service' (from Blg *vojnik* 'soldier'). So Bulgarian usually has no reflection of Turkish vowel harmony alternations, although it does copy some Turkish words complete with Turkish vowel harmony, e.g., *kusurluk* 'deficiency' (cf. Turkish *kusurluk*). Noting, quite insightfully (p. 178), that in the quest to find the source of a borrowed item,¹¹⁶ "the historical linguist must try to avoid anachronism," Johanson goes on to label as "a fallacy" – one all too easily accepted by linguists looking at Turkish elements in Bulgarian – "the silent assumption that the phonetic shapes of the Turkish originals are more or less the same as in modern Turkish." In fact, as Johanson points out, four-fold vowel harmony in this suffix, with *-IVK* surfacing as *-lik/-lük/-līk/-luk*, is an innovation of post-seventeenth century Istanbul Turkish that Turkish dialects of Western Bulgaria did not share. Thus, the invariant *-lāk* in Bulgarian now most likely reflects a Western Bulgarian Turkish (i.e., WRT) source, or else a pre-seventeenth-century chronology (or both). Again, therefore, relying on facts from modern Istanbul usage, as a reflection of Standard Turkish, and generalizing from there to all Turkish both diatopically and diachronically, is simply misleading, and in this case obscures the actual historical situation for Bulgarian *-lāk*.

Let us turn now briefly to a relevant aspect of the external history of our discipline. It was precisely at the end of the eighteenth and beginning of the nineteenth centuries, when Western Europe first began taking notice of the Balkans and their languages in modern times, that speakers of those languages were first beginning the endeavors that would lead to the end of the political and social conditions that had created the Balkan sprachbund as such in the first place, i.e., the creation of Balkan nation states and their standardized national languages. One of the problems of much of modern Balkan linguistics, especially in recent decades, is that it relies on data from these standards.¹¹⁷ It is thus using

114 Thomason 2001: 108 calls vowel raising a "less widespread phonological Balkanism" but does not mention Northern Greek or southeastern Macedonian in this context, but only Albanian, Bulgarian, and Romanian in her list of languages with raising. See §5.4.1.5 (and footnotes 45, 46, 47) and §5.4.3.9 for more details on mid-vowel raising.

115 This is Johanson's notation, an abstract form with the "capital letters indicating morphophonological variation and V standing for an unspecified vowel segment" (p. 177).

116 In Johanson's terms, a "copied" item; see §3.2.1.7 and footnote 61.

117 As Joseph 1987a points out, a significant problem with Schaller 1975 is that he "seems constantly to be treating the regional dialects as somehow being beside the point, though his treatment of the southeastern Serbian (*Torlak*) dialects is a noteworthy exception, for he generally refers to the properties of the written standard languages in discussing common Balkan features." Asenova

dialectological methodology on standardized data. Such an approach is not without merit, insofar as “Balkanization” has penetrated so deeply into the total structure of all the relevant diasystems that even a standardized one will reflect this (e.g., Aronson 1981, 1994; Kramer 1988, 1995; Nikolaeva 1996), but in many instances the standards have purposefully, sometimes for ideological reasons, excluded specifically Balkan features as low style, sub-standard, or dialectal. The lexical example of Turkisms is the most obvious case in point (Kazazis 1972, and see §4.4), but a variety of other features can be included here as well. Thus, for example, object reduplication (see §7.5.1), which is a typical Balkan syntactic feature, is subject to greater constraints in literary Bulgarian than in colloquial Bulgarian (Friedman 1994a), and the leveling of relative pronouns – another Balkan syntactic feature (see §7.7.2.2.1) – is greater in Macedonian dialects than in the literary language (Topolińska 1995a).

3.2.2.10 Dialects and Diffusion

Attention to regional and nonstandard dialects is therefore critical in making the most salient and fruitful comparisons in the Balkans. It plays a role as well in a consideration of how innovations spread throughout a language, for in the typical case, an innovation originates in a localized part of a speech community and spreads from there, if it spreads at all.¹¹⁸ As a feature spreads, it becomes a differentiating characteristic between those speakers that have adopted it and those that have not, so that the spread of a feature can become a basis for creating new dialectal divisions (*sensu lato*) or reinforcing and maintaining old ones. Moreover, since, as argued previously (§3.1), there is no essential difference other than degree between dialect contact and language contact, understanding the mechanisms for the spread of a feature within a language provides insights into the nature of linguistic diffusion more generally in a contact situation, whether contact between speakers of different languages or contact between speakers of different dialects is involved.

The dimensions along which the spread takes place can be quite varied. It is often the case that the spread is over a contiguous area geographically, so that dialectal and geographic factors come together. In such cases, where the spread radiates out from an original locus, one typically finds – an important notion in linguistic geography due to Bartoli 1925 – an innovating core where the innovation has been adopted as opposed to a conservative periphery. A good example of this phenomenon at work in the Balkans is provided by the Greek-internal distribution

2002 also relies mostly on standard language data. On the other hand, various post-2000 works by Domoselickaja, Kisilier, Makartsev, Markovikj, Morozova, Rusakov, Sobolev, and many others cited in the bibliography have taken up the call to pay attention to dialects.

118 We say this since not all innovations manage to spread and become generalized outside of their original locus; historical linguists tend to look only at the ones that have had some “robustness” in that sense, but undoubtedly dozens of novel utterances and pronunciations arise in a speech community every day, even if they might be only marginally and subtly different from other prevailing forms.

of the replacement of the earlier infinitive by finite clauses. While discussed more fully in §7.7.2, briefly put, this feature involved the innovative substitution of a fully finite (person/number-marked) verb for an infinitive in all of its uses, including complementation and nominalization; compare, for instance, the earlier (3.2a), from the *Chronicle of Morea* (fourteenth-century version) with an infinitival complement to ἀρχίζω ‘begin,’ with the later (3.2b) from the same text but a different manuscript version (fifteenth-century) and with a finite clausal complement:¹¹⁹

- (3.2) a. ἀρχασ-αν ... καὶ λέγ-ειν (l. 5261 (Ms. H))
 began-3PL and say-INF
 ‘they began (even) to say’
 b. ἀρχισ-αν ... να λέγ-ουν (l. 5261 (Ms. P))
 began-3PL SBJV say-3PL
 ‘they began to say’

This innovative extended use of finite complementation is found throughout the Greek-speaking world starting in the Hellenistic period, and it diffused slowly throughout the grammar of Greek, affecting some constructions and some lexical items before others. With regard to geographic diffusion, however, what is interesting about this development is that it is found most robustly in the Greek spoken in the Balkans and much less so in outlying dialects such as Pontic Greek (of Asia Minor) and the Greek of southern Italy; the distribution, therefore, adheres to the previously mentioned generalization from linguistic geography and can be viewed in terms of an innovating core in the Balkans with peripheral relic areas.

It is possible, as well though, for features to spread to noncontiguous regions, especially when discontinuous urban areas are involved. This type of spread is referred to as the “gravity” model by Chambers & Trudgill (1980) and as the “parachuting” model by Joseph (2000a), and actually has a venerable history within traditional dialectology. The basis for the spread would be movement of people from urban center to urban center, presumably for trade, business, and work. If the “sphere of influence” of urban centers grows in an area, the feature then presumably spreads to more outlying parts, filling in the gaps, as it were, between the various cities. While this model can predict results different from what might emerge from the claims concerning an innovating core versus conservative periphery, especially if a particular urban area would not fall into the “core” region, the two approaches can be seen as generally compatible. That is, the “gravity” model can be taken to represent the early stages of the spread, before a whole region – cities and intervening countryside – adopts the feature.

It is quite likely therefore that this pattern of spread was actually the first step in the geographic diffusion of the replacement of the infinitive within Greek,¹²⁰

119 The slight difference in the shape of the main verb – ἀρχάζω vs. ἀρχίζω – is not relevant to the point at issue.

120 We focus on Greek here, even though this is a feature that is found throughout the Balkans, as discussed more fully in §7.7.2.1.1.

especially based on the scenario for the origins of this feature suggested by Joseph 1983a: chapter 7. There, it is argued that the geographic locus of diffusion for this feature within Greek was northern Greece, and especially the multilingual urban center of Thessaloniki, for it was there that speakers of Greek, Slavic (i.e., dialects that would eventually differentiate into Macedonian and Bulgarian), Albanian, and Romance would have come in sustained and close contact with one another on a daily basis, providing opportunities for the sorts of effects that could have contributed to this innovation catching on in any of the languages, and thus to convergence among these languages in general. This is the exact social context for contact that Thomason & Kaufman 1988 (cf. also Aikhenvald 2007a) have pointed to as essential to the development of a *sprachbund*, namely with the relevant speech communities each maintaining their own linguistic identity in spite of the extensive and intimate contact, and thus some members of the groups of necessity being bi- or multilingual. Similar conditions are also evident in smaller *sprachbund* situations such as the relatively small one of Kupwar village in India, as described by Gumperz & Wilson 1971, where intense contact with multilingualism has led to structural convergences among Urdu, Kannada, and Marathi.

The details concerning this claim, drawing in part on Joseph 2000a, are presented in §7.7.2, but some important consequences of this scenario are relevant in the context of the present discussion.

For one thing, it would mean that there was robust spread of this innovation to other parts of the Greek world. If the change spread first, in “gravity model” fashion, to areas with large cities, then even outlying urban areas, e.g., Venice and those in Cyprus – the former having “a large Greek community [and being] an important centre of Greek commercial, religious and cultural activity” during the period of Ottoman rule in Greece (Clogg 1992: 16) – would be expected to be affected by the replacement of the infinitive to the fullest degree. And indeed, such is the case. This, therefore, is an area where the “gravity model” and the innovative-core/conservative-periphery make different predictions, or at least need to be seen as presenting different stages of – or perspectives on – the diffusion process. Within Greece, then, it can be posited that the replacement of the infinitive spread from Thessaloniki, the largest city at that time in what is now Greece, to Athens, which was not then the urban center it is today but in the medieval period was perhaps the second largest city in Greece, and the site of an orthodox metropolitan see, to other sizable cities such as Árgos (in the Peloponnesos) and Candia (i.e., Heráklion, in Crete), which were both on major medieval sea-trade routes, as were also Venice, Thessaloniki, and Árgos-Náupaktos, to judge from the maps in Magocsi 1993, and only after that to the intervening, more rural, areas, some of which had orthodox metropolitan sees, e.g., Lárissa and Ioánnina. The more conservative nature of certain parts of the (geographic) periphery can also be understood in this model, as there would have been less robust spread to peripheral areas without urban centers, e.g., southern Italy (Griko) and the Black Sea coast (i.e., Pontic Greek); medieval Cyprus, however, being part of a major maritime trade network, would not have

been peripheral in this regard. Quintana Rodríguez 2006 makes similar points for innovations in Balkan Judezmo.

Especially interesting is the situation in the largest Greek city of the relevant period, Constantinople (modern Istanbul). For the most part, Constantinople Greek, at least in its colloquial variety in medieval times, was not particularly different from other northern dialects of Greek spoken at that time, and this holds as well for the use of the infinitive, which, by the 1400s, would have been highly restricted and limited to a few contexts, most notably in a future tense formation. This situation is in keeping with the model of spread of the infinitive replacement innovation by parachuting into urban centers, and it should be noted as well that many Greeks left Constantinople, e.g., for Venice, with the coming of the Ottoman Turks in the fourteenth and fifteenth centuries, and that exodus would have provided another path for the spread of the innovation.

The models drawn on in this discussion illustrating spread are aimed at macro-aspects of the phenomenon, covering large stretches of territory, whether continuous or discontinuous. There is still the question, though, of what happens at a more local level, in the actual speaker-to-speaker encounters, interaction, and exposure that lead to the adoption by an individual of linguistic features not originally part of his/her linguistic repertoire. Here the work of the Milroys (especially J. Milroy & L. Milroy 1985, L. Milroy & J. Milroy 1992) and their emphasis on social networks has provided a significant basis for understanding the pathways for the spread and adoption of an innovation. The social network is the set of people that one interacts with on a regular basis, and who, in the typical case, interact with one another as well, thus forming a web of interacting entities. The strength of the connections among those in the network and the density of the network, i.e., its degree of inter-connectedness, are measures of its robustness and thus affect the extent to which speakers adopt new features from others in the network or have aspects of their own usage reinforced by occurrence in the speech of others in the network. It is thus a truism that whom speakers interact with has a profound effect on the choices they make as to how to shape their usage.

The developments with the infinitive in Greek, as it happens, offer another opportunity to see the likely effects of social network interaction in action. As noted above, Constantinople Greek in the Middle Greek period seems in all respects to be like other northern dialects especially with regard to the use of finite clauses in place of earlier infinitives. There is, however, one exception to this statement, discussed in Joseph 2000a, namely the usage found in a 1547 translation of the Hebrew Bible into Greek that was written in Constantinople by a Greek-speaking Jew (using Hebrew characters).

This work was a didactic work, designed to help Jews learn the Hebrew of the Bible via Greek, their native language, and it seems by all accounts (e.g., Bellili 1890; Hesseling 1897; Joseph 2019a) to reflect sixteenth-century colloquial spoken Greek of Constantinople. For one thing, it represents some fine phonetic detail (e.g., /mj/ is written as <mnj>, presumably representing something like [mj] – see §5.2) that is characteristic of the spoken Greek of that period. Yet, despite the colloquial character of the translation, infinitives in this text occur in greater

numbers and in a wider range of uses than can generally be found in Greek in that period, including nominalized infinitives as the object of prepositions, infinitives as complements to perception verbs, a future formation with $\theta\acute{\epsilon}\lambda\omega$, and an entirely novel use in a rendering of a Hebraism (e.g., “And God spoke, saying . . .”).¹²¹ Except for the Hebraism, these uses themselves are not surprising in the overall historical context of Greek, in that they are found in earlier stages of Greek, but they are unusual for colloquial Greek of the sixteenth century. In fact, the infinitival usage in this text seems to be somewhat conservative in nature compared to the rest of colloquial Greek of the period, and recall that the attention to phonetic detail in the spelling suggests that this too is indeed colloquial Greek.

Thus, the evidence of this text, taken at face value, would suggest that Constantinople was a peripheral area with regard to the spread of the innovative replacement of the infinitive, even though the preeminence of Constantinople as the leading Greek urban center in that period would suggest that Greek speakers there would have participated fully in the infinitive replacement process, as they indeed seem to have outside of this text.

A solution to this dilemma concerning conservative infinitival usage in the text is to be found in social network theory, coupled with the fact that the translation into Greek was made by a member of the Jewish Greek community of Constantinople, someone who apparently spoke Greek natively but did not have any knowledge of the Classical or Biblical Greek traditions; the translation was from the Hebrew, for instance, not from the Greek Septuagint. That is, it can be hypothesized that the Jewish Greek speech of Constantinople was distinguished from the Greek of Orthodox Christians and that the differences in infinitival usage evident here are a function of the text reflecting an apparently more conservative Jewish Constantinople Greek as opposed to the more innovative and more mainstream Orthodox Christian Constantinople Greek. Indeed, as Wexler 1981: 102, footnote 5 has noted, and see also Marcos 2000: 184, it is often claimed that Jewish languages in general tend to be conservative (though Wexler doubts this is always so), and indeed, the segregation of Jewish communities would certainly have created situations in which Jewish speakers might have had less access to linguistic innovations found in the usage of co-territorial non-Jewish speakers. Here is where social network theory comes into play, for the relative isolation of the Jewish community would have meant that Jews’ main interactions – the building blocks of their main social network – would have been with other Jews, and not with the Christian Greeks. Access to the innovative usage, therefore, would not have been straightforward and the use of the innovation would not have been reinforced by members of the social network, leading thus to conservatism. But the conservatism may also have been a purposeful archaizing tendency chosen to separate the Holy Text from the secular with a known, but archaic, feature, as with usages of *thou* in Modern English. Then, as now, observant Jews were especially concerned with maintaining boundaries. And infinitive usage may well have been emblematic at the time, while phonological changes went unnoticed.

121 See Joseph 2000a, 2019a for examples and further discussion.

As indicated above, there may well have been more at issue here than undefined “social networks,” for it seems likely that religious group identification could have played a role too. In this context, two typological parallels for a religion-based linguistic identification involving the infinitive in Greece are worth mentioning. For example, Judezmo (Judeo-Spanish) of Thessaloniki still showed an infinitive into the late twentieth century (Joseph 1983a: 252ff.), and even now, despite the fact that its (Jewish) speakers are (now, at least) bilingual in infinitive-less Standard Greek and are (now, at least) in constant contact with monolingual speakers of Standard Greek, which is (today) by legal definition in Greece, Orthodox Christian Greek; preserving an infinitive in Judezmo can be taken as a reflection of maintaining boundaries based on religion but using linguistic material as the basis for differentiation from some other group. A converse sort of parallel is seen with regard to Tsakonian, of the Peloponnesos. Tsakonian has no infinitive, despite stemming from a different source from the rest of the Modern Greek world, deriving instead directly from the ancient Doric dialect, which had the infinitive quite productively, rather than the Hellenistic Koine, in which the infinitive was beginning to recede; presumably, Tsakonian has lost its infinitive through contact with Standard Greek (cf. Kisilier 2021 on Greek-Tsakonian contact) and while it is important that most Tsakonian speakers also speak Standard Greek, more relevant is the fact that Tsakonians are virtually all Greek Orthodox and moreover show “allegiance to the greater Greek culture” (Vlamiš 1996). Thus, the sociologically relevant factors of religion and degree of assimilation into mainstream Greek culture and society seem to have had linguistic consequences for Tsakonian and its loss of the infinitive paralleling that of Standard Modern Greek.

What this extended example shows is that at the local level as well, consideration must also be given to the factors leading a particular speaker to adopt – or reject – a given innovation presented to him/her by another speaker. These factors include considerations of prestige and utility associated with the language or dialect with the innovative form, as well as accommodation to the usage of others (wanting to facilitate communication with them, to put them at ease, to express solidarity with them, to not put them down or on the defensive, etc.). All of these factors can influence a speaker’s choice. Moreover, especially regarding the rejection of innovations, as suggested by the Judezmo case, there is also the issue of using language to differentiate oneself from others, what may be called contrastive differentiation or linguistic resistance or boundary marking.¹²²

122 For instance, a Greek who did not want to sound like a Turk could actively change a Turkish *ç* in a word to a *ts* and thus say τσάι ([tsaj]) instead of çay ([tʃaj]) for ‘tea.’ While one might be tempted to say that because (certain dialects of) Greek did not (as the standard language today does not) have a palatal [tʃ], Greek speakers could only deal with Turkish *ç* by rendering it as a *ts*, it is rather the case that speakers can change their speech habits, or at least can exert some degree of conscious control over how they pronounce foreign words, and that the degree of conscious control is even greater when it comes to choosing which foreign words to borrow at all, and which to pronounce in the manner of a foreign speaker. Relevant here too is Matras’s 2009: 225, 228 notion of “authentication,” as discussed in §3.2.1.3; that is, a speaker’s control can be a mechanism both for avoiding certain sounds, and thus adapting a loanword, and for adopting certain sounds too, even if foreign.

To a large extent, these cases are basically the same: innovations are evaluated by speakers encountering them for the first time, and choices are made as to the viability of that innovation with regard to the speakers as individuals and as group members. Group membership and identity are expressed, among other ways, through linguistic usage. Even the use of economically useful languages can be viewed in this light, inasmuch as the marketplace (in the broadest sense) or audience one chooses to go to or aim at says something about the person as an individual, whom s/he shops with, what types of goods s/he trades in, and more generally who s/he is.

Thus numerous factors can play a role in the resolution of contact situations.

3.3 Methodological Consequences of the “Speaker-Plus-Dialect” Approach

All the foregoing considerations and discussion add up to, as essentially a natural extension of our focus on speakers in §3.1, what we can call a “speaker-plus-dialect” approach to language contact, especially in multilingual situations. Although geography is relevant in contact among speakers of different languages, it is not the sole determinant;¹²³ as Campbell 2006: 16 points out:

[diffused] shared linguistic traits are not brought into existence by, nor somehow explained by, the geographical region, in spite of the fact that the notion ‘linguistic area’ is often presented, at least implicitly, as some entity where the geography is prime and the linguistic traits themselves are just reflections of some sort of vague geographic determinism. There is no geographical determinism; the linguistic borrowings are prime, and the geographical areas are only a reflection of these, with no significant causal force of their own.

To which we can add: depending on social conditions. We can thus go one step further and say that it is not the borrowings, *per se*, that are prime, but the speakers who carry out the borrowings. That is, speakers are central to the understanding of language contact; moreover, the contact takes place at the dialectal level in its origins: this is the crucial locus of contact. Hence: a “speaker-plus-dialect” approach.

While focusing on speakers and dialects may seem to be an obvious step to take, it is an important one, and it has certain key consequences for methodology, some of which are evident from previous discussion but all of which bear summarizing here. First, it means that one has to take a localized approach; contact takes place at the local level, involving regional dialects and not standard

123 Campbell 2006: 14 notes that some scholars go so far as to say that geography is not necessarily relevant at all. Dahl 2001: 1460 observes that the notion that “each language has a specific location in space, that no more than one language is spoken in each place, and that language contact takes place between [spatially] adjacent languages” is simply “erroneous.” We give some credence to this view: after all, we refer in §3.1 and §3.2.1.1 to contact potentially taking place within the bilingual mind, thus affecting in such a case languages that are psychologically, but not spatially, adjacent. Still, we take issue with Dahl’s formulation, to some extent, since whether discrete or not, languages are necessarily located in some space and language contact involving speakers (as opposed to learned or literature-based influences which need not involve speaker contact – see §3.1) necessarily occurs in some place.

languages (see §3.2.2.9). Second, local dialect details matter (again, see §3.2.2.9) and thus the dialectology of small areas and with it the examination of fine-grained difference take on particular significance. Such a localized dialectological view leads, third, to looking at clusters of features in concentrated locales, what Hamp 1989a: 44 has called “more restricted coherences” that can overlap one another with regard to particular features but which add up to a wider geographic domain overall. Fourth and finally, then, the sprachbund becomes not a far-flung area defined by some set of features shared over the entire space, but rather a collection of these restricted coherent groupings of locales each with clusters of features. In a sense, as Hamp’s very title, “A crossroads of Sprachbünde” suggests, each concentrated locale is in principle a sort of sprachbund, one of several “linguistic unities,” as Hamp characterizes these “coherences.”¹²⁴

For the Balkans, according to Hamp, these overlapping clusters include the regions he labels the Carpathian arc, the eastern Balkans, the Pannonian area, the eastern Alps, the Dalmatian Adriatic, Illyria, the western middle Balkans, and the Aegean. He states further (p. 47) that “we are dealing here with a spectrum of differential bindings, a spectrum that extends in different densities across the whole of Europe and beyond.” But it starts with diffusion in small areas, and these in turn are built on speaker-to-speaker interactions involving bi- and multilingualism on the part of individuals. Hence: a “speaker-plus-dialect” approach.

The importance of looking to localized convergences in understanding areal features and diffusion has been noted by others. Matisoff 2001: 300 asks, rhetorically, “Does not every ‘linguistic area’ arise from an accumulation of individual cases of ‘localized diffusion,’” and this is a sentiment which Campbell 2006 not only cites (p. 10) but also echoes. Moreover, he endorses it by repeating (p. 18) Matisoff’s dictum about the importance of “localized diffusion.” In this way, too, he echoes Koptjevskaja-Tamm, whom he also cites (p. 16) and who makes reference (2002: 219) to “micro-contacts” as the source of linguistic diffusion. And long before these scholars, Jakobson 1931a/1962, 1931b/1962, 1938/1962, in extending Trubetzkoy’s notion of a sprachbund, envisioned an extensive Eurasian expanse of territory showing convergence essentially via a “chain” of overlapping smaller areas of convergence (cf. §3.4.1.3 on Eurology and also Hamp 1989a on differential bindings). Thus our “speaker-plus-dialect” approach simply gives a name to what is a well-recognized truism about language contact, linguistic diffusion, and linguistic areas. It is arguably the only realistic way to approach these phenomena. Accordingly, in the presentation in later chapters of the convergent features in the Balkans, we pay particular attention to the local level of analysis and local patterns of diffusion, for they are the key to understanding the Balkan sprachbund.

There is another reason for focusing in such a detailed and fine-grained way in an analysis of Balkan convergences. There are several scholars who have eschewed

124 See §3.4 and §8.2 for more definitional discussion regarding the notion of “sprachbund.” Relevant too is the discussion in §3.4.1.2 on the relationship between the Balkans and a putative European linguistic area. See Stolz & Levkovych 2017 for a discussion of the phonology of Europe in an areal perspective.

fine granularity and instead have taken a coarser approach to analyzing the Balkan sprachbund, boiling the relations among the languages down to checklists recording the presence or absence of various features, *grosso modo*. We would argue, to the contrary, that once the importance of speaker-to-speaker interactions in highly localized spheres is recognized, an approach to understanding the Balkan sprachbund in terms of such gross tabulations, as opposed to a finely nuanced appreciation of differences and similarities, can be found to be wanting.¹²⁵

What we have in mind is the several typological and other studies, e.g., Haspelmath 1998, van der Auwera 1998, and Lindstedt 2000, but also earlier works such as Campbell et al. 1986 and Reiter 1994 (and, to some extent, Schaller 1975), since they all treat complex Balkan phenomena as unitary, ticking off points or drawing lines in such a way that the facts “on the ground” disappear from view.

More particularly, they attempt to “weight” or quantify Balkanisms, and depend almost exclusively either on standard language data, or in some cases, oversimplified dialect regions. Their tallies come out skewed in part because they do not take dialectal data into account (see Friedman 2008a for details). For instance, in van der Auwera’s count, Bulgarian comes out as “more” Balkan than Macedonian owing to the presence of stressed schwa (assuming that one even wants to accept this as a Balkanism, see §§5.4.1.1, 5.4.1.6), yet the majority of Macedonian dialects do have stressed schwa even though the standard language (on which he based his enumeration) does not; interestingly, on the Bulgarian side, the standard language does have stressed schwa, while only a minority of Bulgarian dialects (Teteven-Erkeč, some Rhodopian) do not. Similarly, analytic accusative marking is typical of Balkan Romance (via use of the direct object marker *pe*) but, while absent from the standard Balkan Slavic languages, it is also found in the western Macedonian dialects (with *na* as the relevant marker) in close contact with Aromanian. And, Aromanian itself is often missing from more superficial Balkan linguistic accounts, which are satisfied with Romanian as the single representative of Balkan Romance (with the Moldavian dialects almost never figuring in any accounts). And within Aromanian, the dialects of the northwest of Aromanian territory, which is in the Balkan heartland, such as Bela di Supră (Mac Gorna Belica) are considerably more Balkanized (e.g., in possessing distinct evidentials and more simplified nominals) than those of other areas. Thus, as seen in §3.2.2.9, dialect evidence is critical for getting the fullest picture of the effects of Balkan language contact.

More generally, any numerologically based assessment is only as good as the input to the tally is. Thus, when Campbell et al. 1986: 561 evaluate Romanian as the most Balkan language and Macedonian as “lack[ing] several of the areal traits,” their assessment has to be viewed as flawed since it is based on an incomplete assembling of the relevant Macedonian data. A more detailed look reveals a very different picture. In particular, Macedonian has a fully grammatically embedded perfect in ‘have’ (while Standard Bulgarian does not), it uses the same prepositions

125 See also §2.4 and especially footnote 30 for an additional point of criticism regarding such approaches.

and adverbs for both location and direction, and it has genitive/dative syncretism, all traits for which it was not given appropriate credit in the Campbell et al. enumeration. Further, one Balkan feature that they do count, the disparate vocalic alternations in Romanian, Bulgarian, which they call “Vowel harmony (or umlaut),” not only is misnamed and mislocated (Greek, *pace* the initial sentence in their presentation of the feature, shows no such effects) but also, as a result, does not really make for a compelling Balkanism (see §5.4.3.7 for a detailed discussion).

Moreover, it is not clear what “presence” of a feature actually means. Object doubling, for instance, is fully part of the grammar in (western) Macedonian, but it is a pragmatically controlled feature in all of Bulgarian. Thus the structure is “present” in both languages but occupies a very different place in each, so that one could say that *grammatical* doubling is absent from Bulgarian but present in Macedonian. Thus, the relative degree of embedding into the grammar of a language needs to be controlled for, and this is not done in any of the numerological approaches we are aware of.

In addition, the dialect-versus-language question is also relevant here. In some enumerations, e.g., that of van der Auwera, Tosk and Geg Albanian are both considered to count towards the score of “Albanian.” Thus, on the feature of the use of finite complementation instead of infinitives, where Tosk makes greater use of finite forms than Geg does, Albanian is given a score of 0.5, halfway between full presence (1.0) and full absence (0) of the feature. Southern Geg, however, goes with Tosk on this feature.¹²⁶

When the spurious Balkanism is dropped and Macedonian’s features are included (and here we include the majority of Macedonian dialects), then Macedonian, even using such a partial list, is indeed the most Balkanized, as has often been observed in the Balkanist literature (see also, e.g., Hamp 1977a). Still, the point is not to prove that one language or another is more Balkan; rather, we would insist that counting is not the path to insight about the Balkans. What is needed is not a listing of language names matched up with a score vis-à-vis some feature, but rather a “mapping” (in a figurative sense) of locales, and greater attention to the grammatical realities in the languages and dialects of each region. In that way, whether Geg and Tosk count towards “Albanian” would not matter, since the importance would be placed on what the patterns are of diffusion (or not) of features into the locales for which *Geg* and *Tosk* serve as convenient labels that can also be subdivided, where relevant. As with all dialect classifications, the choice of features is relevant (cf. Friedman 1996b).

As it happens, van der Auwera does provide some maps and graphic displays in which the positions of the languages are represented in an approximate geographic fashion (borrowing the notion of “isopleth” from meteorology, much as “isogloss” represents an adaptation), but they are presented with the following caveat (p. 262):

Note finally that the isopleth map is not intended to say anything specific about the cause or diachrony of the convergence; Though Map 2 depicts a core and

¹²⁶ See also §3.5 below for another way in which dialects prove problematic for one type of number-based notion in the classification of Balkan features.

a periphery there is no suggestion that the Balkan type started in the core nor that all isoglosses necessarily include the more central languages. For Map 2 Bulgarian does happen to be included in all isoglosses, but Macedonian is only half-in for the [schwa] isogloss whereas Tosk Albanian, which has the same score as Macedonian, is half-in for the [infinitive loss] isogloss.

And just before that he notes:

The agreement is not perfect, the reason being that other linguists implicitly or explicitly base their judgment on slightly different feature lists. Thus we see Hamp 1977a:281 saying that Macedonian is the most Balkan language, and in the feature counting of Campbell [et al.] 1986:561 Romanian is the winner.

The more general problem here is that one has always to keep in mind that the Balkan sprachbund is precisely an historical phenomenon, shaped by countless instances of speaker interaction. Typological approaches to the Balkans do not take account of diachrony, a problem that was first voiced by Hamp 1977a: 279, 281 and cited in Campbell et al. 1986: 534–535:

Yet while the comparative method is unquestionably an historical study, the field of areal linguistics is no less so; for it too is occupied with analyzing the result of specific, if multiple, linguistic events of the past. Both the comparative method and areal linguistics are historical disciplines – twin faces of diachronic linguistics, if you will.

Hamp 1977a: 282 cogently further observes:

Pavle Ivić has pointed incisively to the difficulty in drawing compact borders to a sprachbund; the configuration is much more that of a spectrum. Yet here we have a multiple offender in Albanian, which in other ways seems to lie near the heart of the Balkan sprachbund. A gross inventorizing would never catch this important textural aspect. [...] [A]real questions can be approached meaningfully and fruitfully only if they are treated in specific terms for what they are – the results of developments with historical depth and specificity.

It follows then that describing a language in the Balkans in terms of being more or less Balkanized should not be divorced from the historical process of (linguistic) Balkanization, which is a phenomenon of contact-induced linguistic convergence that has historical causes and is thus diachronically motivated. And it is precisely in the dialectological and historical facts that these processes can be traced and accurate representations achieved. In this regard, the choice and representation of features are not arbitrary, and their historical origins need to be understood (cf. also Bisang 2006).

This is the problem, moreover, with approaches such as those of Schaller 1975, Campbell et al. 1986, van der Auwera 1998, and others (a most egregious example being Tomić 2004) and their attempts to rank languages as “more” or “less” Balkan (cf. as well the critique of this approach in Joseph 1987a and Friedman 2008a). Admittedly, most Balkanists in practice do impressionistic “gradient” Balkanistics to some extent: note the reference in Hamp 1977a: 281 and Topolińska 2014 to Macedonian as the most thoroughly Balkan of the Balkan languages, and the

observation of Hamp 1989a: 44 that “Tosk Albanian is notably Balkanized and furnishes the full range of Sandfeld’s features, while Geg (especially northern Geg) stands at a conservative remove.” It is no coincidence that in such matters Balkanists focus on Macedonian and Tosk Albanian, as well as Aromanian and northern Greek, as these languages/dialects are spoken in what was (and in some cases still is) the most complex multilingual convergence zone within the Balkans (see Friedman 1994a, 2018a). Still, the crucial point is that even if we operate informally at least with intuitions about gradience, this should not be the theoretical basis for evaluating the nature of the sprachbund: counting does not yield greater understanding, and even assessing “core” versus “periphery” for an area therefore becomes problematic.

As in dialectology, so too in language-contact studies, each feature in a given area can have its own history, and the discussions in the chapters that follow adduce instance after instance where that is the case. Therefore, it can hardly be revealing to just give lists of features and note their occurrence or lack thereof. A “speaker-plus-dialect” approach leads to, and thus requires, more attention to the dynamics and history of language contact in the Balkans and less to numbers and ranking of languages. What emerges out of such attention, then, is not only the uncovering of convergences and similarities in various locales and in various features, but also an appreciation for divergences and differences in them as well (cf. Friedman 1983 on comparative Balkan grammar).

3.4 Sprachbunds Beyond the Balkans

In Chapters 1 and 2 we were concerned with defining the Balkan sprachbund and giving historical background to the development of the concept of the sprachbund as a Balkan phenomenon. Because it was the first sprachbund to be identified as such – in fact, as noted in §2.3.1, the very term was invented by Trubetzkoy 1923, 1930 to capture the relationship of Balkan Slavic to neighboring non-Slavic languages – a history of the Balkan sprachbund is, to some extent, a history and definition of the concept of sprachbund itself. Nonetheless, the larger history of the notion of a sprachbund is an important part of locating the present work in the broader context of areal or contact linguistics. In this section, therefore, we examine the sprachbund as a general and theoretical linguistic construct vis-à-vis other areas and linguistic sub-disciplines, and also explore how the notion and its exemplary illustration, the Balkans, have fared in this larger academic context.

3.4.1 “External” History of the Sprachbund

The sprachbund as a construct has its own particular history within historical linguistic investigations that is important to take into consideration. Moreover, the historical background to the development of this notion is also significant.

3.4.1.1 From the Balkans to North America and Beyond

As already noted in Chapter 2, the commonalities of the Balkan languages contributed to discussions of the “family tree” and “wave” models of language change and relatedness, and they were the original example of the concept of *sprachbund* as proposed by Trubetzkoy. Over this same period, however, a debate continued to rage over what features of a language could be “borrowed,” i.e., were subject to change owing to the influence of contact with another language. Having discussed these issues conceptually in §3.2.1.7, we turn now to some of the historical background that informs these issues.

Scholars such as Müller 1885: 83ff., Whitney 1868: 199, and even Sapir 1921: 220 insisted that the type of structural borrowing that characterizes the Balkan sprachbund could not occur (see also §3.2.1.7).¹²⁷ Jespersen 1922: 213–215 characterizes Whitney’s statement (“Such a thing as a language with a mixed grammatical apparatus has never come under the cognizance of linguistic students: it would be to them a monstrosity”) as “an exaggeration” that “cannot be justified” although he attributes “some truth” to the statement insofar as loanwords are generally adapted to the borrowing language, or at least not borrowed with the grammatical inflections of the source (a statement that we now know to be untrue with respect to the Balkans).¹²⁸ Bloomfield 1933: 470 actually cites the Balkans as an example of what he calls “intimate borrowing” and specifies the postposed definite article and infinitive replacement as concrete instances. More recently, Labov 2007: 349 has resuscitated the old claim that structural borrowing is rare or nonexistent, although he ultimately recognizes that it is possible and even cites Balkan infinitive replacement as a counterexample (see footnote 67). Even before Jespersen and Bloomfield, however, challenges to what we can call the “genetocentric” view of language change – a view that recent attempts to apply computer modeling to language change¹²⁹ have brought back to renewed, and arguably undeserved, prominence – were coming from linguistic anthropologists studying native North American languages, apparently in ignorance of developments in Europe.

Emeneau 1962 (see 1980: 56) refers to evidence of contact-induced change in the indigenous languages of California in earlier work, specifically Dixon & Kroeber 1903, 1907, and in his classic article on India as a linguistic area, Emeneau 1956 (see 1980: 105) even goes so far as to suggest the possibility that

127 At stake was not just the limitations of the family-tree model of genealogical linguistic methodology in general, but, as Emeneau 1962 (see 1980: 55) points out, even the very classification of English as Germanic (Müller 1885: 84, see also §3.2.1.8). It can be argued that Müller’s oft-quoted dictum denying “the possibility of a mixed language” (1885: 86), cited by Thomason & Kaufman 1988: 1 as *Es gibt keine Mischsprache* (‘there are no mixed languages’), reflects in a statement about language (which in that period was often conflated with notions of race) the kinds of anxieties about racial purity that are also reflected in the concept of miscegenation and laws against it, as well as the embracing of eugenics and the Holocaust it led to.

128 A classic example is the set of Turkish conjugations in Romani dialects (see Elšik & Matras 2006: 134–137; Iglá 1996: 214–219; and Friedman 2013b).

129 For instance, Nakhleh et al. 2005; by focusing on phylogeny, such models ignore diffusion or treat it as an undifferentiated intrusion into typical generational language transmission (cf. Donohue 2012).

Boas 1911: 47–53 may have been the first to raise the problem of “the diffusion of linguistic traits across genetic boundaries.”¹³⁰ In that same article, he cites Kroeber’s presidential address to the Linguistic Society of America in 1940 (Kroeber 1941), in which he states that: “words can be borrowed freely between distinct languages, but grammar with difficulty if at all,” as well as Hoijer’s 1948: 335 comment that: “Traits of language are not readily borrowed.”¹³¹ Emeneau 1956 (see 1980: 107) does indicate in passing an awareness of Trubetzkoy 1939 and Jakobson 1944/1971, but it is noteworthy that Jakobson 1944/1971 is about Franz Boas, and describes, among other things, the growth of Boas’s conviction – based on study of Northwest Coast Native North American languages such as Haida and Tlingit – that the methodology of genetic linguistics is inadequate for explaining contact-induced change.¹³² In that article, Jakobson 1944/1971: 485–486 relates that it was Boas’s reading of Trubetzkoy 1939 that brought him (Boas) to the realization that he was not the only scholar who held such views, and that work on language contact had been going on for quite some time in Europe.¹³³ As Jakobson 1944/1971: 486 tells it: “The bitterness of loneliness had disappeared.” Even Emeneau 1980: 1, in characterizing the term *sprachbund* as “popularized [if not invented] by Trubetzkoy and Jakobson as early as 1931,” indicates the lack of communication between North American and European linguists, since Emeneau was apparently unaware of Trubetzkoy’s 1930 formulation for the First International Congress of Linguists.

As noted in §2.3.1, Jakobson took Trubetzkoy’s idea somewhat further in both geographic and conceptual terms. Whereas Trubetzkoy’s model envisioned an area in which attested multilingualism resulted in structural change (as exemplified at the morphosyntactic level), Jakobson’s concept involved vast areas where such levels of multilingualism do not occur. Rather, Jakobson’s more general phenomena were phonological or typological (palatalization, monotonic prosody, the existence of case in nominal inflection) that could be attributed to (remote) genetic inheritance, universal tendencies, or possibly a “chain” of overlapping areas of convergence (on this last idea, see also Hamp 1989a, as discussed in §3.3, as well as Enfield’s 2003 epidemiological model, as discussed in §3.2.1.4). The spread of

130 To be sure, Emeneau covers himself, after a fashion, by writing “whether it was original with him, does not matter – for I do not intend to be bibliographically complete” and he does mention Trubetzkoy’s *Sprachbund*, via secondary sources, in his footnote 26 (Emeneau 1956; see 1980: 124). However, it is symptomatic of the lack of scholarly contact between Europe and North America that at that time he was not aware of European works he later cited in Emeneau 1962 (see 1980: 56).

131 Cf. Sankoff’s 2002: 658 claim, approvingly endorsed by Labov 2007: 348, that “Morphology and syntax are clearly the domains of linguistic structure least susceptible to the influence of contact, and this statistical generalization is not vitiated by a few exceptional cases.” *Plus c’est la même chose*.

132 See Mithun 2017 on an areal linguistic view of North America, where there does appear to be a substantial number of “extensive parallelisms in grammatical categories and structures” (p. 878).

133 As Emeneau 1956 (see 1980: 121) archly but correctly observes, Trubetzkoy 1939: 84 gets a bit carried away with the importance of the preservation of typological features in language classification, but this does not vitiate the basic principle that languages can and do change at all levels as the result of contact, i.e., all is not drift.

glottalization to Ossetian, East Armenian, to some extent to Karachay-Balkar (Pritsak 1959: 350) and some dialects of Kumyk (Bunajksk and Xajdak [Kajtak], Benzing 1959: 396; Gadžimahmedov 1985) and Azeri (e.g., in contact with Tabasaran, Efendiev 1985) in the Caucasus, and the use of ingressive velaric airstream (“clicks”) in all the Nguni languages of the Bantu family except Sotho as a result of a combination of contact with Khoisan languages and the institution of *hlonipha* ‘respect speech’ (see also §5.2 and especially footnote 16), which required extensive avoidance of significant amounts of tabooed vocabulary for large numbers of speakers (Irvine & Gal 2000), all demonstrate that phonological features can spread independently of other parts of grammar.¹³⁴

During the same post-war decade as Emeneau 1956, Weinreich 1953 addressed the question of language change through bilingualism, where he explicitly avoided multilingual situations (§3.1), and in Weinreich 1958 he also addressed the concept of *sprachbund*. In this latter work (pp. 378–379), he does not dispute the phenomenon denoted by the term *sprachbund* but rather complains that the term itself is faulty in that it implies a unit to which a language may or may not belong.¹³⁵ He proposes the term *convergence area* as “more specifically meaningful” than Emeneau’s 1956: 16, 1980: 124 “linguistic area,” a term first used by Velten 1943 as a translation of German *Sprachbund*. Emeneau 1965 (see 1980: 127) restates his definition from 1956: “an area which includes languages belonging to more than one family but showing traits in common which are not found to belong to the other members of (at least) one of the families,” and argues that *convergence area* denotes the process and that perhaps *diffusion area* might be “an equivalent

134 In a sense, phonological features seem more closely allied to the lexicon than morphosyntactic features insofar as both are able to diffuse across genetic linguistic boundaries without being accompanied by the diffusion of structural features (except, perhaps, insofar as these can be contained in phraseological calques). Tuite 1999 makes the point that even ergativity does not really function as an areal rather than merely typological feature in the Caucasus, since its realization is so radically different in the indigenous language families. Masica’s 2001 map, nonetheless, is suggestive that some sort of special treatment of some transitive subjects is typologically areal. In the end, the only feature we are left with that is truly a pan-Caucasian areal feature is glottalization, and perhaps some phraseological calques; still, see now Grawunder 2017 on the Caucasus as a linguistic area. On the spread of retroflexion from Dravidian to Indic and some contiguous Iranian, see Emeneau 1965 (see 1980: 127–129), Hock 1975, and Mitchell 2015.

135 Some accounts even make the “Balkan linguistic union” seem like the “European Union” where official membership is decided on by existing members and certain benefits and obligations follow from membership. Thus, for example, in a misleadingly ambiguous footnote, Tomić 1996: 814 argues that all of Serbo-Croatian can be described as *Balkan Slavic*. She bases this in part on a mischaracterization of the arguments for the existence of the Balkan sprachbund as involving the number of “structural properties” or “areal typological properties” (she uses these two interchangeably; on the problem of conflating *areal* and *typological* see Hamp 1977a) “required for granting membership in the Balkan language union.” She then argues that since not all Balkan languages have all Balkanisms to the same extent, and since all of Serbo-Croatian supposedly has some Balkan features, all of Serbo-Croatian can be included as Balkan Slavic. She then writes: “Let me remark that I have discussed the issue with Eric Hamp, Victor Friedman, Zuzanna Topolińska and Lili Laškova, and they agree with me.” While it is true that not all Balkan languages and dialects have exactly the same configuration of contact-induced changes – a fact with which all of the scholars listed would agree – the structural differences between the Torlak dialects of the former Serbo-Croatian and the rest of Serbo-Croatian are so fundamental, at all linguistic levels, that none of the scholars cited agree with Tomić’s final conflation.

term of specific meaningfulness.” But then he notes that a dialect area can also be a diffusion or convergence area and ends by preferring his original term.¹³⁶ (See Alexander 1981, 1983, 1984–1985 for a discussion of dialect-level diffusion in Balkan Slavic.)

More recently, some scholars (e.g., Stolz 2006) have suggested eliminating the concept of *sprachbund* altogether. Such a suggestion, however, misses Trubetzkoy’s original point, which is the fact that languages can come to resemble one another over time owing to contact-induced change as opposed to showing resemblances due either to divergence from a common (genealogical) ancestor or to universal (typological) tendencies of human language, and that such contact-induced change can be conceptualized as an historically grounded relationship. It is these historically grounded, contact-induced changes that we seek to study here, and it is the term *sprachbund* that refers to the collective languages participating in such changes. While it may well be the case that in various parts of the world, determining the source of a resemblance is a difficult matter to tease out, nonetheless in the Balkans we have rich sources of historical and textual records. Thus, the abuse or misconstrual of the concept of *sprachbund* does not in and of itself constitute a justification for its elimination.

3.4.1.2 Beyond the Sprachbund: Contact and Creolization

In §3.2.1.6, we mention the correlation between social movements and legislation, on the one hand, and the identification of codeswitching as an actually occurring normal phenomenon in various sociolinguistic contexts of bilingualism, on the other. Beginning in the 1960s and 1970s, a change in linguistics that coincided with the civil rights movement in the United States and the decolonialization of Africa and the British Caribbean (and the concomitant rise of associated nationalisms) was the establishment of pidgin and creole studies as a distinct discipline. Hymes 1971a: 3 cites Bloomfield 1933: 472–474 as being the first to distinguish effectively pidgins and creoles as types of languages.¹³⁷ Hymes 1971c was itself a significant collection in the history of pidgin and creole studies, being the publication of the proceedings of a conference held in 1968 which, as Hymes 1971b: 65 notes, marked the end of “the ‘formative’ period of modern pidgin and creole stud[ies]” and their establishment as part of the general investigation of language change.¹³⁸ Although the suggestion that African-American Vernacular English is descended from, or represents the last stages of, a creole goes back at least to Bloomfield 1933: 474, Hymes 1971b: 33 was explicit in connecting the

136 Emeneau 1974 (see 1980: 127) argues that “the same process produces the same result whether between dialects of a language, languages of a family, or languages of several families.” This is similar to the argument made by Mufwene 2000 concerning creole genesis. Some linguists see dialect borrowing as a different process however; see Nakhleh et al. 2005, for instance. Our position is given in §3.1.

137 Jespersen’s 1922: 216–236 account does not distinguish the two, nor does he recognize them as genuine linguistic systems.

138 Wurm 1970 is another important publication worth noting in this context.

significance of creolistics to the rise of Black nationalism in the United States in the 1960s.¹³⁹ The 1970s and 1980s saw the emergence of pidgin and creole linguistics into debates about language change, language origins, and language universals, the details of which are impossible to chronicle here, and the literature has grown exponentially since then (see Mufwene 2001b, 2004, 2008).

The matter of most concern to Balkan linguistics is the relationship of creolization to contact-induced change in general. In their 1988 study of language contact, creolization, and genetic linguistics, Thomason and Kaufman – who, like Weinreich thirty-five years before them (see §2.3.3), explicitly focus on bilingual situations and exclude sprachbunds since the directionality of contact-induced change can be difficult to establish because, in their words, “sprachbund situations are notoriously messy” (1988: 95–96) – distinguish creoles from other types of contact-induced linguistic phenomena as being the result of “abnormal” or “abrupt” transmission.¹⁴⁰ They thus argue for a fundamental difference between the process of creolization, on the one hand, and other types of contact-induced change, including structural and lexical borrowing, on the other. In such an analysis, the Balkan languages as we have them today would presumably be excluded, although in some respects the external histories of these languages and their development show precisely the same types of linguistic change.

It is these similarities that have attracted the attention of some Balkan linguists at the turn of the twentieth/twenty-first centuries, as shown by a mention in Topolińska 1995b (cf. also Topolińska 2000) and Lindstedt 2000, as well as Hinrich’s 2004 attempt to describe Balkanization as a kind of creolization. On the other hand, Mufwene’s 1996 point about language ecology (a term borrowed from Haugen 1972 and used by Cyxun 1998) and the social rather than structural features that are responsible for creolistics as such (Mufwene 1997) is applicable to Balkan linguistics as well. According to these arguments against creole exceptionalism, as noted in §3.2.1.8, there is nothing that defines creolization as a process of linguistic change differing from those phenomena that characterize “normal” transmission, i.e., changes owing to drift, borrowing, interference, metatypy, fusion, code-copying, language shift, etc., as attested in other settings; that is, there is nothing in the structure of creole languages that identifies them as a unique linguistic type. As Mufwene (2003) asks:

Can creolists have been prevented from learning informative lessons about the role of socio-economic ecology in an otherwise normal differential language evolution anywhere by the following factors among a host of others: 1) too much eagerness to treat creoles as interesting deviations from “normal” language

139 The debate over the origins of African-American Vernacular English (AAVE) continues; see Mufwene 2001b and Winford 1997, 1998 for summaries and references on creole origins.

140 Matisoff 2019: 122 agrees with Thomason and Kaufman’s assessment, saying “A true sprachbund is messy ... in the sense that the directionality of the sharing can be difficult to determine.” Matisoff himself has elsewhere (e.g., in 1978: 2) used similar language in describing the Tibeto-Burman family, saying, as a description of the characterization of Tibeto-Burman advocated by Paul Benedict 1972, that it was “an interlocking network of fuzzy-edged clots of languages, emitting waves of mutual influence from their various nuclear ganglia. A mess, in other words.”

evolution; 2) precocious zeal to account uniformly for developments that have really varied from one setting to another; and 3) disregard for the complex history of population movements and for the kinds of interactions among the relocated populations when the nature of the research on the development of creoles militates that such factors be incorporated in our analyses?

The first question resonates with nineteenth-century views noted above, the second with debates within Balkanistics over contact-induced versus internally motivated developments, and the third with issues of local areas within the Balkans.

Related to these questions is the debate between substrate and superstrate hypotheses in explanations of creole genesis, which echoes debate in Balkan linguistics between substrate and superstrate explanations of convergent developments. In Balkanistics, the roles of Latin and Greek (Slavic has only rarely been invoked since Leake 1814), on the one hand, and the variously poorly attested ancient Indo-European languages one of which must be the ancestor of Albanian, on the other, mirror to some extent the debate surrounding Indo-European lexifiers and African languages in creolistics. Just as knowledge of the specific early modern dialects of English, French, Spanish, etc. that affected creole formation is a crucial but poorly attested factor, so, too, in the Balkans, the thousand-year gap between the last Balkan Latin inscription and the first attested Balkan Romance (Romanian) document and the fact that most medieval and early modern Greek documents do not reflect vernacular practice limit our knowledge concerning the roles of these languages. The invocation of African substrata in creolistics bears similarities to the invocation of unattested ancient Balkan languages insofar as (1) the documentation of the relevant African languages post-dates the formation of creoles and (2) the circumstances of creole formation are such that the diversity of possible African sources for them is similar to the diversity of possible ancient Indo-European sources in the Balkans, with the difference that we presume the African sources all have living relatives in Africa while in the Balkans Albanian is the sole survivor of what was once a considerably more diverse linguistic situation.

3.4.1.3 Balkanology and Eurology

As noted in §3.3, Hamp 1977a points out that unlike genetic and areal linguistics – his “twin faces of diachronic linguistics,” elucidating sources of similarities and differences – typology is achronic and seeks to explain resemblances among languages through the nature of language itself, the ideal realizations of which are universals, although the complex realities of which are usually tendencies. In that same article, Hamp cautions against the conflation of areal and typological linguistics as seen, e.g., in Sherzer 1976. More recently, the conflation of the areal and the typological has been seen in projects such as EUROTYPE and the *World Atlas of Linguistic Structures* (WALS). With regard to the former, König 1998: v–vi writes:

Language typology is the study of regularities, patterns and limits in cross-linguistic variation. The major goal of EUROTYPE was to study the patterns and

limits of variation in [...] the languages of Europe [...] by characterizing the specific features of European languages against the background of non-European languages and by identifying areal phenomena (*Sprachbünde*) within Europe [...] and thus contribute to the characterization of Europe as a linguistic area (*Sprachbund*).

For the latter, the editors of *WALS* described it as “a kind of dialect atlas of the ‘dialects’ of Human Language” (Haspelmath et al. 2005: 1) thus taking each language as a dialect of the phenomenon of human language rather than for a single language and claiming that this approach suited their typological purposes. Problems of omission can of course occur when all the dialects of a language are subsumed under whatever dialect happens to be the standard (or whatever dialect happens to be best described); for instance, to take just one example, in the section on perfects, Russian is said to have no perfect, because perfect constructions only occur in North Russian dialects not in the standard language (wals.info/feature/68A#2/25.9/133.5). However, leaving aside such concerns, a fundamental problem here is with the conflation of the areal and the typological; as Hamp 1977a notes, the former is grounded in history while the latter is not. We have, then, a basic issue of slippage not unlike that occasioned by Jakobson’s 1931ab/1962, 1938/1962 extension of Trubetzkoy’s concept from the Balkans to all of Eurasia except the eastern and western extremities (see §§2.3.1, 3.4.1.2).

As exemplified in Haspelmath 1998, who resuscitates Whorf’s concept of Standard Average European, EU-funded approaches such as EURO-TYP have reconceived “Europe” as a linguistic area whose core is located, as if by coincidence, at the Romance-Germanic border from Holland to northern Italy, i.e., roughly the territory of the old European Economic Community (a.k.a. the European Common Market), the post-World War Two precursor of today’s European Union. Moreover, this territory corresponds roughly to that of the Holy Roman Empire, which was fragmented into dozens of autonomous and warring polities precisely when the territory of the Ottoman Empire provided the unified conditions under which the Balkan sprachbund reached its distinctive shape. Such approaches that conflate areal and typological linguistics in Europe can be described as *Eurological* (Friedman 2008a). Put another way, recent Eurological approaches to language typology that place what is roughly the territory of Holy Roman Empire at the center of linguistic convergence (e.g., Haspelmath 1998) are ideologically related to Jakobson’s placement of the Holy Russian Empire at the center of a Eurasian convergence area. Thus, in a sense, Haspelmath 1998 has already answered “yes” to the question posed by Feuillet 2000b regarding whether Balkan linguistics could become a branch of Euro-linguistics. The means by which this answer is arrived at, however, do not rely on historical linguistic methodology.

The idea that the Balkan languages are part of some larger European neo-sprachbund is not new (see Aronson 2007; Feuillet 2000a; Reiter 1994; Hock 1988), but it has received additional impetus from recent work by Bernd Heine and Tania Kuteva, e.g., Heine & Kuteva 2006: 1–47 and references therein, as well as works such as Hock 1988. Here, too, Europe is taken as a kind of linguistic area

in much the way that Jakobson 1931a/1962 took Eurasia as an area more than half a century ago. Such approaches conflate the typological and the areal and thus ignore or erase the historically attested or reconstructible specificities of contact-induced language change.

3.4.1.4 Grammaticalization and Contact

Mention of Heine and Kuteva's work on what they call, e.g., in their 2006 title, "the changing languages of Europe" leads us to offer some commentary on the framework in which they couch much of their discussion of Europe as a new linguistic area. Their analytic framework is that of grammaticalization, by which is meant the development – and the study thereof – of full grammatical "machinery" (processes, categories, morphemes) from lexical or only-somewhat-grammatical material (see, for instance, Traugott & Heine 1991; Hopper & Traugott 1993, 2003; and Heine & Kuteva 2002, among other notable works). While the considerable body of literature on grammaticalization that has emerged in the past few decades has largely focused on language-internal developments, more recently, attention within grammaticalization studies has turned to language contact. This was perfectly understandable, given that students of grammaticalization are interested in the sources of grammatical material, and language contact is one of the ways in which languages acquire new grammatical morphemes and processes and one of the causes for the alteration of existing ones.

There are numerous studies (e.g., Tagliamonte et al. 2014, to mention just one) that apply tenets of so-called grammaticalization theory to specific contact situations, including dialect contact and pidgin and creole languages, and Heine & Kuteva (2005) devote considerable space to "contact-induced grammaticalization" (pp. 13–21), whereby the impetus for the development of a grammatical form in one language is a model from an adjacent other language. Additionally, a further grammaticalization-based notion of particular relevance to the study of sprachbunds has come in works such as Kuteva 2001; Heine & Kuteva 2001, 2003, 2005, 2006, namely "grammaticalization areas." Heine & Kuteva 2005: 182 define "grammaticalization area" as follows: "a group of geographically contiguous languages that have undergone the same grammaticalization process as a result of language contact." They specifically mention the Balkans as such an area, and point to grammatical parallels among the languages such as a 'want'-based (devolutive) future and object reduplication, among others. More specifically, with regard to those features, they see the parallel deployment of a verb for 'want' for a grammatical future and the parallel reductions in form that are found in the various Balkan languages as a basis for treating the Balkans as a "grammaticalization area," and so also with the other features they consider.

The parallelism in the future formations across the Balkans was already observed by Kopitar 1829, and all general treatments of Balkan linguistics have discussed the relevant data, as indeed do we in §6.2.4.1. It is not clear, however, what the notion of grammaticalization area adds here, and if in fact it is really the case that language

contact is responsible for the grammaticalization “process”¹⁴¹ seen in each language. The basic mechanism that Heine and Kuteva make use of is what they call “replica grammaticalization” whereby (p. 92) “it is not a grammatical concept but rather a grammatical process that is transferred from the model (**M**) to the replica language (**R**).” They envision this transfer as taking place as follows (p. 92):

- (a) Speakers notice that in language **M** there is a grammatical category **M_x**.
- (b) They create an equivalent category **R_x** in language **R**, using material available in **R**.
- (c) To this end, they replicate a grammaticalization process they assume to have taken place in language **M**, using an analogical formula of the kind [**M_y** > **M_x**] : [**R_y** > **R_x**].
- (d) They grammaticalize **R_y** to **R_x**.

This seems to us to involve a sleight-of-hand that does not correspond to what speakers actually do. That is, copying a form, or the function of an equivalent form, from another language is something that speakers in contact situations commonly do; that is what is involved in calquing (see §3.2.1.7, for instance). But it is unclear how speakers could take a step back, as it were, and recognize how to create the equivalent of **M_x** in their language (the **R_x** that is created via the “material available in **R**”), since they would somehow need to reconstitute a starting point within their language (**R_y**) in order to generate (via “grammaticalization”) the form they copy into their language (**R_x**). Why, one might ask, if material is available to model **M_x** as **R_x**, must there be an **R_y** at all? We do not deny that this sort of cross-language analogy is possible and can see how it would work if the source language shows variation between, say, a free form and a functionally equivalent bound form;¹⁴² however, we do not see it as always being an essential part of the replication of forms and functions across languages.¹⁴³

Moreover, the notion of transferring a grammatical process across languages seems to be a difficult one. We see how the material and even a relation between different manifestations of that material could be transferred across languages, but the actual *process* by which a relationship is effected in a language (e.g., sound changes that reduce a form or semantic changes that extend a form into a wider context, both “change events” that are often associated with grammaticalization) would seem to have to be established within each language. In fact, although

141 On the use of scare quotes here, see below, where we question the conceptualization of grammaticalization as a process.

142 In fact, as we state in §7.6.2, we suspect a cross-language process like this was involved in the creation of a new independent prohibitive function for a certain class of negation marker in Greek and Albanian.

143 To some extent, Heine and Kuteva’s view is intimately tied to their beliefs about possible sources of grammatical material; if one is committed to the view that lexical items are the only source of grammatical morphemes, a position that many grammaticalizationists take (see, e.g., Hopper & Traugott 2003: 132: “there is no evidence that grammatical items . . . can be innovated without a prior lexical history”), then one has to start with such a source even in contact situations of replication. We reject this view and consequently see grammatical copying in language contact in an accordingly different way.

proponents of so-called grammaticalization theory routinely characterize it as a process, there are others (e.g., Newmeyer 1998; Campbell 2001a; Janda 2001; Joseph 2001d, 2003b) that view grammaticalization as a result, a by-product of other processes and mechanisms of change, and to our way of thinking, the emergence of grammatical material in contact situations adds further evidence to the result-interpretation of grammaticalization and not the process-interpretation. In other words, *grammaticalization* is best understood as a cover term for the result of a series of processes which themselves are quite heterogeneous.

It thus appears that so-called grammaticalization theory is not so much a theory as an ideology. Starting from the amply attested fact that certain kinds of lexical items have a tendency to become semantically bleached and become grammatical markers, proponents of this so-called theory argue that such transformations can only be unidirectional. Such an approach to historical linguistics posits an invisible hand, somewhere, e.g., in the human brain, that prohibits developments in the other direction, i.e., from grammatical element to independent lexical item. Proponents of this ideology expend considerable effort redefining the rules (often with circular results) in an attempt to exclude the numerous exceptions that have been adduced (see Janda 2001; Joseph 2006, 2011a, 2014, 2019b; and Norde 2009 for relevant examples and discussion). Unlike Neogrammarian theory, which productively applied the principle that an exception to a sound law can be formulated as a regularity in its own right, grammaticalization ideology has no tools for handling exceptions aside from *a priori* exclusion. Given these problems, attempting to redefine the sprachbund as a “grammaticalization area” not only misses the cultural richness and linguistic complexity of areal phenomena but also erases the specificity of language change taking place in a particular socio-political context.

This however is not the place for a full critique of various approaches to grammaticalization. It suffices to say here that while grammaticalization does have a place in the understanding of language contact, in the sense that contact between languages is a source of new grammatical material, functions, and categories, the added notion of “contact grammaticalization” seems to involve the traditional notion of calquing, and the further notion of “grammaticalization area” simply labels the observation that languages in contact can come to share grammar. We thus do not address it further here, though we take note of it as a relatively recent addition to the repertoire of constructs posited within contact linguistics. That is, we acknowledge that this is a direction that one school of thought is moving in, but we do not see anything in it that alters what we believe is going on in the Balkans.

3.4.2 Does Size Really Matter? On Members and Membership

Among the questions about sprachbunds are several that pertain to various minima that might be brought to bear in determining whether some group of languages constitutes a sprachbund. For instance, one can wonder: How many languages are

needed? How are boundaries to be determined (see also §1.1)? How many features are required? Must they be evenly distributed? We address these and related issues here.

3.4.2.1 On Numbers of Languages and Locations of Boundaries

Since language contact is involved in the basic definition of a sprachbund, clearly the answer is greater than one, but how much greater? Two languages would necessarily constitute the logical minimum for a sprachbund just as the minimum needed in genetic linguistics for a language family – a situation where by definition, essentially, no contact is involved – is one, as in the case of so-called *isolates*, languages not demonstrably related to any other.¹⁴⁴ While for a sprachbund, one has to have at least two to tango, Thomason 2000: 312, 2001: 99 is among those who insist that a sprachbund must be at least a *ménage à trois*, a point to which we return below.

Trubetzkoy 1930, in his formulation of the difference between the language family and the sprachbund, made no mention of boundaries or numbers. He was attempting both to account for and to distinguish the two diachronic ways languages come to resemble one another: what Labov 2007 has distinguished as *transmission* and *diffusion* (cf. also Hamp 1977a). As Trubetzkoy recognized, the language family is distinguished by regular sound correspondences that can be recognized using the Comparative Method, and this in what we can call grammatical morphemes and core vocabulary. The sprachbund on the other hand, as Trubetzkoy defined it, was characterized by shared syntax, morphosyntax, non-systematic phonological correspondences, and “culture words.”

While the regularity of sound correspondence has a predictability that neatly parallels the scientific method (as clearly demonstrated by the discovery of Indo-European laryngeals),¹⁴⁵ the distinction between *core vocabulary* and *culture words* is to a certain extent vague, notional, and not immune to social manipulation. Thus, for example, in the Pomak (Balkan Slavic) dialects of Greece, numerals and kinship terms are Turkish despite the fact that the dialects are clearly of Slavic origin, arguably connected to these speakers’ view of Turkish as having importance to their identity as Muslims. Any given body part or basic verb of motion, feeling, bodily function, etc., has the potential to be replaced by a loan. Nonetheless, even Romani, which has been subject to massive multilateral contact, has an Indic core that accords remarkably well with the notional concept in its basics.

144 For example, Sumerian is generally held to be an isolate, as are Zuñi and Basque, to name a few. Such isolates of course can have internal dialect diversity – quite rich in the case of Basque (cf. Trask 1996b) – a situation that stretches the notion of “language isolate” through its intersection with the vexing language-versus-dialect question (on which see §1.2.3). Also, it may well be that these isolates do in fact form a stock or phylum with some other existing language, but such connections are not demonstrable given our current state of knowledge and methodology.

145 For a compact but readable account of the methodological steps that led to the positing of laryngeal consonants for Proto-Indo-European, and the subsequent confirmation of their existence, see Anttila 1989: 264–273 and Hock 1991: 545–550.

The usefulness of the concept of language family is considered to be self-evident since it gives us an historical basis for accounting for language resemblances and relations. At the same time, as noted above, if we cannot demonstrate a relationship of a given language to any other, then the existence of a family with only a single member poses no problem to the concept of language family. Similarly, when seeking the defining characteristics of this or that language family, it is precisely the shared history of regular sound change combined with notions of core vocabulary and basic grammar that enable us to speak of boundaries, although Thomason & Kaufman 1988 interrogate the rigidity of such conceptions. It can be argued that nineteenth- (and early twentieth-) century ideas connected with the need to establish purities of lineage in “races” were carried over to languages as well, whence Schleicher’s characterization of the Balkan languages as “corrupt” and Whitney’s characterization of structural borrowing as a “monstrosity” (see §3.4.1.1). In a world suffering anxieties about “purity” of race and origin, and one in which political (national) boundaries were in the process of being drawn and redrawn, it is understandable that such concerns would permeate academic discourse. Moreover, the difference between a language and a dialect or the definition of a dialect boundary remains, to some extent, a social or political artifact, e.g., the privileging of an isogloss that marks the development of a nasal over that marking the treatment of certain consonant clusters (see §1.2.3).

In the case of the sprachbund, however, the original point that Trubetzkoy was trying to make is sometimes forgotten or misunderstood. Trubetzkoy was not talking about any situation of bilingual contact but rather situations in which there was a range of similarities in syntax, lexicon, morphosyntax, and even phonology, but precisely without regular sound correspondences and shared core vocabulary. Absent from Trubetzkoy’s original formulation but constituting an underlying assumption was areal contiguity, but it is the very nature of areality that raises the question of defining the “area.” Masica 1976: 11 writes: “Some [instances of convergence] . . . involve only two or three contiguous languages. These may merely be instances of what is possibly a tendency for contiguous languages anywhere in the world – or at least contiguous dialects of contiguous languages – to resemble each other in some way or another. Even if every Indian language turns out to be linked to its neighbors by special two-by-two relationships, forming a continuous network covering the subcontinent, this in itself would not establish India as a special area, especially if similar **arbitrary** [our emphasis (VAF/BDJ)] linkages continue beyond India.” Thomason 2001: 99 writes: “The general idea is clear enough: a linguistic area is a geographical region containing a group of three or more languages that share some structural features as a result of contact rather than as a result of accident or inheritance from a common ancestor. The reason for requiring three or more languages is that calling two-language contact situations linguistic areas would trivialize the notion of a linguistic area, which would then include all of the world’s contact situations except long-distance contacts (via religious language . . . , etc.) . . . the linguistic results of contact [among more than two languages] may differ in certain respects.” But, as Hamp 1989a had already pointed out with respect to what is now

former Yugoslavia, even the Balkans can be understood as a “crossroads of *Sprachbünde*,” with “a spectrum of differential bindings, a spectrum that extends in different densities across the whole of Europe and beyond.” If we look to clusterings of smaller convergence areas that can be said to add up to a larger area, then it could well be that as few as two languages will be involved in some clusterings; note, for example, the Macedonian and Albanian of the Debar town dialect in North Macedonia regarding nasalization (see §5.4.1.4).

We are faced then with two problems: The problem of “boundaries” (which subsumes both the territorial implication of “area” and the membership implications of “union” and “league”) and the problem of “number.” These problems are reminiscent of the difficulties in defining concepts such as “nation,” “empire,” “state,” “ethnicity” or “culture,” as well as “language,” “dialect,” “pidgin,” and “creole.” What level of control constitutes a “state”? How big must a “state” be in order to be an “empire”? How many “nations” must it comprise? When is the speech of a community a “dialect” of another “language” and when is it a separate “language”? What is “separate” and how much intelligibility is required before it is “mutual” (cf. Haugen 1966)? From a general theoretical point of view, it does not actually appear to be the case that the kind of diffusion that takes place among three or more languages is in any way qualitatively different from diffusion that is possible between two languages. Moreover, contact phenomena are never arbitrary. They are embedded in social relations as well as the structures of the languages that manifest them. In a sense, the village of Kupwar (Gumperz & Wilson 1971; also Masica 1976: 11) is a linguistic area, albeit one that is part of a larger area just as the dialects spoken in it are parts of larger languages.

If we keep in mind Trubetzkoy’s original motivation in proposing the terminological distinction between the sprachbund and the family, then two languages related by diffusion can constitute a sprachbund in the same way that two languages related by transmission from an ultimately common source can constitute a language family. The crucial difference between the language family and the sprachbund is that the former can be an isolate with but a single member, while the latter by definition requires more than one member in order for diffusion to take place. There is, however, another issue in the definition of a sprachbund as understood by Thomason & Kaufman 1988: 95 – who do not impose a tripartite requirement on the concept – namely that of directionality. It is generally agreed that in bilateral language contact situations, there is usually asymmetry in the direction of transference. In principle, if two languages were demonstrably genetically different enough such that similarities resulting from diffusion could be identified but the directionality could not (or could be shown to be symmetrical), then such a unit could arguably be described as a sprachbund, but such situations appear to be poorly or not attested. On the other hand, a situation such as that found in the Balkans clearly involves diffusion, but directionality can be variable. For our purposes here, determining directionality is desirable but not requisite, and it can even be argued that it is ultimately irrelevant (cf. Ilievski 1973 on the question of

internal versus external factors). Moreover, in the end the size question does not affect the Balkans – regardless of the minimum number that might be determined for talking about a sprachbund, it is clear that the Balkans, with up to six distinct contributing language groups (Albanic, Indic, Hellenic, Italic (Romance), Slavic, and Turkic), would meet any minimum. Similarly, in terms of the number of features, the scope of discussion in this book makes it clear that there are dozens of features to consider, even if not all of them turn out to be due to language contact (cf. the definition of sprachbund).

3.4.2.2 On the Quality and Quantity of Features and Their Distribution

Just as it is reasonable to question the minimum number of languages needed for a sprachbund, one might well ask also how many features must be present in these languages for talk of a sprachbund to be warranted, a question also asked and discussed (and similarly answered) in Thomason 2000: 313–314. Here also, looking to genetic linguistics offers some useful insights. Basically, just as numerical quantifications fail to provide an accurate picture of a linguistic area, so too when speaking of features, we cannot really identify a quantifiable threshold, a magic number metric for how much shared vocabulary or how many shared features are needed to decide relatedness. In this, areal linguistics is like its Janus-twin genetic/genealogical linguistics: the criterion is too many similarities to be a coincidence (cf. Jones 1786, where this criterion formed the bedrock of what became modern historical linguistics). In the case of the Balkans, we also have the advantage of historical documentation.

As noted in §3.3, checklist approaches fail to capture the complexities of Balkan linguistic realities. This is not to say, however, with Andriotis & Kourmoulis 1968: 30, that the Balkan sprachbund is “*une fiction qui n’est perceptible que de très loin*” (‘a fiction that is perceptible only from very far’) and that the commonalities are “*tout à fait inorganiques et superficielles*” (‘completely inorganic and superficial’). To the contrary, Balkan linguistic diversity occurs within the context of a set of structural similarities that comprise a framework of contact-induced change. Moreover, we must distinguish between “superficial” and “surface.” As Joseph 2001a has argued, surface realizations constitute the locus of language contact, and explanations that appeal to typological aspects of universal grammar (including so-called formalist “explanations,” which are, in fact, a type of description) tell us nothing about language contact; so also Aikhenvald 2007a. Also, surface realizations are by no means “inorganic”; they represent convergences that are evidence of the multilingualism that we know existed for centuries and even millennia. Nonetheless, while rejecting the notion that the Balkan sprachbund is a fiction, we must place the differences in the context of the similarities. Where relevant, then, we Balkanize the Balkans by examining certain cleavages, e.g., in future marking (§6.2.4.1) and referentiality (§6.1.2), and discuss the nature of their areality. In so doing, we hope not only to produce a more nuanced picture of the

most famous sprachbund, but also to argue that, like all language change, degrees of convergence can take place with varying speeds. At the same time, however, based on our available documentation, processes that may have been set in motion, or at the very least begun to be reinforced during the middle ages, achieved their current state during the *Pax Ottomana* (and it is telling that this same period is referred to in Bulgarian as *turskoto igo* ‘the Turkish yoke’).¹⁴⁶ Moreover, the effects of the end of that historical period have shown a combination of mutability and resiliency, which, at this early stage, can only be hinted at.

With regard to the quality of features, an interesting view is presented in Heath 1984: 378: “It now seems that the extent of borrowing in the Balkans is not especially spectacular; ongoing mixing involving superimposed European languages vs. native vernaculars in (former) colonies such as Philippines and Morocco is, overall, at least as extensive as in the Balkan case even when (as in Morocco) the diffusion only began in earnest in the present century.” On the one hand, the point that significant change can take place very rapidly concurs with our view that it was precisely during the Ottoman empire that the Balkan sprachbund as we have come to know it was formed.¹⁴⁷ The examples from Morocco and the Philippines, however, all involve lexical items or reinterpreted morphemes rather than morphosyntactic patterns. Moreover, the relationship of the colonial languages to the indigenous is roughly equivalent to that of Turkish to the Balkan Indo-European languages at the time of the Ottoman conquest. While Turkish did maintain a certain social prestige owing to unequal power relations, there is nonetheless a significant difference between recent European colonial settings lasting a century or so and the five centuries of Turkish settlement in the Balkans during which the language became indigenized and members of all social classes were Turkish speakers. To this we can add that the complexity of indigenous power relations prior to conquest is another part of the picture that is easier for us to tease out in the Balkans than in European colonies owing to longer histories of documentation of the languages prior to conquest. It is precisely this background of long-term, stable language contact with significant documentary history that makes the Balkans an interesting and important model for comparing and contrasting with other contact situations.

Related to the issue of the kinds of features is the question of what distribution of features is needed in the group to permit classification as a sprachbund. In particular, must the features identified as diagnostic be present in all of the languages? A. Belić 1936 and S. Mladenov 1939 adduce the piecemeal geographic distribution of some Balkan features as a problem for the sprachbund construct (so also Birnbaum 1968, cited in Schaller 1975: 100, footnote 18); our discussion above addresses that criticism, since a cluster-based approach means that features need not be widespread to be relevant (see also Thomason 2000: 314).¹⁴⁸ Here, too,

146 In a similarly telling example of differences of point of view, the Byzantine Empire is sometimes referred to in Bulgarian as *vizantijskoto igo* ‘the Byzantine yoke.’

147 On rapidity of change, see Mufwene 2004: 203, Dixon 1997, but cf. also Joseph 2001e.

148 Cf. also articles such as Steinke 1999 and Reiter 1999 and the discussion of defining features mentioned in §1.1.1. Note also Masica 1976, whose approach to mapping South Asian areal features was to map each identified feature as far as possible geographically, without, however, examining historicity.

a comparison with genealogical linguistics is instructive. According to Bird's 1982 compilation of the distribution of roots reconstructed for Proto-Indo-European in the various branches of that family – Bird operates with fourteen such branches – only one root, **tēu-* 'swell,' is found in all fourteen branches. Moreover, there are only eight roots that occur in thirteen branches. The numbers of nonisolated roots increases as the threshold for distributions decreases, so there are twenty-eight roots attested in twelve of the branches, and so on. Taken in this light, the absence of postposing for the definite article in Greek, for example, is much less important than its absence in the non-Torlak dialects of the former Serbo-Croatian. Similarly, the distributions of 'have' and 'want' futures take on different significances in different geographic contexts. The point is that it is not the absolute totality of features but rather the cumulative effect that justifies the concept of sprachbund.

Having determined that distribution need not be uniform and that the quest for an absolute minimum of features is unrealistic, the next question is whether certain types of features are more relevant than others. The question of the methodological issue of the sprachbund as a consistently definable unit is that despite the parallel first drawn by Trubetzkoy 1923 between the genetic linguistic family defined by common descent and the areal linguistic league defined by subsequent contact, the manner of selecting the correspondences used to define the latter have not been systematized. Contact phenomena, however, do not have the type of systemic invariance found in phenomena such as regular sound change and shared morphology, which serve as the bedrock of demonstrable genetic/genealogical origin. Contact-induced change, by its very nature, involves a complex ecology of choices among competing systems (cf. Mufwene 2001a). In his original formulation, Trubetzkoy allowed for all types of features (other than those used to define the language family), and, as noted in §4.1, earlier works gave prominence to the lexicon while more recent works give primacy to structural commonalities, e.g., Thomason 2001: 100 who bases her definition on "structural features." Within the group of nonlexical features that are taken to be more important, calques are especially significant.¹⁴⁹ Campbell et al. 1986 use evidence from calques and shared metaphors in Meso-American languages to argue for a Meso-American sprachbund, since some degree of bilingualism is needed for calquing to occur and spread (cf. also Ross 2001 on metatypy).¹⁵⁰ Likewise, intimate borrowings (as discussed in §3.1 and §4.3) are especially significant since they give direct

149 We call calques (defined in §3.2.1.7) "nonlexical" since they involve not so much the transfer/copying of specific lexical material as rather the transfer/copying of conceptual structures; we do in Chapter 4, however, consider both calquing and the transfer of specific lexical material as grist for our mill there, namely the discussion of the effects of contact in the Balkans on the lexical stock, the vocabulary, of the languages involved. We do not see this as a contradiction, as our purpose in this chapter and our purpose in Chapter 4 differ somewhat.

150 Other cases where calques have been used in arguing for a sprachbund include Nuckols 2000 with regard to a Central European area taking in Hungarian, Czech, Slovak, and German, and Gil 2015 with regard to a Mekong-Mamberamo linguistic area (covering part of Southeast Asia – the Mekong river area – and down the Malaysian archipelago into western New Guinea, to the Mamberambo River), where part of the evidence is the shared phraseology of 'eye-day' meaning 'sun.'

evidence of communication between speakers that is not “object-oriented,” not purely for the end of satisfying needs one speaker may have, but rather, in the parlance of §3.1 above, is “human-oriented.”

Since contact is involved, except for cases of shared retentions, defining a sprachbund presupposes an innovation and thus a drift not only away from a prior state but also toward a state that resembles that occurring in another language; in this way, sprachbund phenomena typically involve both convergence on a new type by two or more languages but concomitant divergence from earlier types (which may be preserved in genetically related languages or dialects outside the sprachbund) as well. Sobolev 2011 has claimed that the definition of Balkanisms is circular: Balkan languages have certain features and those features constitute the Balkan sprachbund. This is basically an inaccurate characterization – one might even say caricature. We begin with the fact that a variety of languages are spoken in a multilingual environment over a long period of time. For most of these languages we have the previous stages well attested. If we leave out this starting point, then it is indeed possible to accuse Balkan linguistics of circularity. However, it is precisely the diachronicity at our disposal that enables us to identify convergent features among the Balkan languages (cf. Ilievski 1973, where he states that what is important is not the source of convergence but the fact of convergence).

3.5 Conclusion – Defining “Balkanism”

The discussion in the preceding sections allows us now to address a key remaining question, namely to define what we mean by “Balkanism.” This term, as noted in §2.3.2.1, appears to have been first introduced by Seliščev 1925, who uses it in his subtitle “*une balkanisme ancien en bulgare*” (‘an ancient Balkanism in Bulgarian’); in the article, he does not define the term per se, but rather characterizes it as follows:

Par suite des rapports étroits de culture et de langue qui existent entre des divers peuples balkaniques et des influences mutuelles qu’ils ont eues les uns sur les autres, on constate l’existence chez eux de nombreux phénomènes communs, traits de civilisation ou faits linguistiques. (p. 43)

As a result of the close cultural and linguistic relationships to be found among the various Balkan peoples and the mutual influences that they have exerted upon one another, we can observe the existence of a number of phenomena common to them, both features of civilization and linguistic facts.

phénomènes communs aux langues de l’est et du sud-est de la Péninsule balkanique (p. 50)

phenomena common to the languages of the east and southeast of the Balkan peninsula.

Since then, while many scholars seemingly take the notion for granted as referring to similarities of some sort among Balkan languages and so offer no particular definition, instead often choosing to identify it via examples,¹⁵¹ a few scholars have attempted a definition. For instance, Schaller 1975: 36 gives the following characterization of the term:

In Rahmen einer typologischen Klassifizierung lassen sich nun bestimmte Merkmale des Sprachbaues der einen Sprache mit Merkmalen einer anderen Sprache vergleichen, wobei sich gerade bei den Balkansprachen bestimmte Merkmale typologischer Art feststellen lassen, die nicht nur in einer, sonder gleichzeitig in mehreren Balkansprachen parallel auftreten und daher als “Balkanismen” bezeichnet werden.

Within the framework of a typological classification, certain characteristics of the linguistic structure of one language now allow for a comparison with characteristics of another language, in such a way that just among the Balkan languages can certain characteristics of a typological nature be established which occur in parallel fashion not in only one but rather at the same time in several Balkan languages and thus are designated as ‘Balkanisms.’

Steinke 1999: 80, basing himself on Steinke 1976: 32, defines the notion as follows:

Beim Balkanismus handelt es sich

- (1) um einen gemeinsam ähnlichen sprachlichen Zug der Balkansprachen, der
- (2) in allen Bereichen der Sprachen anzutreffen ist, der
- (3) nicht zum indogermanischen Erbgut gehört, der
- (4) geographisch gesehen, in einer für den Balkanraum typischen Häufung vorkommt, der
- (5) sich gewöhnlich nicht synchronisch, sondern nur diachronisch richtig erkennen läßt und der
- (6) vom jeweiligen Sprachsystem adaptiert wird.

With Balkanisms, it is a matter of:

- (1) a common similar linguistic feature of Balkan languages, which
- (2) is to be found in all regions of the language, which
- (3) does not belong to the Indo-European inheritance, which
- (4) viewed geographically, occurs in a typical clustering for the Balkan region, which
- (5) usually can be properly recognized not synchronically but only diachronically, and which
- (6) is adapted by each respective language system.

151 Hinrichs 1999b: 430 explicitly says, for example, that “*eine Lösung des fast nur noch scholastischen Problems einer endgültigen ‘Balkanismus’-Definition nicht abzusehen ist*” (‘a solution to the almost scholastic problem of a definitive definition for “Balkanism” cannot be foreseen’).

Finally, Joseph 1983a: 247 provides these glosses and comments on the term and its use: This term [“Balkanism”] can be taken in (at least) four ways:

- (1) Any similarity between Balkan languages
- (2) Any similarity between Balkan languages that is an innovation vis-à-vis their common ancestor language (in this case, Proto-Indo-European)
- (3) Any similarity between Balkan languages that is due to language contact
- (4) Any similarity between Balkan languages that is unique to the Balkans.

and then notes that of these, (1) “is the least interesting,” because it does not distinguish different sources (genetic/genealogical relationship, universality, etc.) for similarities between languages, and that the others are attempts to make some further specifications to (1).

What these definitions have in common is the obvious mention of similarities across languages, as well as some concern for the origin and/or distribution of the similarities. Where they differ is in specifying, or not, the source of the similarities; the definitions of Schaller and Steinke, for instance, do not distinguish explicitly between similarities due to independent innovation in each language and those due to language contact (though in practice the authors surely would and in fact do).

It may well be that, as Hinrichs says (see footnote 151), that a conclusive definition is not in sight for this “scholastic problem” (i.e., one akin to determining how many angels can dance on the head of a pin), but nonetheless we make our own attempt here. From the discussion in this chapter, it should be clear that for us, a key element in understanding this notion is language contact, and more particularly a specific type of language contact. Since we refer to the contact situation in the Balkans and the results of contact with the designation “Balkanization,” we could, somewhat tautologically and thus perhaps not too revealingly, define “Balkanism” as a linguistic feature that resulted from some degree of Balkanization. In a more pointedly helpful way, though, we define “Balkanism” as follows:

a linguistic feature resulting from the bi- or multilingualism born of language contact that is common to two or more unrelated or only distantly related languages within the Balkans.

A few comments on this definition are called for. First, by locating the elements of this definition in the Balkans, we are saying what might be an obvious point but one that bears mention, namely that “Balkanisms” are a feature of the Balkans; shared features found in other sprachbunds are entitled to their own name (e.g., “South Asianisms”). It may well be that a generic term for any such shared features in any linguistic area would be useful, and an extension of “Balkanism” into that function could well be envisioned for that (just as “Grimm’s Law,” while referring to a particular set of sound changes between Proto-Indo-European and Proto-Germanic, is sometimes used to characterize wholesale consonant series shifts found elsewhere).

Second, by referring to “unrelated or only distantly related” languages, we are excluding features found only in, say, Macedonian and Bulgarian or Aromanian and Romanian. But the exclusion is not necessarily a matter of principle but rather is one of convenience, since with any such characteristics, it is not always clear that they are contact-induced as opposed to part of the languages’ inheritance from a common proto-language. If contact can be demonstrated to be the cause, then the feature would certainly qualify as a “Balkanism,” with the vagueness of “only distantly related” coming to the rescue in such a case.

Third, by talking in terms of a “linguistic feature” in this definition, we have avoided restricting the notion to only structural characteristics. One reason for doing this is that some of the features we discuss of a phonological nature, especially in Chapter 5, are not so much a matter of phonological structure as of phonetic realization, i.e., pronunciation. While one could argue that any aspect of the realization of sounds can have repercussions at an abstract level (e.g., in that the content of a phoneme is altered or an opposition is neutralized), we prefer to focus on the feature itself and not necessarily on systemic consequences.¹⁵²

Another reason for using a broader term like “linguistic feature” has to do with the question of how lexical items fit in with any conception of “Balkanism.” As noted above in §3.1 and §3.2.1.7, shared lexical items per se are not necessarily revealing of contact or of a special type of relationship or interaction among speakers of different languages of the sort associated with structural convergence, since words can spread without real contact (as in the case of learnedisms) and since even very casual contact involving little or no bilingualism can lead to borrowing (especially in the case of “concrete,” i.e., need-based and culturally based, borrowings). Thus, one might want to exclude lexical parallels from any conceptualization of “Balkanism,” so that loanwords in and of themselves would not be “Balkanisms.”

However, as with the “unrelated language” criterion discussed above, any such exclusion would be more a matter of convenience than principle. That is, as emphasized above in §3.1 and developed further in Chapter 4, certain *classes* of lexical items, e.g., discourse markers (see §4.3.4), are highly diagnostic of intimate, intense, and sustained contact, and such is true also for calques (loan translations); by their very nature (see §3.2.1.7), they signal some degree of knowledge of the other language, and in the case of highly colloquial sorts of calques, especially phraseological ones, as opposed to those restricted to higher registers, they certainly point to intimate contact. Such lexical parallels, therefore, we would want to be able to call Balkanisms.

In the case of concrete vocabulary, however, the real issue is that one cannot tell for sure whether they are indicative of and derivative from the sort of language contact that can yield structural convergence. Such words can and do diffuse within the context of a sprachbund and so they may well have spread in the intense contact

152 Inasmuch as linguistic reality for speakers consists of the forms and uses of the language, we suspect that speakers care less about systems than linguists do.

we posit as essential for the formation of a sprachbund, but they are not diagnostic and thus cannot be reliable indicators; one would not be inclined to think of an intense sort of contact relationship if concrete loans were all that one could find in common among a number of areally related languages. However, if there are positive indications of a sprachbund for other reasons, as is the case within the Balkans, we see no reason not to speak of “lexical Balkanisms,” even for concrete lexical items.¹⁵³

Further, this definition excludes features that are a matter of independent but parallel development in the languages in question (for whatever reason, e.g., universality or accident) and those that are due to common inheritance from a proto-language.¹⁵⁴ On the other hand, this definition includes shared retentions (in a conservation zone), features that have spread from one language to another via (so-called) borrowing (including copying, in the sense of Gołąb 1976 and Johanson 2002, and calquing, and such; see §3.2.1.7), and features that are part of a speaker’s linguistic repertoire that carry over into or are caused by his/her learning of another language (so-called substratum features and reverse interference; see §3.2.1.3–4).

Finally, this definition also means that some of the classifications that have been given for Balkanisms, such as the distinction in Schaller 1975: 100–101, 192 between “primary” and “secondary” Balkanisms, represent unnecessary overclassification.¹⁵⁵ For Schaller, the number of languages showing a particular feature was the basis for distinguishing “primary” – appearing in three or more languages in the Balkans – from “secondary” – appearing in only two languages. It is not at all clear that these numbers in practice represent a meaningful criterion: one can ask, e.g., whether two dialects of the same language (e.g., Geg and Tosk Albanian) would count as two “hits” or as one. Moreover, given our cluster approach to the sprachbund, the number of languages in which a feature is found is not in and of itself diagnostic or revealing, a point made in various reviews of Schaller 1975 (e.g., Bevington 1977; Joseph 1987a).¹⁵⁶

153 The term “lexical Balkanism” can be found in the literature: Birnbaum 1965: 42–43, for instance, refers to “*die Anzahl der hier anzutreffenden lexicalischen Balkanismen*” (‘the number of lexical Balkanisms being found here’) in Serbo-Croatian. Some scholars, however, stop short of actually using it: Schaller 1975: 172 talks about “*Übereinstimmungen der Balkansprachen im lexikalischen Bereich*” (‘correspondences of the Balkan languages in the lexical domain’) and “*gemeinsam Lehnwörter*” (‘common loanwords’) which are “*neben den als Balkanismen behandelten Besonderheiten aus dem lautlichen, morphologischen und syntaktischen Bereich*” (‘near the particularities from the phonological, morphological, and syntactic domains treated as Balkanisms’), and Skok 1971, 1972, 1973, 1974 uses the phrase “*Balkanski turcizam*” (‘Balkan Turkism’) to label Turkish loanwords in Serbo-Croatian that show a wide distribution in the Balkans, e.g., *džámija* ‘mosque.’

154 This is relevant only when the languages in question could be genetically related; thus within the Balkans, it is an issue for features shared between Greek and Albanian, for instance, but not Greek and Turkish.

155 Though it might seem relevant, given the labels used, the distinction Hinrichs 1999b: 431 makes between “*Makrobalkanismen*” (‘macro-Balkanism’) and “*Mikrobalkanismen*” (‘micro-Balkanism’) is not relevant here, as it is not based on numbers and counting but rather refers to broad, over-arching structural characteristics. See also §5.2 footnote 7, for some further discussion.

156 See also §3.3, where the dialect-versus-language issue proves problematic in another way for a numerological approach. Cf. also Duridanov 1983, cited approvingly in Asenova 2002: 292–

In this chapter, therefore, we have laid out what we see as a sensible framework for discussing the intriguing features that link various of the languages of the Balkans in the long-recognized special way that the designation “sprachbund” implies. In the chapters that follow, details about these various features, grouped by grammatical component – lexicon, in Chapter 4, phonology, in Chapter 5, morphosyntax, in Chapter 6, and syntax, in Chapter 7) – are discussed against the backdrop of the key notions developed here.

293, who also mentions Schaller 1975: 123, Steinke 1977, and Bernštejn et al. 1963. Duridanov distinguishes “total” and “partial” Balkanisms, the former appearing in “all the languages” of the Balkan sprachbund and the latter appearing in “three of the languages.” This formulation represents a combination of the past and current approaches to Balkan linguistics practiced in Bulgaria and by some Bulgarians in other countries, according to which all of Balkan Slavic is “Bulgarian” and all of Balkan Romance is “Romanian” and, as in Sandfeld’s day, the relevant dialects of Romani, Judezmo, and Turkic are excluded. Asenova 2002: 293 rejects bilingual correspondences as Balkanisms, but see §3.4.2.1 for our discussion of the number of languages needed in determining a sprachbund.