Considerations upon Romanian-Albanian Linguistics Reports in Phonetics and Phonology

Cristina-Valentina Dafinoiu
"Ovidius" University Constanta

Abstract

This article focuses on some phonetic phenomena common to the Romanian and Albanian languages. First, it argues the cogency of the topic and reviews the literature in this field. Then, it identifies the phonetic phenomena (covering both the vocalic system and the consonant one) common to both above mentioned languages, by collecting, synthesizing and analyzing the information so generously offered by the literature. Next, we attempt to answer the question of whether or not these phenomena share a common origin, by highlighting the important role of the substratum often regarded as being common to both languages. Last but not least, this article studies the diachronic evolution of the phonetic phenomena shared by Romanian and Albanian. Regarding the vocalic system, we discuss issues pertaining to the "ă" vowel in Romanian and "ë" in Albanian, as well as the emergence of some other vocal phonemes in the phonological system of either language. The analysis of the consonant system of the two language has raised many issues related to the emergence and evolution of certain consonants, as well as the existence of some consonant phenomena common to Romanian and Albanian such as nasalization and rhotacism.

Keywords: *Romanian, Albanian, vocalic system, consonant system, influence, common*

The necessity of a historical study of a language results from the fact that the language is a living phenomenon, which therefore evolves in time, inevitably having to go through changes. For the Romanian language, when linguistic phenomena cannot be explained in any other way, experts often appeal to the comparison with the Albanian language, based on the fact that Romanian and Albanian descend from languages spoken by the indigenous peoples who had first inhabited the Balkan Peninsula. Thus, Romanian linguists have reached the conclusion that in the case of the two languages brought to attention, Romanian and Albanian, referring to a linguistic relationship may entail either the way one language has influenced the other or their shared Thraco-Illyrian substratum.

The earliest studies of the linguistic concordance between Romanian and Albanian are the product of foreign linguists (B. Kopitar, Fr. Miklosich, N.S. Trubezkoy, G.Mayer), who tried, sometimes even exaggerating, to clarify through appeal to Illyrian everything that was less clear in Romanian, without being able to bring enough convincing evidence to support their hypothesis; thus principles were formulated claiming that "everything that is common to Romanian and Albanian and cannot be of Latin or Slavic origin must be considered as originating from a Balkan, Illyrian or Thracian, old idiom" (Russu 168). To the hypothesis advanced by these foreign linguists have subscribed important figures in the history of the Romanian language, such as Al. Rosseti, Ov. Denusianu, B. P. Hasdeu, Al. Philippide, Th. Capidan and last but not least, Gr. Brâncuș. These researchers have all agreed, to various degrees, on the notion that Romanian has a Thraco-Dacian substratum which corresponds to, or is even identical with, the Illyrian one of Albanian (Brâncuș

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

259). In this sense, there are obvious linguistic concordances between the two languages, Romanian and Albanian, not only in phonetics but also in morphology and vocabulary.

Concerning the analysis of the phonological systems of the two languages upon which we will focus now, things have evolved a lot, by trying throughout the twentieth century to be given the most exact and complete description of the concordances, but also of the distinctions between them; amongst the common phenomena of Romanian and Albanian worth mentioning are the nasalization, rhotacism and labialization of some consonant groups.

Taking an account of the two languages' vocalic phonemes it can be observed that Albanian as well as Romanian know a number of 7 phonemes; most of them were preserved and passed on in this form from Thraco-Dacian and Illyrian; the innovations are little "typical and explicable inside the satem group of European idioms" (Russu 172).

1. Common vowels from Albanian and Romanian are *a, e, i, o, u* and respectively the central non-open vowel *ă/ë,* to which add up for Romanian *â/î* and for Albanian *y* [*ü*]; the last graphemes *î/â* and respectively *y* [*ü*] represents "different vocalic phonemes from the medial series, because the Romanian vowel is unrounded, while the Albanian one is rounded" (Russu 38).

The Romanian vowel *ă,* Albanian *ë* but also Bulgarian ъ are considered by most linguists as identical, being included in the Balkans's category. Presently, in three of the Balkan languages, this specific sound determined linguists to come up with many opinions, thus their theories regarding the history of this sound split up. On one hand are the ones that support the idea that the existence of the *ă* vowel is due to the presence of the common Romanian-Albanian substratum (Philippide 55), and on the other hand are those who consider this vowel as being a spontaneous and independent evolution (Densusianu 29), but still, "it is not due to hazard the fact that in languages as Albanian and Romanian, which are based on a common Thraco-Illyrian substratum, a vowel was developed in identical phonetic conditions, a vowel which has almost the same timbre" (Rosetti 78); the hypothesis of spontaneous evolution is less likely probable for the Bulgarian language. Some linguists also bring into discussion the possibility that the presence of the *ă* vowel in Bulgarian is due to the influence of the Romanian language. S Pușcariu brought up the idea according to which "the coincidence in closing the unstressed vowels *a* > *ă, o* > *u* and *e* > *i* which is present in Romanian, Albanian, Bulgarian and partially at neo-Greeks, is […] a phonological rule which extended father than the borders of one language without us knowing inside which language to search for the originating point is without us having to explain the genesis of this closed utterance through the influence of the substratum" (Pușcariu 267).

Analyzing the contexts of ă/ë vowel evolution in Romanian and Albanian it is noticeable that it appears only in certain conditions as part of Latin origin words as well as part of Slavic origin ones. Hence, ă/ë originates from:

- Latin words which have *a* preceded or followed by some consonants (the rule does not initially function for *a,* which remains untainted: Lat. *adancus* > Rom. *adânc*) as in the examples: Lat. *camisia* > Rom. *cămașă* and in alb. *këmishë,* Lat. *laudare* > Rom. *lăuda* and in alb. *lëvdoj*.
- For Romanian from Latin words which have *o* in stressed position, in contrast to Albanian where *o* unstressed from Latin words has become *ë*; for example: Lat. *contra* > Rom. *cătră,* Lat. *depóst* > Rom. *după*; Lat. *consillium* > alb. *këshille.*
- From Slavic origin words, which have an unstressed *o,* as in the example: Sl. *cumo* > Rom. *sită* and Alb. *sitë.*
- From Latin words who have an unstressed *e*; the articulation of the Romanian language presents analogies with the Albanian one, and the vowels' duration loss was followed by their isolation in an unstressed position; precisely this isolation tendency also reaches the *e*

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

vowel, which in an unstressed position becomes *ă*; therefore Lat. *paecatum* > Rom. *păcat* and alb. *mëkat.*

It should be mentioned that for the Romanian language, *ă* also comes from an unstressed *i* in Latin origin words as: Lat. *video* > Rom. *văd,* Lat. *capistrum* > Rom. *căpăstru.*

2. Albanian has a vocalic phoneme *y* [*ü*] which, however, does not exist in common Albanian. This sound comes from the Latin *u*, a phenomenon common to the entire Western and Central-European area of vulgar Latin. Such words as Lat. *fructus* were brought into the old Albanian *frujt*, pronounced just as the French *fruit*, while in Romanian we have *fruct*. It can be noticed that being a crossing area, Albanian takes from Latin not only particularities specific to the Western region, but also specific to the Eastern one of the Latin language; hence, a word such as the Latin *furca* became in Albanian *furkë* and in Romanian *furcă*, preserving *u*, and different from the French *forche* or the Italian *forca*, where *u* changed to *o*.

Regarding the consonant system, after a closer look from the perspective of Romanian-Albanian similarities, we will focus on some phenomena as *rhotacism*, the evolution of some consonant groups, the history of some consonants and the *nasalization*. Starting from the idea according to which "most phonetically transformations occurred in Albanian, while in Romanian the phonetic aspects of the substratum elements were better preserved" (Brancuș 217), it can be easily noticed that some terms from Albanian went under ancient phonetic transformations, which made them distinguishable from their Romanian correspondents.

a. Thus, occlusive and spirant consonants from the end of Albanian words have suffered a deafening process which represents "in all probability" (Brancuș 219) a very early dialect feature, specific to the *Tosca* Albanian dialect, and considered an innovation for this dialect. It can be noticed that this phenomenon does not occur in the *Ghega* Albanian dialect, considered more preserving than the prototype, where occlusive and final spirants were sound / voiced. Also, the final consonant preserves its sonority for forms preceded by definite article: *bredh* "fir tree" – *bredhi* "the fir tree", *bredha* "fir trees" – *bredhat* "the fir trees". This thing can be explained through the fact that the articulation phenomenon is chronologically situated before the aphonic phenomenon of the consonants. For the Romanian language, the corresponding points which support the examples from above, are usually older phonetisms and have at the end the sound consonant: *brad* (in common Romanian *bradzu*) and *mânz* (in common Romanian *mândzu*). This fact supports our statement that in Romanian, the type of words as those mentioned before must be reported to the Albanian version with sonor consonant.

b. After a long evolution, Romanian knows today two phonemes *z* and *f* originated from *dz* and *th*; these two last phonemes are known also in the contemporary Albanian which has *x* [dz] and *th*: *xixë* "spark", *xixëloj* "firefly", *thërime* "piece", *them* "to tell/to speak". In Romanian, inside words considered autochthonous, in a common phrase, *d* followed by *i* (*iot*) brought to the emergence of *dz*, consonant which, along with the separation of dialects, was kept in Armenian (*dzic, dzună, dzinere*) and became *z* in Daco-Romanian (Lat. *dico, -ere* > old Rom. *dzic* > Rom. *zic*).

The Albanian phoneme *th* knows in Romanian more graphical and sonorous transformations: *s, ț, ci/ce, f,* as it follows: Alb. *tharbët* > Rom. *sarbăd,* Alb. *thark* > Rom. *țarc,* Alb. *thump* > Rom. *ciump,* Alb. *thërrimë* > Rom. *fărâmă.*

c. Latin words, which had in their composition the consonant group *rd* (fluid followed by dental occlusive), have distinctively evolved in the two languages; while in Romanian, the *rd* group is preserved, in Albanian the *rd* group becomes *dh* or *x* [dz], as in the examples: Lat. *surdus* > Rom. *surd* and in Alb. *shurdh,* Lat. *viridis* > Rom. *verde* and in Alb. *verdh,* Lat. *meridiare* > Rom. *merizare* and in Alb. *mërxej*.

d. Concerning the Romanian-Albanian linguistic rapport in phonetics and phonology, an important problem occurs due to the sound [g'], represented in Albanian by the *gj* grapheme (as in

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

gjuhë “language”, *gjiashtë* “six”) and in Romanian through consonant groups *ghe, ghi*. In Romanian, this sound also appears in words of Latin origin (such as *gheață*), as well as in words recently emerged in the language through Turkish influence (such as *ghiaur, gherghef*), French (*ghid, ghilimele*) and German (*gherghină, gheșeft*). The fact that this sound is common in both Balkan languages and that it occurs in old autochthonous words as *ghimpe* in Romanian and *gjemp* in Albanian, again caught the attention of the specialists regarding the problem of existence and not of a common origin. The existence of a common origin of the [g’] sound in Romanian and Albanian is accepted just in so far as it admits the hypothesis that the Macedo-Romanians (who inhabited a vast area to the south of the Danube in the Balkans) might have lived near the Albanian who knew this sound and from whom they might have taken it too. Because linguists as Al. Philippide and Ov. Densusianu considered that palatalization of sounds (*p, b*) in Romanian was an early phenomenon, since the common phase of the language, brought by Romanians from the South of Danube. There are also opinions which support the idea that the phenomenon had evolved and spread independently inside the two language, and as chronology, the palatalization would be a more recent phenomenon, appeared after the break of linguistic contacts (Rosseti 263).

e. A consonant which had disappeared early in Latin, being well certified in all Romanian dialects, is *h*. The presence of this consonant in the Romanian language was long time ago based on the Slavic influence, through some Slavic words entering in the Romanian vocabulary, words such as: *har* < old Sl. *харь, hrană* < old Sl. *храна, prah* (in old Romanian) < old Sl. *прахъ* (Rosetti 302). Other linguists consider that this consonant resisted thanks to the substratum, because she occurs in some common words in Romanian and Albanian (Rom. *hameș* “lazy man” – Alb. *hamës,* Rom. *hututui* “scatterbrained / confused” – Alb. *hutoj*, Rom. *lehăi* “to speak nonsense” – Alb. *leh* “to bark”) and goes through some parallel evolutions in both languages (1. the birth of a prosthetic *h* in words that begin with a vowel, especially in Romanian language for Moldavian regionalisms: *harc, hargat, haripă, hodaie, harmăsar*, and for Albanian in both Tosca and Ghega dialects: *hark* “arch”, *hardhi* “grapevine”, *harrin* “to arrive”, *hik* “to go”*;* 2. The disappearance of *h* initially in words such as *aide, eleșteu, aiduc*, in Southern dialects of Daco – Romanian, and *eq* for *heq* “to move”, er for *herë* “date”, *a* for *ha* “to eat”); nevertheless, the Slavic influence cannot be ignored either in the birth nor in the maintenance of *h*.

However, we have to mention that a *velar* consonant *h* also exists in *ancient Greek*; the transformation of this Greek consonant *χ* [h] was made in Latin into *ch*, which brought after to consonants *c* and *k* in Romanian and Albanian: gr. *χάρτης* > Lat. *charta* > Rom. *carte,* Alb. *kartë;* old Gr. *χριστιανός* > Lat. *christianus* > Rom. *creştin,* Alb. *krishterë*; old Gr. *χορδή* > Lat. *chorda* > Alb. *kordë,* Rom. *coardă;* this transformation represents an old phonetical phenomenon often met in mediated by Latin, because the *χ* consonant is preserved untainted inside terms which occur later in languages as Romanian and Albanian (directly or through old Slavic). In this regard, the history of the Greek word *χορός* is very eloquent, a word which in an ancient evolution, brought through Latin succession *chorus* became in Romanian *cor*, and in Albanian *kor*, while, the same word old Gr. *χορός*, brought through Slavic succession, old Sl. *хорo*, became in Romanian *horă* and in Albanian *horë* (Çabej 35), so by preserving the *h* consonant untainted, mediated by Slavic. Equally important to mention is that in the kept Thracian–Dacian language material, many words which contain *h* can be found, name of plants (*hodela*) or placenames (*Histria*, *Hierasus*), names that the Latins wrote *Carsium* (*Hârșova*), *Gerasus* (*Hierasus*), which proves that, encountering the *h* sound in the language of the autochthonous, Latin wrote it using *c*, or *g*; that is why Al. Graur’s statement seems highly entitled, who said that “sounds are hard to borrow, but easily to adopt from the substratum” (Graur 401).

Regarding the origin and age of this consonant in Romanian and Albanian, specialists have divided opinions. On one side Gr. Brâncuși admitted that *h* “which disappeared in Latin since the

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

republican age, a consonant which Romanians had since Slavs' migration, comes from the language of the autochthonous. They were not taught by Slavs how to spell *h*, but borrowed from them many words containing this sound" (Brâncuși 128); on the other hand I.I. Russu does not consider pertinent the idea that *h* comes from the autochthonous substratum, replying that Brâncuși's statement of terms "as *hameș, hututui, lehăi* are entirely recent in Romanian" (Russu 429); The *History of the Romanian language*, the work of the Romanian Academy, also agrees with him. It claims that "the existence of *h* at the beginning in a few words, without usage in morphology, gave it a precarious position in common Romanian; it would have disappeared, if it had not been strengthened by many borrowings containing one *h* from Slavic or from other languages" (*Istoria limbii române* 325). However, we all accept the idea that the *h* consonant had disappeared from Latin even since the classic age. For the rest of the Romanic languages which know the *h* consonant, this is considered an innovation with various sources: it is labelled as being either an etymological script, as in French and rarely in Spanish, or a graphic sign which notes the hiatus, in French, Spanish and Italian.

f. A similar phenomenon in Romanian and Albanian which affected the words of Latin origin is *nasalization*. From a physiological point of view, it can be easily noticed that during the course of the vocal emission, the air gets out through the nasal passages and produces a special resonance, called *nasality*. The phenomenon of nasalization consists of the fact that a vowel followed by a nasal consonant is always nasalized (phenomenon mentioned in the syntactical Phonetics). In Romanian, Latin words which had *n* in their phonetic composition have evolved in such a way that *e* closed into *i*, and *o* into *u*. Hence, words such as Lat. *longus* > Rom. *lung*, Lat. *bene* > Rom. *bine*. In Albanian, the nasalization phenomenon has differently evolved in its two dialects: Ghega forms a series of nasal vowels like the ones existing in French (*â, ô, e, u*), while Tosca, on the contrary, creates an accentuated *ë* nasal vowel, like the Romanian *ă*; for example in Ghega dialect: *zâni*, *gjuni,* and in Tosca dialect: *zëni.* The *a* or *e* vowel is stressed, followed by an *n* or *m*, from words of Latin origin passed into the Tosca dialect as *ë* or *e*, and into Ghega dialect as *â* or *e* : Lat. *gamba* > Tosca dial. *këmbë,* Ghega dial. *kâmbë,* Lat. *canticum* > Tosca dial. *këngë,* and in Ghega dial. *kângë.* In Romanian, the nasalization phenomenon of the vowel placed in the same syllable with *n*, went further, so as to stressed *a* followed *n* became *ă*, then closed in further to *â*, a darker timbre, passing into the Romanian vocalic system into the central closed vowel category: Lat. *canis* > * *căne* > Rom. *câne* (dialectal), Lat. *canticum* > Rom. *cântec* > and finally *cântec*. When we take into consideration the Slavic element, it can be easily noticed that Romanian splits from Albanian, referring to the fact that inside Slavic words later introduced in Romanian, *n* does not taint the timbre of previously stressed vowel: old Sl. *рана* > Rom. *rană*; However, Albanian applies to these words a treatment similar to Latin elements: old Sl. *вранъ* > alb. *i vrërë* "black". Concerning the age of the phenomenon in two languages, Romanian specialists agree to the idea that this phenomenon belongs to the period after the fourth and fifth century, but previous to the Slavic language influence (Avram 57).

g. Another phenomenon which is similar to Romanian and Albanian is *rhotacism*, i.e. the change of intervocalic *-n-* to *-r-*. It is a phenomenon encountered in Albanian just in the Tosca dialect (in Southern Albania) and represents a differentiation criterion of the two Albanian dialects: Tosca *vera* and Ghega *vena*; this fact has led to the idea that rhotacism is a recent phenomenon in Albanian (dating since the end of the eleventh century), subsequent to the dialect splitting, which is also recent. In Romanian, rhotacism appears not just for intervocalic *-n-* but also for intervocalic *-l-* in words of Latin origin: Lat. *gelu* > Rom. *ger*, Lat. *populus* > Rom. *popor*, Lat. *mola* > Rom. *moară*; in Istro-Romanian *bire* for the Daco-Romanian *bine*, or *cire* for *cine.* It is apparent that the phenomenon does not affect Romanian words of Slavic origin (Sl. *сило* > Rom. *silă,* Sl. *мило* > Rom. *milă*), which entitles us to consider this phenomenon as already completed when the Slav

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

migration started. By trying to establish a functioning period of this phonetic law, specialists issued different opinions; some claim that rhotacism is due to the Thraco-Dacian substratum, because it occurs also in Thracian places names such as *Δώλον, Dolo, Dolus > Doris*, so it is a phenomenon belonging to the common Romanian or primitive Romanian; others claim that rhotacism is due to the impossible pronunciation of an intervocalic *-n-*, fact which led to its adaptation to joined vowels. Other researchers have stated the hypothesis that rhotacism can be considered an effect of nasalization: initially, the *n* consonant nasalized the previous vowel, *n* weakened remaining with only its dental part and then, thanks to its vibrant character, the free remaining part evolved to *-r-*; through *n*'s debilitation it has come to an intermediate form *-nr-*, encountered only in some Daco-Romanian dialects and only from this intermediate stage has the final form *-r-* emerged; therefore, the transformation of the nasal occlusive into vibrant (± nasalization) occurred only after the nasalization of the previous vowel (Avram 79).

Hence, as it was previously mentioned, the phenomenon of rhotacism is present nowadays in the Western and North-Western area (corresponding to Transylvania and "Țara Moților," viz. the Apuseni Mountains) of the Romanian language (in the Daco-Romanian dialect where nasalization seems to have first emerged), in Istro-Romanian, in Albanian (only in the Tosca dialect), but, in some way also in Greek, where, however, it does not have an unitary character. (In Greek, this phenomenon is explained as rather a dissimilation) Trying to find an origin of this phenomenon, specialists have again different opinions; some stated the opinion that Istro-Romanian must have known rhotacism before splitting from the other Romanian dialects (Daco-Romanian, Macedo-Romanian and Megleno-Romanian), because Istro-Romanians would have gone in their current homeland, Dalmatia, from North of Danube. Other linguists support the idea that the phenomenon of rhotacism emerged independently in both languages, Romanian and Albanian, considering the larger context of the rhotacism's spread in the Balkan Peninsula, in Southern Italy (this phenomenon is found also in some dialects of southern Italy) and in some Northernmost regions, in linguistic areas between which cannot be yet established any connection. To these adds up the supporters of the idea that in Albanian rhotacism is subsequent to Slavic influence, because this phenomenon also affects Albanian words of Slavic origin (Sl. **трачина* > alb. *tërsirë*). The presence in Albanian of some words such as *anë* "border", *punë* "work", *hanë* "moon" which were not affected by rhotacism because *-n-* comes from a germinated consonant *nn*, proves that in Albanian there must have existed the distinction between a hard *n* and a weak *n*.

Concerning the transformation of the intervocalic *l* in *r*, linguists (B. Kopitar, Fr. Miklosich) have tied the phenomenon since the beginning with the substratum; but its absence in Albanian and in certified Dacia words, leaves no room for the possibility to demonstrate its autochthonous Daco – Moesia origin (*Istoria limbii române* 257). This Romanian *r* resulted from the transformation of the intervocalic *l*, has in Albanian the corresponding velar *l*: alb. *mugull* the equivalent Rom. *mugure*, alb. *dhallë* for Rom. *zare*. But in Romanian *mal* is kept (term considered to be autochthonous), in Albanian *malj*; and if we also admit that the Romanian *abure* is related to the Albanian term *avull*, as their common meaning implies, then it may be claimed that Romanian has on one hand a rhotacized *l*, and on the other hand an untainted preserved *l*.

Also, it has been ascertained that, if for some phenomena a Romanian-Albanian concordance is certain, because they have a common origin, for others the Romanian-Albanian parallelism is no longer that certain, and hypotheses which were once considered relevant to proving one phenomenon or another, have been refuted by recent interpretations.

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)

References

AVRAM, Mioara. *Nazalitatea şi rotacismul în limba română.* Bucureşti, 1995
CAPIDAN, Theodor. ”*Raporturile albano-române”,* Dacoromania. II, 1921-1922.
ÇABEJ, Eqrem. *Studime etimologjike në fushë të shqipes.* Tirana, 1996
DENSUSIANU, Ovid. *Istoria limbii române.* I, Bucureşti. 1961.
Istoria limbii române, II, Bucureşti, 1969.
NICULESCU, Alexandru. *Individualitatea limbii române*. Bucureşti, 1965.
PHILIPIDE, Alexandru. *Originea românilor,* II. Iaşi, 1927.
POGHIRC, C. ”*Vocalele româneşti a, albaneze ë, bulgăreşti ъ şi oscilaţia a/e în grafia cuvintelor trace”,* SCL, XI, 1960.
ROSETTI, Alexandru. *Istoria limbii române.* Bucureşti, 1968.
RUSSU, I.I. *Etnogeneaza românilor.* Bucureşti, 1981.

BDD-A23985 © 2015 Ovidius University Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 02:59:53 UTC)