The typology of Balkan evidentiality and areal linguistics

Victor A. Friedman
University of Chicago

An explanation of evidentiality in the Balkan languages must take both areal and typological factors into account. There are two types of complexes of grammaticalized meanings in Balkan evidentials: confirmative and nonconfirmative. Presence of a marked confirmative implies presence of a marked nonconfirmative, but marked nonconfirmatives can occur in a system without marked confirmatives. The complex of meanings associated with marked nonconfirmativity has a functional basis and occurs in widely separated, unrelated languages, but the grammaticalization of those meanings into verbal paradigms in the Balkans is due to areal factors. The distinction between meanings calqued onto existing forms and the creation of new forms to convey new content does not map isomorphically onto the distribution of semantic-pragmatic isoglosses. Moreover, the geographic distribution of evidentiality in Balkan Slavic argues for a sociolinguistically based explanation for the spread of the phenomenon from Southwestern Macedonia and Eastern Bulgaria. Analyzing the rise and spread of grammaticalized evidentiality in the Balkans, therefore, requires an account that considers sociolinguistic factors and dialectal distribution at various levels of the grammar. The picture that emerges is one in which form, content, narrative structure, and social factors all have roles to play, and a typologically informed areal explanation appears to be the most adequate.

1. Introduction

In comparing linguistic phenomena, when two or more languages display an observable degree of similarity there are three major sources of explanation: 1) descent from a reconstructable common ancestor, 2) contact between two speech communities that had developed previously in isolation from one another, and 3) the workings of natural tendencies in human languages that result in similar phenomena. The first and second of these – the former frequently called “genetic”
or "genealogical", the latter "areal" – both involve historical accounts (cf. Hamp 1977:279). The third explanation, which can use the tools of typology to deduce the rules of universal grammar (cf. Comrie 1981:30–31), is, in Hamp's terms, achronic rather than diachronic, although Croft (1990:203–259) illustrates more recent applications of typology to diachronic questions. With regard to formal-generative explanations, Joseph (2000:24–25) argues that surface realizations, not generative rules, are the points of contact between speakers that result in the spread of features the accumulation of which are said to characterize a Sprachbund. Thus, areal versus typological explanations emerge as a key issue in Balkan linguistics. Grammaticalized evidentiality is often cited as an example of the influence of Turkish on the structure of Balkan Slavic and/or of Albanian (Conve 1910/1911; Golab 1960; Konesi 1965:148) although accounts imputing native origin also exist (Demiraj 1971; cf. also Friedman 1978; as Van Wijk 1933; Huntley 1979). However, various general works on Balkan linguistics either fail to mention this category or, at best, treat it as a bilateral correspondence between Turkish and Balkan Slavic or Albanian, as in Sandfeld 1930:119–20; Schaller 1975:79, 94; Feuillet 1986:67; Asenova 1989:203–204; and Demiraj 1994:169–171. In fact, the category also occurs in other Balkan languages and shows different manifestations in various dialects. In this paper I shall address the expression of evidentiality in the Balkan languages as a case study of a contact phenomenon, specifically a Balkanism, that must take both areal and typological factors into account.

Given the recent increase of interest in evidentiality (e.g. Guentcheva 1996; Johanson & Utos 2000; Mushin 2001; Squarzini 2001; Alexander 2002; Alkhnawal & Dixon 2003, to name just a few), as well as the varied phenomena to which the term has been applied and the lacunae in the history of its study, I shall address the term "evidential" itself before continuing to the exposition and analysis of areal versus typological explanations of the Balkan evidentiality.

2. General considerations

The year 2002 marks the fifty-fifth anniversary of the first appearance of the term "evidential" (for Kwakiutl, Boas 1947), the fiftieth anniversary of Lunt's (1952) description of Macedonian as having a 'vouched for'/'distanced' opposition, and the forty-fifth anniversary of the application of the term "evidential" to a Balkan language. In his ground-breaking article, Jakobson (1957) applied "evidential" to a Bulgarian example in which he glossed the old perfect (PI) (without auxiliary) za-mina as 'it is claimed to have sailed' and the simple preterite (PD) za-mina as 'I bear witness it sailed,' Jakobson referred to Lunt (1952). The year 2002 is also the thirty-fifth anniversary of Aronson's (1967:87) observation that 'witnessed' does not capture the actual meaning of the Bulgarian simple preterite, and the twenty-fifth anniversary of Aronson's (1977) arguments for a redefinition of "mood" and "status" such that the relevant Balkan verbal category is a type of status rather than evidential, taking Jakobson's definition of the term (see also Friedman 1977:55–81). In addition, 2002 marks the twentieth anniversary of Friedman 1982a, in which I demonstrated that "reported" could not be taken as the literal meaning of the Balkan old perfect (past indefinite) without the third person auxiliary. Guentcheva (1996:13), among others, has pointed out that Jakobson's original definition of evidential is too closely tied to "reportedness", which is only one of the possible meanings generally associated with the evidential complex, and she suggests using Lazard's "mediatif". Johanson (2000:61) suggests "indirectivity". Mushin (2001:21) argues for a distinction between a "narrow" and a "broad" interpretation of evidentiality. The narrow interpretation is limited to literal source of evidence, whereas the broad interpretation includes the attitude of the speaker – which is most frequently but not necessarily the result of source of evidence. The fact that grammaticalized evidentials in the Balkans (and some but not all other regions where evidentials are grammaticalized) encode speaker attitude relates them to epistemic modals. This in turn sometimes leads to a confusion between evidential and epistemic modality as exemplified by Palmer (1985:56), whose second-hand accounts of Tajik, Turkish, and Abkhaz completely miss the fact that these languages (and others) have both epistemic modals and a distinct verbal category that involves the subjective relation of the speaker to the narrated event, i.e. the opposition "confirmative/nonconfirmative" in Aronson's (1967) terms, "epistemological stance" in Mushin's (2001:xii), Jacobsen (1986:5), without, unfortunately, giving a citation, reproduces exactly Aronson's (1977:14) insight (using Jakobson's terminology) that the category in question involves the subjective relationship of the participant in the speech event to the narrated event. Jacobsen (1986:3) traces the history of the study of evidentiality only as far back as Boas (1911:42–43, 124–125, 246–247, 348–49, 443, 496), although in fact the earliest attested reference to the relevant grammatical phenomenon dates to the eleventh century. Moreover, this reference is in a description of Turkic grammar and therefore deserves citation here:

The difference between these two forms [in -d and -m] is that DY [-d] on preterite verbs indicate that the action occurred in the presence of the speaker. The action was verified by its occurrence in his presence. For example, if someone one says bardı the meaning is, 'He went and I saw him go with my own eyes.' M S [-m] on the other hand, indicate that the action occurred in the absence of the speaker. Thus ol barmış 'He went but I did not see him go'; ol kalmış 'He came but I did not see him.' This is a general rule holding good for all preterite verbs ...

(AL-Kaşgārī, cited in Dankoff 1982:412)
Al-Kašgari's formulation remains fairly typical for oversimplified accounts of evidentiality even today:6 Although I continue to agree with Aronson (1991:129–130), who writes: "The notional meaning of "evidential" is easily derivable from the grammatical category of status...". I shall use the term "evidential" in this article since I will be dealing with precisely those notional meanings of "status" that are elsewhere understood under this term.

Thorough accounts recognize various uses (contextual variant meanings) for evidentials, which in the case of the Balkans can be grouped as in Table 1:

<table>
<thead>
<tr>
<th>A. CONFIRMATIVE</th>
<th>B. NONCONFIRMATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. witnessed</td>
<td>3. felicitous</td>
</tr>
<tr>
<td>a. confirmable by speaker</td>
<td>a. neutral, resultative</td>
</tr>
<tr>
<td>b. confirmable by speaker and addressee</td>
<td>b. reported, hearsay</td>
</tr>
<tr>
<td>(general, historical, gnomic, personal facts)</td>
<td>c. deduction, inference</td>
</tr>
<tr>
<td>2. unwitnessed</td>
<td>4. infelicitous</td>
</tr>
<tr>
<td>a. dubitative (irony, sarcasm)</td>
<td>b. admimistrative (surprise)</td>
</tr>
</tbody>
</table>

The exposition which follows explicates the meaning and organization of Table 1.

3. Balkan Slavic

Not all the Balkan languages show all the possible meanings given in Table 1. I shall refer to a system that encodes all the meanings listed in Table 1 as a Type A system. Any system that encodes the meaning listed under A also encodes the meaning listed under B, but not vice versa, i.e. there exists an implicational hierarchy: Systems that only encode the meanings listed under B will be called Type B systems. I shall begin with Macedonian, which has a Type A system. In many respects, Macedonian is the most Balkanized of the Balkan languages insofar as it displays the most consistent realization of morphosyntactic convergence features (Joseph 1983). Standard Macedonian, which in this respect represents its west-central dialectal base (Veles-Prilep-Kičevo-Brod), has a basic opposition between two inherited past tense forms (each having a subordinate aspectual opposition aorist/imperfect). These two past tenses are traditionally labeled "past definite" (PD) and "past indefinite" (PI), although, as I have shown (Friedman 1977:21–81), the PD is marked for confirmativity, while the PI is an unmarked past. Thus, for example, although the PD ti igrale 'you were playing' usually refers to a definite time in the past, whereas the PI ti si igral 'you were/have been playing' will normally lack a definite time reference, either form can actually refer to either a definite or indefinite time in the past. That which differentiates the two forms is the degree of commitment of the speaker to the truth of the statement (see examples (1)–(7) below).

The PI is descended from the Common Slavic perfect, which used the auxiliary 'be' (Modern Macedonian 2sg present si) plus a resultative participle in -l. Both Standard Macedonian and the western and southern dialects now have a new perfect using the auxiliary ina 'have' plus the invariant neuter verbal adjective of the main verb (e.g. 2sg. ti inal igrao 'you have played/been playing'). This perfect is marked for resultativity (see Friedman 1977:82–113). In opposition to the markedly confirmative PD on the one hand, and the markedly resultative perfect in 'have' on the other, the chief contextual variant meaning of the PI consists of the nonconfirmative complex, although, like the unmarked past, it is also the choice in contexts where complete neutrality is required. While the confirmativity of the PD normally implies that the speaker has first-hand knowledge of the narrated event, this form can also be used for unverified, reported, and other non-first-hand knowledge as long as the speaker has reason to be convinced of the truth of the statement. The following examples illustrate the various facts and possibilities:

1) No podocna se slučja raboti za koi ne [Macedonian]
   but later...PD works about which NEG
   znaev.
   I.knew.PD
   'But later things happened which I didn’t know about.'
   (Nov. Macedonija 19.VI.74:5)

2) Zara: Blaže bil vo Moskva.
   Blaže was.PD in Moscow
   'Blaže was in Moscow.'
   Kosta: Da, beše.
   yes he.was.PD
   'Yes, [I know] he was.' (10/86)

3) Mu se javiv na vukjo mi. Ne beše
   him IN I.called.PD to uncle me NEG was.PD
   doma, na plaža bil.
   at.home on beach was.PD
   'I called my uncle. He was not at home, [apparently] he was at the beach.'
   (8/92)
(4) Ne veroam deko joj go napravi! [*napravi](Macedonian) 
that he it did.PI did.PD

'I don’t believe that he did it.' (Friedman 1978:110)

(5) Za-toa što forma-tu BISE REŠENO odozvara na [Macedonian]
for-that what form-the was.PD decided answers to
SE REŠI, i.e. na minato-to opredeleno vreme, koe go
in decided.PD i.e. to past-the definite tense which it
upotrebljavamo za označavanje dejstva što se vrše ili se
we.use for indication actions what in performed.PRI in
izvršile vsopredelen moment vs minato-to.
accomplished.PRI in definite moment in past-the
'It is because the form had been decided corresponds to was decided,
i.e., to the past definite tense, which we use to indicate actions that
were performed [PI] or accomplished [PI] at a definite time in the past.'
(Nova Makedonija 19.XII.84.11)

(6) Ti si bil vo Amerika. [Macedonian]
you AUX.2SG was.PD in America.
'You are/were/have been in America.'

(7) Kade bil majstor-ot! [Macedonian]
where was.PD master-the
'Where “is/was the boss’"

Because the use of the PD for witnessed past acts is uncontroversial, I have not
given any examples. Example (1) illustrates the use of the PD for something that
the speaker clearly did not witness but is now certain happened. In example (2), Zuza
chooses the PI because she was not in Moscow herself, but Kosta, who was also
not in Moscow, responds in the PD because he accepts the fact as generally known
(pace Lunt 1952:93). In example (3), both statements about the uncle are based on
the same telephone call. The first statement is in the PD because the speaker is per-
sonally convinced that his uncle would have come to the telephone had he been at
home. He chooses the PI for where his uncle was said to be, however, since neither
he nor his interlocutor could actually know for sure. Example (4) illustrates the
fact that a PD cannot be used in a context that contradicts the notion of confirma-
tion. The one exception to this—which is the reason the asterisk is in parentheses—is
if the statement is infelicitous (Austin 1962:14), i.e. if the speaker really is con-
vinced that he did it, but is expressing dismay, amazement, etc. (cf. the English I
can’t believe I ate the whole thing). Example (5) illustrates the fact that the PI is the
form of choice in such completely neutral contexts as definitions. Neutral usage

is more common in the first and second persons for pragmatic reasons. Example
(6) can illustrate any of the nonconfirmative uses of the PI depending on the con-
text in which it occurs. The illustrative contexts pertain to each of the types given in
Table 2:

Table 2. Possible interpretations of example (6)

(a) The speaker has seen the addressee in America and is making a general or resulative
statement of fact.

(b) Someone told the speaker, who then tells the addressee that his trip has been reported.

(c) The speaker sees the addressee wearing a Chicago Bulls T-shirt and makes a deduction.

(d) The addressee tells the speaker that he was in America and the speaker does not be-
lieve him. Another possible context would be if the addressee calls the speaker on the
telephone and says he is in America at the moment of speaking and the speaker retorts
with example (6). In this context, the correct English translation would be a present
tense form.

(e) The speaker is in America and unexpectedly meets the addressee. Here, again, the cor-
correct English translation would be present tense. It is important to note that in this last
type of usage, the nonconfirmativity is based on the fact that surprise is something
which requires a state of mind in which the speaker would not have been willing to
confirm something until the moment of its (unexpected) discovery (Friedman 1981). 8

Example (7) illustrates the fact that uses of the PI with apparent present meaning,
i.e. in contexts where the correct English translation will be present tense, must
always refer to an ontologically past speech event or state of affairs. If a customer
walks into a store and does not see the boss and is surprised by that discovery, he
cannot use (7) as his question, because he is expressing surprise about the present
moment of discovery itself and not about the pre-existing state of affairs that oc-
casioned it. If he were expressing astonishment at what someone else had told
him, then he could use the PI, since there would be a previous event to which the
form could refer. In reported or dubitative usage, there is always a real or puta-
tive statement that is being referenced or refuted, while in admirative usage there
must always be a pre-existing state that the speaker has just discovered. Thus, the
question in example (7) is limited to some sort of past tense meaning.

In terms of evidentiality, mutatis mutandis and ceteris paribus, the Bulgar-
ian situation is basically the same as the Macedonian situation with regard to
markedness relations and semantics. Although Bulgarian differs from Macedonian
in certain morphological and paradigmatic manifestations of evidentiality, these
differences do not affect the basic arguments. Thus, for example, both languages
have markedly nonconfirmative pasts (i.e. paradigms showing only reported, in-
ferred, dubitative, and admirative meanings) using the old resulative participle as
an auxiliary, but in standard Macedonian the new marked nonconfirmative uses
imal ‘have’ plus the neuter verbal adjective of the main verb, whereas standard
Bulgarian uses bil ‘be’ plus the I-form of the main verb, e.g. Macedonian inal
napraveno Bulgarian *bil napravil* 'he has/had done.' The only other difference that need be noted here is that the claim in traditional Bulgarian grammar according to which the old perfect with the third person auxiliary is a perfect, while the absence or the auxiliary signals 'reportedness,' is belied by both written and spoken usage, of which example (8) is typical:

(8) *Imalo e edin pop, i go ozenil.* [Bulgarian]
    had.PT AUX.3SG one priest and him married.PT
    'There was a priest, and he married him (to her).'

The speaker was recounting a legend from the 17th century about a wealthy Aromanian boy who had run off to Korçë (Albania) with a poor Macedonian girl and had gotten married and built a church there. From the context, it is clear that the source of all the information is a single report, and yet the auxiliary is present for the backgrounded information and absent for the foregrounded information. This is entirely in keeping with Fielder's (2002b and references therein) arguments concerning the fact that auxiliary omission in Bulgarian encodes the opposition foregrounded/backgrounded rather than some form of evidentiality.

For Serbian, including the Balkan Serbian (Torlak) dialects of southeastern Serbia, it is clear that, as in the rest of Slavic, the descendant of the Common Slavic perfect (corresponding to the Macedonian and Bulgarian PI) has become the unmarked past (typologically, this agrees with developments in e.g. Romance and Germanic). It is sometimes claimed that the synthetic pasts (corresponding to the Macedonian and Bulgarian PD) encode a meaning such as "witnessed" (Samilov 1957).

4. Albanian

Just as Macedonian exemplifies a Type A system with a confirmative/nonconfirmative opposition in which there is a marked confirmative member, so, too, Albanian exemplifies a Type B system in which only nonconfirmativity is marked. In such a system, the marked forms are never neutral but either reported, deduced, dubitative or confirmative. In traditional Albanian grammar, the name of the series of paradigmatic sets is *habitore* from *habi* 'surprise'; this corresponds to the French *admiratif* introduced by Dozon (1879:226–227) as a translation for Kristorforidhi’s Greek *apodóketo* ‘unexpected’, all of which refer to the most common use of these forms. Hence, the most common meaning gives its name to the entire set of paradigms. Table 3 gives an illustration of the relevant forms, which are derived historically from an inverted perfect and pluperfect.

<table>
<thead>
<tr>
<th>Present</th>
<th>Nonadmirative</th>
<th>Admirative</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td></td>
<td>paska</td>
</tr>
<tr>
<td>Perfect</td>
<td>ka pasur</td>
<td>paska pasur</td>
</tr>
<tr>
<td>Imperfect</td>
<td>kish(te)</td>
<td>paskësh</td>
</tr>
<tr>
<td>Pluperfect (imperfect)</td>
<td>kish(te) pasur</td>
<td>paskësh pasur</td>
</tr>
<tr>
<td>Double Perfect</td>
<td>ka pastë pasur</td>
<td>paskësh pasur</td>
</tr>
<tr>
<td>Double Pluperfect</td>
<td>kish(te) pastë pasur</td>
<td>paskësh pastë pasur</td>
</tr>
</tbody>
</table>

As was just indicated, the Albanian meaning of the admirative can correspond to any of the marked nonconfirmative meanings found in Macedonian. Thus, for example, the Albanian equivalent of example (6) would be (9a) or (9b):

(9) a. *Ti qenë n’Amerikë.* [Albanian]
    you are.ADM.2SG in America
    'You are in America.'

b. *Ti paske qenë n’Amerikë.*
    you have.ADM.2SG been.PT in America
    'You have been/were in America.'

A crucial difference between the Albanian evidential system and the Balkan Slavic is that while a Baltic Slavic evidential always has some sort of past reference, even when the apparent meaning is present, the Albanian system has a true present. This is illustrated in the difference between (9a) and (9b). Moreover, in Albanian it is possible to have a sentence like (10):

(10) *Ku qenka mjeshtr-i?* [Albanian]
    where IS.ADM.3SG boss-the
    'Where is the boss?'
    (Rilindja 8.VI.82.8)

Additional evidence for the analysis of Macedonian admirable meaning as always having a past reference is to be found in the fact that Albanian permits past admiratives to describe newly discovered pre-existing states, as in example (11), which was uttered by a Kosovan colleague upon taking a sip of a local unaged grape brandy (raki) whose quality was reputed to have declined in recent years:

(11) *Paska qenë e mirë!* [Albanian]
    have.ADM.3SG be.PT AGR good
    'Why, it turns out that it is good!'
    (6/01)
5. Balkan Romance

The situation in Balkan Romance represents a divided picture. For the most part, Aromanian does not have marked evidential forms, but the Frasieriote dialect of Gorna Belica (Bela di Suprâ), on the Albanian border northwest of Struga in southwestern Macedonia, has borrowed the Albanian third person administrative marker -ka, which is attached to a calqued masculine plural participial base, as illustrated in Table 4 (after Marković 1995:80, cf. also Friedman 1994a):10

Table 4. Aromanian and Albanian present and perfect indicatives (3sg 'have')

<table>
<thead>
<tr>
<th>Nonadministrative Aromanian</th>
<th>Albanian</th>
<th>Administrative Aromanian</th>
<th>Albanian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>ari</td>
<td>ka</td>
<td>avushka</td>
</tr>
<tr>
<td>Perfect</td>
<td>ari avushi</td>
<td>ka posar</td>
<td>avushka avush</td>
</tr>
</tbody>
</table>

The Aromanian construction involves the use of a native masculine plural participial base (usually imperfect, but aorist with precisely those verbs that have a distinctive, suppletive aorist participial base in Albanian). Such a base occurs nowhere else in the language and is a clear calque on the way Albanian forms its administrative.11 Moreover, the Frasieriotes of Bela di Suprâ migrated to that village from Myzeqe in central Albania at the beginning of the twentieth century and were, at that time, all bilingual in Albanian. After the Balkan Wars (1912–1913), Bela di Suprâ was part of what became Yugoslavia and the Republic of Macedonia (except during 1941–1944, when it was part of Albania), and knowledge of Albanian was not maintained by subsequent generations, who were, instead, bilingual in Macedonian. This shift in bilingualism is reflected in the behavior of the Aromanian administrative. Like the Albanian administrative, and unlike the Balkan Slavic evidentials, the Aromanian administrative does not occur in connected narratives, i.e., while it is perfectly normal for a Balkan Slavic narrative to be related entirely in the PI (although for stylistic reasons other tense forms are likely to occur in lengthier narratives), neither Albanian nor Aromanian narratives would use more than the occasional administrative. On the other hand, according to my consultants, the Aromanian administrative would not be used in questions. Thus, for example, in (12), which is the Aromanian equivalent of (7) and (10), a plain present is required:

(12) [Aromanian]
    *Iu iesti*/fusa majstor-lu?
    'Where is the boss?'
participle of the main verb. Examples (14) and (15), which both use conditional markers and are translations from Bulgarian (given here with the original), are illustrative of the markedly nonconfirmative evidential meanings of the present and past presumptive, respectively:

(14) imi spuse ca tă băiat-ul ar fi având în [Romanian] me said how that boy-the CO.AUX.3SG be having in Bulgaria un frat ofițer
Bulgaria a brother officer
cuva mi, cе tuj momcă uVALO u vas, u Bălgiaria, [Bulgarian] told me that this boy had.PI by you in Bulgaria brat činovnik
brother official (in Bulgaria)
hec told me that this boy had a brother, an official'

(15) Bai Ganiu imi ceru să-i fac cinste, deoarece [Romanian] B. G. me sought su-him make treat since pe drum aș fi fumat din tutun-ul lui. on road CO.AUX.1SG be smoked from tobacco-the him Băj Ganiu poștea aș da poerja, poenča îz [Bulgarian]
Păța a învăță ce negoi-va pața. road-the AUX.1SG smoked,pt from his the tobacco
'Bai Ganiu wanted me to treat him, since on the road I had smoked some of his tobacco.'

In these examples, the Romanian presumptive corresponds to a Bulgarian and Macedonian PI and a Turkish mi̇ş-past but a plain indicative in the Albanian translation. I have demonstrated elsewhere (Friedman 1982b) that Albanian present admiratives correspond to the Balkan Slavic PI and the Turkish mi̇ş-past in only about 10% of the cases in comparable translations, and do not correspond at all to Romanian presumptive usage, despite the fact that the presumptive and the admirative encode the same complex of marked nonconfirmative meanings (Friedman 1998). The Romanian presumptive can be used in questions, but is not used for connected narratives. There is also some debate concerning whether the presumptive constitutes a separate paradigm or a syntactic construction, especially in the case of the past presumptive which is identical to the subjunctive, future, orconditional perfect. Although Dimitriu (1979:269–271) argues that evidential uses of the past presumptive are clearly different from other modal uses, evidential-like uses of similar modals in Western Romance (cf. Squartini 2001) can be added as an argument against the independent status of the Romanian past presumptive. On the other hand, however, none of the Western Romance languages possess a construction equivalent to the Romanian present presumptive, nor do they use ‘be’ as an auxiliary in active constructions (Manoliu–Manea 1994:270).

Before moving to our last Balkan Romance example, it is worth citing a Slavic modal paradigmatic set that appears to calque the Romanian presumptive, namely the probabilistic mood of the Bulgarian village of Novo Selo, in northwestern Bulgaria across the Danube from Romania and a few kilometers from the Serbian border (Mladenov 1967:108–116). In this dialect, the ordinary future is formed, as in Balkan Slavic in general, by means of a preposed invariable particle derived from a verb meaning ‘want’ (хă) plus a conjugating main verb. The non-Balkan Serbian future type—which is formed by means of a postposed conjugating clitic auxiliary of the same origin (‘want’) attached to a stem based on the infinitive—has been retained in this dialect but reinterpreted as an nonconfirmative evidential with present, future, and past tenses, as illustrated in Table 6:

<table>
<thead>
<tr>
<th>Present</th>
<th>Indicative</th>
<th>Probabilistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>2sg</td>
<td>3sg</td>
</tr>
<tr>
<td>gledăm</td>
<td>gledăs</td>
<td>gledă</td>
</tr>
<tr>
<td>gledăs</td>
<td>gledă</td>
<td>gledicăm</td>
</tr>
<tr>
<td>gledă</td>
<td>gledă</td>
<td>gledicăm</td>
</tr>
<tr>
<td>gledićă</td>
<td>gledićă</td>
<td>gledićă</td>
</tr>
</tbody>
</table>

The Table 6. Novo Selo probabilistic present and future sg conjugation of ‘see’ and 3sg PI

The Novo Selo presumptive has the same complex of meanings as other marked nonconfirmatives, and according to Mladenov (1967:111) is especially common in questions as in example (16):

(16) Ĉa ikā-ĉā ĝjatra ĉū-grad | [Dialectal Bulgarian] ŽU go-FO-2SG tomorrow in-town ‘Will you go to town tomorrow?’ 
(Mladenov 1967:111)

Located in the border region of Bulgarian and Torlak (Balkan) Serbian, the population of the village is mixed Romanian and Slavic, and much of the Slavic-speaking population is said to have migrated from Macedonia (Mladenov 1967:8–9). In the past there was debate over whether to assign the dialect (and, by implication, the territory) to Serbian or Bulgarian. I shall return to this case below. Our final Balkan Romance example is the Istanbul dialect of Judeo-Serbian, which calques the Turkish mi̇ş-past by means of the pluperfect, as in example (17):

(17) Dos ermanos eran, uno salýu doktor diği, el otro [Judeo-Serbian] two brothers were one became doctor dentist the other salýu dahiška, despuñ s' ariya etcho doktor de bebés became internist afterwards s' had become doctor of babies
There were two brothers, one became a dentist and the other became an internist, afterwards he became (lit. had become) a pediatrician.'

(Varlo 2001:91)

According to Varol, the narrator knew the two brothers and had watched them become doctors, but he had only heard about the second brother's change of specialization. Were the narrative in Turkish, the ordinary choice for the last clause would be the miy-past. The narrator attempts to render the difference by using a pluperfect, thus iconically indicating narrative distance by means of an anterior tense form.13

6. Romani

Like Aromanian, Romani does not generally distinguish evidential categories, but according to Kostov (1973), the dialect of Sliven in Bulgaria has borrowed the -I of the I-participle as a marker of evidentiality (much like the Aromanian phenomenon described above) as in example (18), given in both Romani and Bulgarian:

(18) Oda vakeggas mangi, či tu phiras-l. [Romani] he told.po me that you go.IM-EV
Taj mi kaza če ti si xodel. [Bulgarian] he me told.pfd that you AUX.2SG go.PI
'He told me that you were going.' (Kostov 1963:108)

On the basis of Kaldersash Romani data, Matras (1995:100) argues that the distinction between third person singular intransitive preterites that agree in either gender or person with their subjects (e.g. masc. gelo, fem. geli personal gelas 'come') encode an evidential strategy, namely the forms agreeing with the subject in person mark personal knowledge while those agreeing in gender, which are therefore morphologically indistinguishable from past participles, mark a resultative deduction in which the addressee is invited to share in the conclusion, as in example (19):

(19) a. But zumnavenas te integrin pe, arakhenas [Romani] many tried.PI su they.integrate in found.PI buja and-e fabriki. Ale či parižilhas khanči pa jobs in-the factories but NEG changed.AO nothing on gindari sar train e Rom thoughts how.they.live the Roms
'Many tried to integrate, they found jobs in factories. But nothing changed with regard to prejudices about how the Roms live.'

b. Sa khetan i saj phenav tumenge kado. Ke sar saš [Romani] all together can I tell you.dat this that how was de katar avilam ame andre Europa či zdvies, či from whence we came.PD we in Europe till today NEG parižili pa amari situacija khanči. changed.pt on our situation nothing 'All together I can tell you this: That as it has been since we arrived in Europe and until this day, nothing has changed with regard to our situation.' (Matras 1995:100)

According to Matras, (1995:101), the person-agreeing form parižilas in (19a) expresses the speaker's personal knowledge as an expert on Romani history that nothing has changed, whereas in (19b), by using the gender-agreeing form parižili, the speaker is emphasizing the resultant state and inviting the audience to join him in reaching this conclusion on the basis of the shared evidence of the situation (in this case, the events he has related).

7. Greek

Finally, it is necessary to mention Greek, which does not have any evidential uses of paradigmatic tense forms, but does have two features that deserve notice here. First, as illustrated by example (20), the perfect is not acceptable in non-volitional expressions, which is the exact opposite of Balkan Slavic (illustrated here by Macedonian) where the PI would be the normal choice and the PD would sound strange:

(20) Κήρωσι μέ το κατά λόγο, ἐπάσα ὑπ' είκοσι τί [Greek] without su it L-understand I.broke.PD /I.have broken the bukula.
bottle
Νίτικο δα νε σεταμ, γο δρέψω / συμ γο [Macedonian] nothing su NEG I.notice it I.broke.PD AUX.1SG it skrip išle-to.
broke.PI bottle-the
'Without noticing it, I broke/have broken the bottle.' (Friedman 1977:127)

The association of nonvolitionality with nonconfirmativity occurs in other languages. In Georgian, for example, as in Balkan Slavic and Turkish, the aorist or the simple preterite is associated with confirmative meanings while the perfect is associated with nonconfirmativity.14 Example (21), however, illustrates the contrast of the aorist with volitionality and the perfect with non-volitionality in Georgian:

[Georgian text]
The typology of Balkan evidentiality and areal linguistics

8. Explanations

It was Conev (1910/1911:15–16) who first proposed the classic areal explanation for grammaticalized evidentiality in the Balkan languages when he noted the correspondence between Bulgarian and Turkish admimative expressions; however, he did not offer any semantic or terminological elucidations: "The use of the past indefinite is due to Turkish influence in instances such as: toj bil dobár čovek! [It turns out that] he is a good man!" To bil далеко! [It turns out that it] is far!" Cf. Turkish ey [modern standard iyi] adam iniz, azaq [modern standard uzak] iniz.\footnote{Colgh (1960) made a similar but more extensive set of observations for Macedonian, Weigand (1923/1924) was the first to connect Bulgarian admimative usage of the PI with the Albanian admimative paradigm, even going so far as to suggest an areal explanation, but he missed the fact that the Albanian admimative contains a true present which is only historically an inverted present, whereas the Bulgarian PI is a genuine past tense being used with only apparent present meaning (cf. Friedman 1981). Many handbooks treat these issues only in passing (Sandfeld 1930; 119–20; Schaller 1975; 79, 94; Feuillet 1986; 67; Asenova 1989: 203–204; and Demiraj 1994: 169–171). The Balkan Slavic : Albanian correspondence is treated as typological, while the Balkan Slavic : Turkish, when mentioned at all, is treated as areal. Among the similarities adduced is the parallel between Bulgarian and Turkish shown in Table 7 (using the Bulgarian l-form napisdal ‘wrote’ and the Turkish equivalent yazımız), which illustrates the assumption frequently found in Balkan linguistic literature that there are separate reported and non-reported paragraphs which are homonymous in the first two persons and marked by auxiliary omission for the reported paradigm in the third person (but see example (8)):}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Bulgarian & Reported & Turkish & Reported \\
\hline
Past Indefinite & & &  \\
\hline
1sg & napisdal sâm & yazımız &  \\
2sg & napisdal si & yazmışsan &  \\
3sg & napisdal e & napisdal & yazmış & yazımız \\
\hline
\end{tabular}
\caption{Parallel paradigms in Bulgarian and Turkish 1–3sg ‘write’}
\end{table}

\begin{thebibliography}{9}

\item Colgh (1960)
\end{thebibliography}
influence of a reinterpretation of the superficial similarity in Eastern Bulgaria (cf. the fate of Albanian -ka in the Frasheriote dialect of Gorna Belica Aromanian).

In terms of typological/universal explanations, a number of arguments can be adduced for the rise of grammaticalized evidentiality. It was Lohmann (1937) – using data from a variety of languages, including Turkish, Georgian, Armenian, Bulgarian, and Lithuanian – who first suggested the connection between the focus of the perfect on the result of a past action leading to a tendency for perfects to express the nonconfirmative evidential complex of meanings. Likewise, Whitney (1889:295–296) notes: “According to the Hindu grammarians, the Sanskrit perfect is used in the narration of facts not witnessed by the narrator; but there is no evidence of its being either exclusively or distinctively so employed at any period.”

In Takelma, a Native American language of southwestern Oregon, the inferential is not used in connected narratives despite its meaning of indirect evidence experience. In part, this is because “the constant use of relatively uncommon inferential forms would have been felt as intrusive and laborious” (Sapir 1922:158). However, the Takelma inferential is used frequently in “rhetorical questions of anger, surprise, wonder and discovery of fact after ignorance of it for some time” (Sapir 1922:200). Thus the Takelma inferential displays the typical combination of communicative and indirect meaning. 17 Worth noting here is the attractive use of the preterite in Nordic languages, e.g. Swedish Det var då en räcker, Finnish Olipas se vonkale! ‘It is [lit. was] a big one!’ (e.g. a fish).18 Educated native speakers of Nordic languages who are fluent in English calque this usage into their English. Thus, for example, upon seeing a handmade chair at a UN observation post, the chief of logistics, a Swede, commented to me in English: “That was a nice chair!” (8/1994) and a Swedish professor once remarked upon the first sip of a white wine: “That was a nice wine!” (1/2001; cf. example (11)). It is clear from the point of view of semantics, that the meanings associated with the nonconfirmative evidential complex (report, inference, sarcasm, surprise) can be expressed by the same verb forms in languages that have not been in contact with one another. This grouping by itself, however, is not an argument against an areal explanation for the development and grammaticalization of such complexes of meaning in specific cases.

Since an areal explanation is by definition diachronic, a counterargument must also make use of diachronic facts. In the case of the evidential complex, the oldest (eighth century) Turkish monuments already display an opposition between the -di and -mi2 pasts much like that of modern Turkish (Tekin 1968:192–193). At the same time, evidentiality was not encoded in Greek, Latin, or Old Church Slavonic, nor, as noted above, is there convincing evidence from Sanskrit. The one scrap of evidence concerning the syntax of the ancestor of Albanian relates to the postposed definite article (Hamp 1982) and so is not useful. Huntley (1979) has observed that the phenomenon of third person auxiliary omission is already attested in Old Church Slavonic, and van Wijk (1933) has cited evidence from a medieval

Church Slavonic paterikon that suggests that alternations between the perfect and simple preterite in Church Slavonic was being used in a proto-evidential fashion, since in the original Greek the aorist is used throughout. A relevant portion is cited here as (24):

(24) Oke navyk' velyx i navyi zavit' je' [Old Slavonic]
O.fatherILearned.prd old and new testament from ust'. i otves'tor starec' rete ume, ispl'n'll nesli
mouth and having,answered old.man said him filled:pt aux:2sg
aer' slonov' i tretti rete i mn' na ogni v'zvalo est'
word,ins and third said and me on fire grew:pt aux:3sg
byele. i otves'tor starec' rete ume. i ty
herbs and having,answered old.man said him and you
stramolubie otnga ot sebe.
hospitality you.chased.prd from self
'Father, I learned [PD] the Old and New Testaments by heart [lit. 'from mouth']. And the old man answering [lit. having answered] said: 'You have filled [PI] the air with word[s]’ […] and the third said to him: 'Grass has grown [PI] in my heart [lit. fire].' And the old man answering [lit. having answered] said: 'And you have expelled [PD] hospitality'

(van Wijk 1933:243)

According to van Wijk, the first two brothers (the second, not cited here, had a dialogue using the same tense forms as the first) use the simple preterite to describe actions that they performed themselves and the old man uses the perfect because he was not present; in the case of the third brother, the action is in the perfect because the brother himself witnessed only the result, whereas the old man answers in the simple preterite thereby demonstrating that the choice of the perfect in the first two encounters was not motivated by a desire to avoid the second person aorist, which is identical to the third person. Van Wijk suggests that the old man’s use of the simple preterite in the third encounter emphasizes the gravity of the sin, i.e. it is a kind of confirmative. For Albanian, Demiraj (1971) cites evidence of evidential usage from the earliest extensive Albanian texts, but since these were not written until the sixteenth century, their evidence is not decisive for the period prior to contact with Turkish. The situation for Romanian is similar since the earliest texts date from the sixteenth century.

The morphological historical developments in Balkan Slavic relevant to the rise of evidentiality were correlated with a series of semantic and discourse-pragmatic developments. These are summarized in Table 8 (based on Friedman 1976, 1988a):
Table 8. Stages in development of Balkan Slavic evidentiality

Stage 1 Old perfect expands to any past tense context and becomes unmarked past in all of Slavic (PI in Balkan Slavic); old perfect of `be' (masc. sg. bil) used as auxiliary in Serbia (15th century), later in Bulgaria and eastern Macedonia (17th century); simple preterite (PD) develops meaning of marked confirmativity in Balkan Slavic; PI develops the chief contextual variant meaning of nonconfirmative evidentiality.

Stage 2 Imperfect l-form develops in Western Macedonia and Eastern Macedonia and Bulgaria (16th century); rise of `have' perfect in Western Macedonia (17th century); complete loss of third person auxiliary in PI in Western Macedonia, partial loss in Eastern Macedonia and Bulgaria, sporadic loss in Serbia; PI as auxiliary used for marked nonconfirmativity.

Stage 3 Competition of PI and `have' perfect (West Central Macedonia: Prilep vs Bitola).

Stage 4 Perfect in `have' expands to auxiliaries (Ohrid-Struga, Kostur-Korča [Kastoria-Korçë]), PI limited to marked nonconfirmativity.

Stage 5 PI disappears (Kostur-Korča [Kastoria-Korçë]).

These developments result in the dialectological distribution of morphological isoglosses given in Table 9 (based on Friedman 1988a):

Table 9. Isoglosses relevant for Balkan Slavic evidentiality illustrated in Map 2

1. Complete absence of the third person auxiliary in the PI (west of Skopje-Velez-Ostrovno [Arnissa] - Kajlar [Pelemazis])
2. Presence of an imperfect l-form (west of an isogloss running southeast from Kumanovo into Aegean [Greek] Macedonia, where it meanders eastward north of Valovište [Sdihirokastron], east of Sere [Srrai], southwest of Zlatibor [Nea Záhi], north of Drama and northward through west-central Bulgaria, where the imperfect l-participle develops east of the isogloss)
3. Use of bil as an auxiliary (east of Delcevo, Kukui [Kilkis], Solun [Thessaloniki])
4. Use of perfect constructions in ima `have'
   b. Pluperfect imate dojdeno (same as above without sporadic spread)
   c. L-form ima dojdeno (Gostivar-Leria [Floria] - Vardar [Aksios]).
   d. Extension to auxiliaries `have' and `be' (south and west of Ohrid-Struga)
   e. Absence of `have except in a few expressive contexts (south and west of Kostur-Korča [Kastoria-Korçë]

The result of these developments is that the expression of evidentiality in Balkan Slavic varies significantly according to region. Western Macedonia and Eastern Bulgaria appear to be two centers of innovation. In both regions, an imperfect l-participle developed, thus making the PI exactly parallel to the PD in terms of the subordinate aspectual distinction aorist/imperfect. In Western Macedonia, however, a new perfect using the auxiliary `have' also developed, the center of innovation arguably being the Ohrid-Struga region. In this region, the PI is now limited to marked nonconfirmative contexts. To the south and west, the new perfect in `have' replaced the PI altogether, resulting in a system that recapitulates that of Common Slavic, the only difference being the survival of a few PI forms in expressive contexts. To the north and east, the PI, which had become the unmarked past in opposition to the PD, retained its neutral and resultative nuances, although it also developed the chief contextual variant meaning of nonconfirmative evidentiality in opposition to the confirmative PD. The PI is thus in competition with the `have' perfect. This is the basic situation in most of western Macedonia. In the Bitola-Resen area, however, the PD tends to replace the PI, a phenomenon which, like the rise of the `have' perfect itself, has been attributed to the influence of Armanian (Golub 1984:135; Koneski 1965:148). Thus, it can happen that a speaker from Bitola will use the PD in contexts such that a speaker from elsewhere in Macedonia will mistakenly assume that the narration is based on first-hand experience. The weakening of the confirmative/nonconfirmative sense of the opposition between the PD and the PI extends beyond Bitola to the surrounding villages and the town of Resen. Thus in Dihovo, 8 km. west of Bitola, l-forms are not consistently used in nonconfirmative contexts so much as in contexts of taxis (Friedman 1988b). Similarly, the new `have' perfect using the PI auxiliary (e.g. toj imaj doj- deno `he has/had come') does not appear to be a marked nonconfirmative. Speakers from Resen and Bitola will accept this form in contexts involving witnessed events whereas speakers from other areas will not. Thus, the Bitola-Resen system, while containing an inventory identical to that of the rest of west-central Macedonia, does not assign the same semantic or discourse-pragmatic values to the forms and seems to be moving toward a Korča-Kostur type of system in which evidentiality has been eliminated. In Eastern Bulgaria, on the other hand, the third person auxiliary is partially retained in the PI, its presence or absence being used to signal backgrounding or foregrounding. In Eastern Macedonia and Western Bulgaria, where the imperfect l-participle did not develop, there is an interplay between the presence versus absence of the auxiliary in the PI and the use of the verbal adjective such that the PI behaves mostly like an unmarked past (see Fielder 2002a).

In Albanian, too, the function of the aspirative is dialectally differentiated. The development of the inverted perfect into a classic aspirative set of paradigms is especially characteristic of central Albania (southern Geg and northern Tosk), where, as in Macedonia, urban centers such as Elbasan, which is in the southernmost Geg region, were dominated by Turkish. Here, as elsewhere in the Ottoman Balkans, Turkish functioned as a marker of urban identity (see Akan 2003). In Northern Albanian (especially rural Gag), the aspirative still retains nuances of its meaning as an inverted perfect (Gage 1979:16-18), and even shows a tendency toward being eliminated via restrictions on its occurrence. Thus, for example, in Dushmani, 30 km east of Shkoder near the Montenegro border, the aspirative only occurs in the perfect, e.g. paska pajis, Standard Albanian paska pasur (Cimo-
The typology of Balkan evidentiality and areal linguistics

identified by Matras is based entirely on native material. In this respect, Turkish is part of a larger Turkic complex in which evidential meaning is always grammatically encoded, but utilizing different material in different languages (see the articles in Johanson & Útás 2000).

Returning now to the question of areal versus typological explanation, it is clear that the meanings associated with the evidential complex, especially marked nonconformativity, are found across a broad spectrum of diverse and widely separated languages. At the same time, it is worth noting that the development of perfects into evidentials seems to take place predominantly in regions with some sort of Turkic contact, and dialects that are especially influenced by a Turkic language (e.g., Istanbul Judezmo, Inglo Georgian, Tajik, cf. Friedman 1988c) are especially likely to have a calqued development. While the development in Turkish itself does not seem to have a contact-based explanation, it is undeniable that in some cases language contact is directly responsible for the rise of the grammatical encoding of evidentiality in a given Balkan language or dialect. Frasherioti Gorna Belica Aromanian, Sliven Romani, and Istanbul Judezmo are all of this type. Moreover, it is highly likely, given the general Slavizcation of Megleno-Romanian (Capidan 1943:21), that the use of an inverted perfect as a marked non-conformative is a conceptual convergence with Slavic.20 In the case of Daco-Romanian, Manoliu-Manea (1994:270) mentions conflicting hypotheses concerning Slavic influence versus native development. The argument in favor of Slavic is the spread of ‘be’ as an auxiliary to active constructions (cf. Golgb 1984:135), which is not found in Western Romance. Similarly, the semantics of the Novo Selo probabilitive, combined with the fact that it involves a reinterpretation and subsequent extension of a native Slavic future construction – particularly salient here being the fact that the dialect is located in a region transitional between Serbian and Bulgarian with a significant Romanian-speaking population – suggest that it is a contact based phenomenon.

This leaves Albanian and Slavic. Taken as a supralinguistic set of diatopics, East South Slavic, i.e. Southern (Torlak) Serbian, Macedonian, and Bulgarian, display a series of isoglosses that are morphological, semantic, and pragmatic, in which the development of evidentiality arguably reflects the differentiation of Macedonian and Bulgarian. The medieval division between two centers of East South Slavic literacy and innovation – one in southwestern Macedonia (Ohrid), the other in northeastern Bulgaria (Preslav, later Târnovo) – continued through the Ottoman period and into the modern period. During the medieval period, these were the two areas of the most complex and intense language contact (cf. Friedman 1995). Ohrid is in the heart of the region where Albanian, Romance, Slavic, and Greek were all in contact and competition and then overlain by Turkish. The region around Preslav did not have the same complexity, but the contact with Turkish was especially strong (Golgb 1960; Conev 1910/1911:13). Moreover, in addition
to the obvious importance of Turkish as the language of administration, Turkish had a specifically urban prestige associated with Ottoman restrictions on urban residency (cf. Akan 2003). A similar situation is reflected in Albanian, where the center of administrative innovation appears to be the central, urbanized area contiguous with southwestern Macedonia. Moreover, the textual evidence for Slavic and the dialectal distribution for Albanian both suggest that the encoding of evidence, while built on native material, did not really enter the grammatical system until after contact with Turkish. Thus, for both Albanian and Balkan Slavic it can be hypothesized that the grammatical encoding of evidentiality began in urban centers where Turkish was widely spoken and had high prestige and spread from there to the countryside. Istanbul Juedezmo pluriform usage seems to reflect the beginnings of the same type of phenomenon. The absence of grammaticalized evidentiality from most Aromanian dialects, on the one hand, and from Greek, on the other, may be connected with patterns of rural transhumancy for much of Aromanian, and competing urban prestige both for Greek as a language of religion and commerce, and, in some urban centers such as Bitola, resistance of Aromanian vis-à-vis Macedonian, which during the Ottoman period was more associated with the countryside.

Balkan evidentiality, like infinitive replacement (Joseph 1983), is not a uniform phenomenon: The complex of meanings associated with marked nonconformativity clearly have a functional basis, but the grammaticalization of those meanings into verbal paradigms in the Balkans clearly requires at least a partial areal account. A mere comparison of forms, however, is inadequate. On the one hand, morphological, semantic, and discourse-pragmatic isoglosses show differential distribution, e.g., in Balkan Slavic and Albanian (cf. also Friedman 1983). Fielder (1999), in a detailed study of narrative strategy in a Bulgarian text and its Turkish translation, suggests that the evidential use of the Bulgarian PI is a conceptual convergence with Turkish (the calquing of new meaning onto existing material, after Otheguy 1991), as opposed to linguistic convergence, i.e., the borrowing of new meaning via new material. This proposal is consistent with Friedman 1994b, where I argue that discourse functions (i.e., the pragmatic component of the grammar) are not merely subject to borrowing but actually serve as entry points for the development of structural change.

In the case of Balkan evidentiality, however, there is a distinction between the Turkish/Balkan Slavic type and the Albanian/Balkan Romance type (to which can be appended Romani and Novo Selo Bulgarian). In the former, both confirmative and nonconfirmative appear as grammaticalized meanings, whereas in the latter, only nonconfirmative meanings are grammaticalized. The distinction between meanings calqued onto existing forms and the creation of new forms to convey new content does not map isomorphically onto the distribution of semantic-pragmatic isoglosses. Moreover, the geographic distribution of evidentiality in Balkan Slavic argues for a sociolinguistically based explanation for the spread of the phenomenon from Southwestern Macedonia (together with the imperfect l-form and the 'have' perfect) and Eastern Bulgaria (together with the imperfect l-participle and the reinterpretation of auxiliary omission). That is, regions with heavy concentrations of Turkish-speakers in urban centers served as sources of innovation. The spread of the Albanian admimrative northward and southward from a heavily Turkish-influenced urbanized center looks quite similar.

Lindstedt (2000), referring to Friedman's (1997) characterization of power and prestige relations among the Balkan languages suggests that it is the languages in the middle of the hierarchy, i.e., Balkan Slavic, Balkan Romance, and Albanian, that show the most mutual convergence owing to the mutual multilingualism, whereas Greek and Romani, at the relative top and bottom, respectively, show less. As Eloeva and Rusakov (1990:8) point out, there is a difference between Romani and the other languages of the Balkans. The characteristics of the Balkan linguistic league developed in environments of mutual multilingualism, i.e., speakers of a given language were influenced not only by the structures of the other languages they learned but also by hearing their languages spoken as second languages. In the case of Romani, however, (like Juedezmo), the nature of sociolinguistic relations was such that multilingualism went only in one direction. Although Romani-speakers learned other languages, their own language was not normally learned by others. Lindstedt argues that Greek is also subject to unidirectional multilingualism, but in the other direction: While speakers of other languages learned Greek, Greek-speakers were less motivated to learn other languages. Nonetheless, it is clear that many Greeks generally knew Turkish (Dizikirikis 1975), which dominated the Ottoman urban environment. At the same time, multilingualism was a characteristic of the Ottoman town. Thus, while Lindsted's hypothesis that the middle of the prestige hierarchy is the most active site of linguistic convergence has attractive features, and indeed involves precisely the languages first observed as "Balkan" by Kopitar in 1829 (see Friedman 1997), when examining evidentiality, the situation seems more complex. It would seem that urban Turkish brought about the spread of evidentials in Balkan Slavic and Albanian, but while the former is Type A like Turkish itself, the latter is Type B. At the same time, it appears that urban Aromanian blocked the development of the Balkan Slavic evidential system in Bitola, while the widespread knowledge of Turkish among Greeks was not enough to bring about the development of a Greek evidential system. On the other hand, specific dialectal phenomena such as those in Istanbul Juedezmo, Sliven Romani, Bela d Supra Frasherioti Aromanian, and Novo Selo Bulgarian show the rise of evidentials that are not attested beyond the confines of small areas. The first two dialects are urban, the second two rural. At the same time, the urban dialects are of languages at the bottom or outside the prestige hierarchy, while the rural dialects are more centrally
located in that hierarchy. Thus, intersecting factors of ethnicity and the opposition urban/rural seem to produce different effects in different environments.

Analyzing the rise and spread of grammaticalized evidentiality in the Balkans, therefore, requires an account that considers sociolinguistic factors and dialectal distribution at various levels of the grammar. The picture that emerges is a nuanced one in which form, content, narrative structure, and social factors all have roles to play, and a typologically informed areal explanation appears to be the most adequate.

Map 1. The Balkans: Places mentioned in the text

Map 2. The Republic of Macedonia and neighboring countries showing places and isoglosses mentioned in Table 9 (Names of towns and rivers are all given in Macedonian)

Notes

1. "Typology consists in the achronic equation of features or relations within grammars; such an activity is clearly a first step in the exploration of human linguistic universals. By definition, then, typology can play only a limiting, but not an initiating or central, role in the idiosyncratic specifics of language history. [...] There is then the type of study which has been called areal. Such phenomena were certainly noticed by 19th century scholars at least intuitively, but it was one of the several systematizing merits of the Pragueans to expound in an orderly and principled way the features, the affinites, characterizing a Sprachbund. [...] 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 1977:279)

2. "Given the fact that these parallels [analytic future formation, postposed definiteness marking, infinitive replacement, and resumptive pronoun object marking = VAF] are not the result of a common inheritance from Proto-Indo-European, and that they represent a divergence from earlier stages of each of these languages, the convergence they show is striking and provides an important starting point for an investigation into the language contact
situation that gave rise to them. Such, however, is not the case with parallels that can be attributed to the workings of Universal Grammar. [...] Languages in [contact] situations may show "deeper" similarities [...], but the surface forms themselves, the output of generative rules of syntax, would seem to be the critical level at which to judge similarities that would reveal the existence of a Sprachbund. [...] Such an importance of surface structures is really to be expected if the basis for the spread of such features – that is, the basis for the development of contact-induced areal convergences – is at least limited bilingualism, transfer, and reverse interference, for surface forms are the point of contact between speakers." (Joseph 2000:24–25). Cf. Croft (1990:246–259) on the debate concerning functional-typological descriptions or explanations versus formal-generative ones.

3. Hinrichs (1999) is an exception in this regard.

4. In subsequent work, Fielder has argued convincingly that auxiliary omission in the Bulgarian third person past indefinite signals the discourse phenomenon of foregrounding vs. backgrounding (see Fielder 2002b and earlier articles cited there).

5. Johanson (2000:67–71) includes the participant in the narrated event as a possible alternative to the participant in the speech event.

6. For a more nuanced account, see the articles in Johanson and Utas (2000).

7. Examples from written sources are given with bibliographic citation. Examples that occurred in ordinary conversation are given with the month and year. All other examples were checked with native speakers. Translations are mine.

8. The notion was subsequently applied to Turkish, using the term "unprepared mind" by Slabin and Akso-Koç (1982), and to Japanese with the term "new information" by Akatsuka (1985).

9. The Albanian perfect uses the present tense of 'have' (3 sg ka) plus the participle, which is usually based on the aorist stem.

10. There are two Aromanian dialects spoken in Gorna Belica, that of the Mballoits (MİbaVolts) and that of the Fashriolts (Fashrolts). The Fashriolts of Gorna Belica emigrated there from Myzeqe in central Albania before the Balkan Wars. At present, the village is uninhabited during the winter, but villagers return to their homes there for the summer.

11. I wish to thank Marjan Markovik for the opportunity to work with his uncle, Tomislav Manovski/Toma Mani, as well as Vasilie Balukoski/Silja Baliu, Andon Labrovski/Idona Jabrè, and Kosta Panovski/Koci Pani. I am grateful to all of them for sharing their knowledge of their native language.

12. In Albanian, the past admiring occurs in the imperfect but not the aorist, and this appears to have influenced the Aromanian choice of participial base. The fact that the Aromanian base is masculine plural appears to reflect a dialectal 'toki' (south Albanian) participle, which ends in shwa and thus looks like a masculine plural adjective.

13. I wish to thank Petar Atanasov and Marjan Markovik for the translation into Macedonian.

14. The prototypical category of distance was originally identified by Lunt (1952) for the Macedonian PI, subsequently added in Chvany (1986) for Bulgarian, and most recently connected with deixis in general by Fielder (2002b; cf. Also Topolińska 1995). Interestingly enough, this analysis comes full circle, in a sense, back to Conev's (1910/11) original characterization of what is general referred to today as evidentiality as part of the opposition definite/indefinite, i.e. as a form of deixis.

15. Balkan Slavic and Turkish have a subordinate aspectual distinction aorist/imperfect that can occur with both confirmatives and non-confirmatives. Georgian, too, has an aorist/imperfect subordinate aspectual distinction, but only the aorist enters into evidential oppositions (with the perfect).

16. I am indebted to César Montoliu for bringing this to my attention and supplying the examples.

17. Turkish has rules of progressive devoicing and vowel harmony that affect the surface shapes of morphemes. Thus, for example, depending on the vowel and consonant in the segment that precedes it, -dir, -dir, -dir, -dir, -dir, -dir, -dir, -dir, -dir, -dir are all possible realizations of the copulative particle for which -dir is the traditional citation form.

18. See Adelaar (1997) on the calquing of a Quechua "sudden discovery tense" as a Spanish perfect (in Ecuador) or pluperfect (in Peru and Bolivia).

19. I wish to thank Hannu Tommola for supplying the examples.

20. Fielder (2002a) has shown that in conjunction with its foregrounding and backgrounding function, absence and presence of the third person auxiliary in eastern Macedonian and in Bulgarian encodes the difference between an actional and a statal perfect, i.e. between a focus on the past action and a focus on the resultant state.

21. It is worth noting that while the form resembles Albanian, the content is Macedonian. The form is more likely the result of a Romance tendency on, perhaps, of southeast Macedonian clitic ordering rules.

22. There were also successive waves of Romani-speakers. Some of these assimilated linguistically to Albanian, Macedonian or Greek (and later, Turkish), while others retained Romani.

References


Novo Makedonija. Daily newspaper published in Skopje, Macedonia. Citation given in the order day/month/year/page.


Rilsë. Daily newspaper published in Pristina, Kosovo [at the time when these examples were published]. Citation given in the order day/month/year/page.


