

Loanwords and Substrate in Romance Languages

DAN UNGUREANU (Prague)

Introduction

In most languages, there is a substrate, a number of words kept from the language spoken before the present one. There is a pre-Greek substrate in Greek, and a non-Indo-European substrate in Germanic. Sometimes, we are only able to identify these terms as belonging to the substrate, without any possibility of further analysis. There is an additional frustration: these words are usually few in number, and their distribution in any given language seems arbitrary.

French and Romanian were heavily influenced by Frankish and Slavic. An important Frankish minority lived among the French for centuries, until it was absorbed or got separated; an equally important Slavic minority lived among the Romanians for just as long, until it was absorbed or got separated. Frankish and Slavic provided a significant number of loanwords to French and Romanian. We will show, however, that the loanwords cluster in certain less stable areas of the vocabulary, and both French and Romanian borrow the same words – more correctly, they replace the same meanings with loanwords.

Why Romanian?

Romanian is the only Romance language developed on a non-Celtic substrate and with no Germanic loanwords. All other Romance languages have a foreseeable, and well-studied, Alpine substrate, Celtic substrate and Germanic superstrate.

1. A typology of the substrate

Some words simply do not travel. For *fir tree*, *oak*, *birch* and some other tree names, it seems as if the words have their own area and do not get replaced by invasive languages: Portuguese, Spanish, French, Romanian took the substrate words they found on the spot, as Latin did before. Tree names, rocks, dens, caves, cliffs are more attached to a place than to languages. Swamps, bogs, and marshes keep their local names.

Loanwords behave like dogs; they can be given away and will follow the new master somewhere else. Substrate words are like cats; cats prefer to stay where they live, they are more attached to their condo than to a master or another. We consider that words ‘belong’ to a language. A few of them, in fact, ‘belong’ to an area, and do not move.

1.1 General concepts

1.1.1 Abstract concepts

Lat. *autumnus*

Rom. *amurg* “evening”, Sanscr. Rig-Veda *śarvarī* “night”

Fr. *gauche* “left”, Irish *ciotóg* “left hand”, Catalan, Spanish, Portuguese *esquerra*, *izquierda* “left”

Rom. *mare* “big” (cf. Irl. *mor* “big”), Lat. *grandis* “big”, Engl. *big* may be a substrate.

Rom. *mic* “small”, Sp. *charro* “small, weak, bad”

Fr. *gaillard*, Rom. *bucuros* “joyful, merry”, Gr. *μάκαρ* “happy” belong to the substrate [and may be related as well]. Finnish *kikama* ‘revelry, merrymaking’, Orkney English *blide* “happy, pleased” (< Norn)
 Germ. Goth. *diups* “deep” etc.
 Rom. *sîmbră* “team, peasants working together”

1.1.2 Actions, verbs

Rom. *a răbda* “to endure”, Northern Saami *gierdat* “to endure” (< Germ. **hardjan*)

1.1.3 Tools and other objects

Usually, plough parts:

Fr. *soc* “plough-share”, Rom. *bîrsă* “share-beam”, Sskr. *linga* “plough”, Sanscr. *phāla-*, Pers. *supār* “plough-share”, Finnish *sabra* “wooden plough”, *saara* “fork in a plough”, Egyptian Arabic *biskha* “part of a plough” (< Coptic *περσο/πικο* “plough handle”), Egyptian Arabic *sikka* “ploughshare” (< Coptic), Egyptian Arabic *?antūt* “peg on the beam of a plough” (< Coptic), Egyptian Arabic *babūda* “plough handle” (< Coptic)

Bahraini Arabic *xašin* “digging tool”

Rom. *mătură* “broom”, Fr. *balai* “broom, Genista tinctoria, Genista sp.”, Sp. Galician *bascullo* “broom” (see also in *Plants*), Egyptian Arabic *bihnāw* “broom made of palm leaves” (< Coptic)

Fr. *braies* “trousers”, Gr. *κόμβος* “belt”

Gr. *κοσύμβη* “cloak used by herders and peasants”, Fr. *saie*, *sayon*, Rom. *țundră* “cloak worn by peasants”, Sscr. **dṛcā-* “coarse garment”

Fr. *bourre* “uncarded wool”, Rom. *bască* “freshly shorn uncarded wool”

Rom. *traistă* “sack”, Gr. *κόουκος* “leather sack”, Egyptian Arabic *shinf* “sack for straw” (< Coptic)

Gr. *πλίθος* “brick”, Old Pers. *isti* “brick”, Egyptian Arabic *īūbah* “brick” (< Coptic)

Lat. *tonsa* “oar”

1.1.4 Baskets and fish-traps

Gr. *πέριπθα* “basket for a cart”, Fr. *benne*, *banne* “basket for grapes”, Fr. *berceau* etc., Bahraini Arabic *zanbil* “basket”

Vazimba (Madagascar) *kimana* “fish-trap”, Beosi (Madagascar) *kitreko* “fish-trap”

Rus. *сырп* “fishing net” (< Mansi)

1.1.5 Foods

Rom. *brînză* “cheese”, Egyptian Arabic *halum* “cheese” (< Coptic *qalwim*), Fr. *tomme* “cheese”

French Savoyard *bléchier* “to milk”, Rom. *zer* “whey”, Rom. *urdă* “ricotta cheese”, Romansch Grisons *tšigrún* “ricotta cheese”, Swiss Glarus *ziger* id.

Indo-Iran. **piūša* “colostrum, biestings”, Sscr. Rig-Veda *kilala* id.

“colostrum” is a substrate in many languages.

1.2 Body parts

1.2.1 Face

Lat. *labium* “lip” (cognate to Germ. *lepjan), Rom. *buză* “lip”, Irula (Dravidian language) *matu* “lip”, (pre-Dravidian) Bengali *thot* “lip”, Gr. *μύσταξ* “upper lip”, *χεῖλος, χελύνη* “lip”

Lat. *bucca* “cheek”, Fr. *joue* “cheek”, proto-Germ. **kaukon* “cheek”, Sscr. *kapola* “cheek”

Sard. *bruncu* “snout”, Rom *bot* “snout”

Fr. *bec* “beak”, Rom. *cioc* “beak”, Portuguese “beak, bill”, Gr. *ράμφος* “beak”

1.2.2 Neck and throat

Fr. *gosier* “throat”, Rom. *grumaz* “neck”, Finnish *niska* “neck”, Gr. *φάρυγξ* “throat”, *λάρυγξ* “windpipe”, proto-Germ. **kbnekkān* “neck”

Sinhalese *bella* “neck” (<Vedda); Sskr. *gala* “throat”, Bengali *ghad* “neck”

Fr. *jabot* “crop”, Rom. *gușă* “crop, throat”

North Saami *suovdi* “gills”, Elmolo-Samburu *tiro* “fish liver”, *suke* “pelvic fin”

1.2.3 Other body parts

Gr. *ἰξύς* “loins”, *πυγή* “buttocks, rump”, North Russian *gurkan* “pelvis, hind part of elk back” (< Komi)

Sscr. **kućsi-* “side of the body, flank”

Finnish *täkkä* “meat on the breast of game birds”, Arm. *erbuc* “breast of animals”

Rom. *rînză* “stomach, curdle”, Occitan *bescle* “spleen”

1.2.4 Related concepts

Rom. *șută, ciută* “hornless”, Gr. *κολοβή* “with short horns”

Rom. *copil* “bastard, child” *băiat* “boy”, proto-Germ. **kiltham* “child” (Engl. *child* etc.), Sskr. Rig-Veda *kumara* “boy”

Fr. *bréhaigne* “barren (animal)”, Rom. *stearpă*, Friulan *sterpe*, Catalan *mana* “barren cow” (Portuguese etc.)

Martin Kossmann proposes to add here Baltic German *Marjelle* “girl” (from Lithuanian).

1.3 Animals

1.3.1 Diseases and endo-parasites

Fr. *douve*, Rom. *gâlbează* “liver fluke”

Portuguese *modorra* “the staggers”: intoxication of cattle with ryegrass

Gr. *ἔλμινθες* “flukes, worms”, *σκώληξ* “worm”, *ἰκτερός* “jaundice”, Sard. *tilingròne* etc. “earthworm”

1.3.2 Birds

Lat. *fulix* “heron”, Rom. *barză* “white stork, Ciconia”, Sp. *garza* “heron”, Gr. *μέροψ* “bee-eater, Merops sp.”, *ἀηδών* “nightingale”, Gr. *κόσσυφος* “blackbird, Turdus merula”. Maghreb Arabic *bellarj* “white stork, Ciconia”

Sinhalese *rera* “wild duck” (< Vedda)

Romanian (Albanian, Slovenian) *rață* “duck” is very probably a substrate.

Arm. *atawni* “dove”, Indo-Iranian **kapauta-* “pigeon” Sanskr. *kapota*

1.3.3 Fishes

In Greek, Germanic and some Saami languages, the fish names are substrates.

Germanic *roe*

1.3.4 Snails

Ital. Lucanian *caracòdio* “snail” (< Greek ?), Sp. *caracol*, Rom. *melc* “snail”

1.3.5 Horses

Celtic *marka* “horse” may be a substrate, Fr. *cheval* “horse” and Irish *pell* “horse” are substrates. So is Ugric **lox* “horse” (Hungarian, Khanti, Mansi, loanword in Russian), proto-Germ. **hursa* “horse”

Rom. *mînz* “foal, colt”, Sanskr. *kiśorā* “foal”

1.3.6 Cervids and goats

Fr. *bouc*, Engl. *buck*, proto-Germ. **lambaz* “lamb”, Rom. *țap* “billy-goat”, Sanskr. *chāga-*, Osset *saegāe* “billy-goat”, Wakhi *čəy* “kid”, and Old Irish *molt* “castrated billy-goat, wether”, Latin *caper* (*capra*) “billy-goat”, Greek *ἰξάλος* (castrated billy-goat”, Ital. Campania *zimmarru* “buck”

Fr. *chamois* “deer”, *daim* “roe deer”, Romanian *ciută* “doe”, English *doe*, Catalan *isard* “chamois”

Germ. **skap* “sheep”, Proto-Celt. **sido-* “elk, stag”, Saami Kildin *puaz* “deer”, *koan'n't* “wild deer”

Fr. Gascon *marro* “ram”, Fr. *mouton* “ram”

1.3.7 Boars

Gr. *πέλας* “boar, *Sus scrofa*”, Engl. *boar*, Indo-Iranian **uarāj^ha-* “wild boar”

1.3.8 Hounds

Engl. *dog*, Sp. *perro* “dog” (< Basque)

1.3.9 Badgers, beavers, moles

Fr. *blaireau* “badger, *Meles meles*”, Rom. *viezure* “badger”, Irish *brocc*, Welsh *broch* (pre-Celtic), North Saami *guoksi* “one year-old beaver”, Finnic **mäkrä* “badger”, Ital. Calabria *milogna* “badger”

Rom. *vătui* “newborn hare”, Irish *pattu* “hare”, *petta* “hare”, Sp. *becerro* “yearling calf” and *gazapo* “young rabbit”. Lat. *lepus* “rabbit, hare”, Ital. North Venetian dialect (Feltrinese) *solva* “talpa, mole”

1.3.10 Frogs

Romanian *broască* “frog”, proto-German **kruddon* “toad”, Portuguese *sapo* “toad”, Gr. *βάτραχος* “frog”, Finnic **konna* “toad”, Egyptian Arabic *bagrur* “frog” (< Coptic *πεκροϋρ* “toad”), Ital. dial. Calabrese *batràci* “frog” (< Greek), Ital. dial Ferrarese *galàna* “turtle, *Testudo* sp.”, French Languedoc *babi* “toad”

1.3.11 Bees and hives

Catalan *rusc* “beehive”, Fr. *ruche*, Galician *colmea*, Sp. *colmena*, Gr. *σίμβλος* “beehive” (< Semit.?)

Cat. Sp. *arna* “beehive”, Mogogodo *sakana* “beehive, log hive” (< Yaaku), *χωνόβ* “nest site in a hollow stem” (BRENZINGER 1992: 235)

Lat. *cera* “wax”

1.3.12 Parasites

Rom. *căpușă* “tick”, Catalan *paparra* “tick” [possibly Gr. *κροτόν* “tick”], Lat. *cimex* “bed-bug”

Catalan etc. *llagasta* “garrapata, tick, Ixodes sp.”, French Languedoc *lagast* “tick, Ixodes sp.”

1.4 Plants

1.4.1 Cereals

Lat. *alica* “emmer groats, Triticum diccocum” cf. Rom. *alac* “Einkorn wheat, Triticum monococcum” > Hung. *alakov* “Einkorn wheat”, id. (the Romanian word may or may not come from Lat.), Arm. *gari* “barley”, proto-Slavic **zobb* “oats, Avena sativa”, Gr. *πίτυρα* “husk, bran”

Lat. *avena* “oats, Avena sativa, straw”

1.4.2 Berries

Fr. *amélanche* “shadberry”, Fr. Gascon *avajon* “myrtille, blueberry, Vaccinium myrtillus”

Rom. *zmeură* “raspberry”, Armenian *moš* “blackberry, bramble”, Sp. *madroño* “strawberry”, Catalan *maduixa* “strawberry, Fragaria vesca”, Gr. *βάτος* “blackberry, Rubus fruticosus”, *κόμαρος* “strawberry tree, Arbutus fruticosus”, *μαντία* “raspberry”, proto-Celt. **subi* “strawberry”, Old Irish *sub*, Italian dial. Veneto *giàsena* “blueberry, Vaccinium myrtillus”, Italian Lombard *magiustra* “berry, Fragaria sp.”, Italian Lombard Ticino *dren* “raspberry” (< Celtic), Port. *arando*, Sp. *arandano*, Ital. Piemontese *arandoss* “blueberry, Vaccinium myrtillus”

Lat. *baca* “berry or other fruit”

Gr. *σῦριγξ* “reed, flute” proto-Germ. **khrudom* “reed”

1.4.3 Wine grape

Indo-Eur. **oinos*, Rom. *strugure* “grape” Lat *acinus* “grape” *racemus* id., Gr. *ἄμπελος* “grape-vine”

1.4.4 Leguminosae / Fabaceae (bean family)

Rom. *mazăre* “peas, Pisum sativum”, *păstaie* “pea pod” Latin *ervum* “vetch, Vicia sp.”, Lat. *faba* “bean”, Vedic *masūra* “lentil”, Nahali *malkā* “pea (pod)”. Gr. *κύαμος* “beans, Phaseolus sp.”, *λάθυρος* “Lathyrus sativus”, *ἔρῆβινθος* “chickpea, Cicer arietinum”, Germanic **arawit* “peas”, Germ. *Erbse*. Alb. *bathë* “broad beans, Vicia faba”, Italian Napoletano *alastra* (Acacia genus)

Catalan *arn* “gorse, Ulex europaeus”

Fr. *balai* “broom, Genista tinctoria, Genista sp.”, Galician *bascullo* “broom”

Lat *tuber* “tuber, lump”

1.4.5 Poisonous plants

Rom. *brîndușă* “Colchicum autumnale”, Gr. *ἴον* “violet”, Sp. Port. **belenium*, *belumäte* “Colchicum autumnale”, Catalan *velesa* “Colchicum autumnale”

Rom. *spînz* “henbane, Hyoscyamus niger”, Ital. Lucca *carongello* “id.” (contra, LEI) Sl. **belenъ* “henbane”

Rom. *zîrnă* “black nightshade, Solanum nigrum” Gr. *κώνειον* “hemlock, Conium maculatum” (both poisonous)

1.4.6 Plant parts

Rom. *iască* “fungus growing on trees”, cf. Gr. ἰσκαί “fungus growing on trees”

Rom. *sîmbure* “pit, kernel”, Gr. κόκκος “kernel of fruit”, Sp. *corozo* “fruit pit, kernel”

Sard *tseúrra* “bud, sprout”, Rom. *mugur* “bud, sprout”

1.4.7 Trees

1.4.7.1 Conifera

Latin *abies*, *abietis* “fir tree”, Fr. *sapin* “fir-tree”, Rom. *brad* “fir tree”, cf. Gr. βράβυ “Juniperus sp.”, Old Icelandic *bar* “fir needles”, *barr-viðr* “fir forest”, Aramaic, Hebr. *beroth* בֵּרוֹת “fir or Cyparissus sp.”¹

Lat. *larix*, Fr. *mélèze*, Rom. *zadă* “larch, Larix decidua”; Compare Fr. *mélèze* with Rom. *molid*, maybe re-made from **moliz* (Picea abies)

Ladin *dasa* “fronda di abete, fir tree branch”, Rom. *cețină* “id.”²

1.4.7.2 Birch

French *bouleau*, Spanish *abedul*, Catalan *bedoll* “birch” (could come from Celtic or from Lat. *betula*, itself of Celtic origin), Finnish *suokko* “downy birch”.

1.4.7.3 Quercus sp.

Lat. *quercus*, *aesculus* “oak”, *suber* “cork-oak”, Lat. *cerrus* “Quercus cerris”, Fr. *chêne*, Galician *caxigo* “oak”, Romanian *stejar*, Bearnais *droulh* “holm-oak, Quercus ilex”, Portuguese *carvalho*, Italian Lombard *rugul* oak, Armenian *t'etawš*, *uši/*(h)oši* ‘holm-oak’ (Hrach Martirosian), Greek *πρῖνος* “holly oak, Quercus ilex, kermes oak, Quercus coccifera” *ἰξός* “Quercus suber”, proto-Germ. **aiks* “oak”. [cf. Gr. ξύα “beech, Fagus sp.”], Egyptian Arabic *qarw* “oak” (< Coptic)

1.4.7.4 Other

Lat. *figus* “fig-tree”, Bulg. *muskura* “medlar, Mespilus germanica”

1.5 Landscape terms

According to BEEKES (2014: 46), in Greek many landscape terms belong to the substrate. The observation is valid for other languages as well.

Lat. *campus* “field”.

1.5.1 Ashes

Rom. *scrum* “anything charred” (cf. Alb. *shkrumb*). Gr. ἀλάβα “black soot for writing”, ἰχνύς “ashes”

1 Plinius *petunt igitur in Elymaeos arborem bratum, cupresso fusae similem, exalbidis ramis, iucundi odoris accensam et cum miraculo historiis Claudi Caesaris praedicatam. folia eius inspergere potionibus Parthos tradit; odorem esse proximum cedro fumumque eius contra ligna alia remedio* (Naturalis Historia, lib. XII, XXXIX) – in the Zagros mountains, in Elam.

2 Unsure. Either a Slavic loanword or a substrate.

1.5.2 *Gravel*

Lat. *sabulum*, “gravel”, Fr. *gravier* “gravel”, Gr. *ψάμμος* “sand”, Indo-Iran. **sikatā-*, Old Pers. *šikā* “sand”, Khotanese *siyatā*, Sogdian *šyktb* “sand, gravel”, Pashto *šəga*, Germanic **sand*. North Saami *čievra* “gravel”

1.5.3 *Stone*

Gr. *πέτρα* “stone”, *βάλλεκα* “*psephon*, stone”, Fr. *caillou* “stone” Catalan *pissarra* “slate”, Port. *lousa* “slate”. Egyptian Arabic *āl* “game of pebbles” (< Coptic *αλ* “pebble”), Rumantsch *crap* “stone”, Italian (Molise) *pjéskje* “stone” (< Oscan, Samnitic), Lat. *lapis* “stone”, North Saami *geadgi* “rock”
Palestinian Arabic *sarār* “pebble”

1.5.4 *Caves, natural shelters and built shelters*

Rom. *stînă* “sheep stall”, *argea* “shed, half-buried room”, Armenian *gom* “fold/stall for sheep or cattle” (Hrach Martirosian), Galician *albó*, *alboio* “shed”, Greek *μάνδρα* “sheep stall”

Sard. *ospile* “natural cover for cattle”

Lat. *favissa* “underground cellar” (if not related to *fossa*, *fodio*)

Fr. *garrenne*, *bauge* “den” (and *arma*, *tuna*, *tana*, *traucu*, *barrenc*, *aven*, *sot*, *balma*, *igue*, *garb*, *gobe*, *cros*, *muche* v. Jean-Yves Bigot, *Vocabulaire français et dialectal des cavités et phénomènes karstiques*), Port. *loira* “rabbit burrow” (Engl. *lair*?)

Ital. Liguria *armella* “cave”

Rom. *vizuină* “den”, Finnish *korju* “bear’s den”, Gr. *εἰλυμός* “den” *φωλεός* “lair, den”, North Russian *пырóm* (*pyrom*) “cover formed by fallen trees” (< Komi)

Gr. *ἄντρον* “cave”, Ital. Nort. dialect. ligur. *barma* “cave” (**balma*)

Gr. *γέφυρα* “bridge”, Arm. *kamurj* “bridge”, Alb. *urë* “bridge”

1.5.5 *Swamp*

Rom. *baltă* “swamp, pond”, Catal. *bassa* “pond”, Finnish *suo* “swamp”, Port. *Barro* “clay”, Sard. *lothiu* “muddy”, Sp. *tollo* “muddy place”, Fr. *boue* “mud”, Engl. *bog*, Sard. *Orgòs* “morass, marsh”, Fr. dial. *noue* “muddy, swampy place, brook”

1.5.6 *Water springs*

Rom. *pîrîu* “brook”, Sp., Port. *arroyo* id., North Russian *šor* “brook” (< Permian), Sard. *mintza*, *mittsa* “water spring”, Old Prov. *toron* “source, fountain” (< Basque *iturri*), Fr. Barcelonette *aven* “water spring”, North Saami *ádjág-* “water spring”, Italian Ladin *festil* “abbeveratoio, abreuvoir, water hole” (*aiva de festil*, “spring water”) may belong here.

Ital. Piemont *rugia* “water pipe”, Bahraini Arabic *skār* “irrigation channel stopper”

1.5.7 *Hills, cliffs and shores*

Lat. *caespes* “sod, turf” Fr. *motte* “clump of soil, mound”

Port. *morro* “hill” and *penedo* “cliff”, Fr. *motte*, *tolon*, and Rom. *măgură* “hill”

Fr. *berge* “embankment”

Fr. Béarnais *cruque* “hillock”, Basque *mokor* “mound, motte de terre” (does not belong here if it means only “clump of soil”), Sardinian *mògoro* “hill”, Armenian

- blur* “hill”, **doyn* “hill” (Hrach Martirosian), Greek *τύμβος* “mound”, Finnish *vaara* “hill”, North Saami *baksi* “cliff”
 Germ. *Fels* “rock”, *Klippe* “cliff, reef”, Engl. *reef*
 Rom. *mal* “shore”
 cf. Old Norse *melr* “sandbank, sand dune” (-r is a masc. sg. suffix).
 French Jërriais (Jersey island) *mielle* “sand dune”, North Saami, *miel’li* “steep, sandy bank on the shore of a river, hill”, Icelandic *melur* “gravel bank”, Welsh *moel* “bare hill”, proto-Germanic **melha* “sandbank”, Russian *мель* “shoal, sand bank”, *садиться на мель* “to run aground”
 Basque *mala* “tierra arrastrada por un torrente”, also means “hill”, Russian dial. *чурок* “coastal rock” (< Komi-Yodzyak)
 Russian dial. *вож* “low place near river” (< Komi-Permyak)
 Russian dial. *тулым* “rapids” (< Mansi)
 Finnish *neemi* “promontory”, Ital. Lucania *timpa* “promontory” (< Oscan)
 Sinhalese *kolamba* “ford, harbor” (< Vedda)
 Gr. *θάλαττα* “sea”, proto-Germ., Engl. *sea* “sea”
 Vedic *kūla* “slope, bank”, Fr. *talus* “slope”, Gr. *κρημνός* “slope, overhanging bank”, *ὄχθη* “shore, bank”, Sicilian *calanna* “slope”
 Germ. *Strand*, North Saami *uffir* “rocky slope near the seashore” [if not from Germ. *Ufer*], Finnish *niemi* “cape, peninsula”, North Russian *буздóm* “slope on river-side” (< Komi), Sp. *vega* “fertile field along a river”

1.5.8 Islands

Saami *suolu*, Finnish *saari*, Lat. *insula* “island”

1.5.9 Fallows

Fr. *jachère* “fallow”, Fr. dial. Centre *bouige* “friche, fallow”
 North Russian *šutem* “abandoned field overgrown by grass” (< Komi id. “fallow land”, Ital. Campania, Lucania *iersu* “fallow land” (< Gr. *χέρσος* “dry”), Fr. Occitan *apsar* “fallow”

1.6 Physical phenomena

Rom. *abure* “vapor”, Finnish *kaasu* “mist, haze, steam”, Gr. *κύφελλα* “clouds of mist”, Engl. *steam*, Egyptian Arabic *shabbura* “fog” (< Coptic), Rom. *negură* “mist”, North Saami *ciebka* “layer of clouds, mist or haze over the sea”, Saami Kildin *cigk* “fog”
 Fr. *galerie* “cold wind”, Germ. **brise* “breeze”, Rom. *a adia* “a breeze blows”, Orcadian English *kuil* “light breeze” (< Norn), Lat. *atabulus* (wind so called by Horatius)
 Fr. *givre* “frost”, North Saami *biegga* “rime, frost”
 Egyptian Arabic *shōb* “heat of the day” (< Coptic)

2. Substrate in French and Romanian

There is a Gaulish, Celtic substrate in French, and a Dacian substrate in Romanian.

There is an interesting parallelism between the two. Despite the relatively small number of substrate words in both languages, about 150 in each, more than a quarter of them are clustered in the same semantic areas.

For the sake of clarity and symmetry some words may be repeated. The long and close relationship of Latin with Celtic (and of Celtic with Germanic) has allowed the circulation of loanwords in all directions. Some dictionaries will list a Celtic word taken over into Latin as Latin (e.g. *caballus*). In some cases, words were shared back and forth between Celtic and Germanic (*ambactos* is just one example).

In both cases, in French as well as in Romanian, some etyma are better established than others. For French, the comparison with other Celtic languages may confirm an etymology; for Romanian, the comparison with Albanian will corroborate one. In some other cases, the Celtic origin in French and the Dacian origin in Romanian are a little more than educated guesses. A word may be assumed to be Celtic or Dacian when:

- a) it has no Latin / Slavic / Germanic etymology
- b) it is dialectal, regional, or specialized
- c) it is connected to a pre-agricultural, marginal, hunting-gathering lifestyle (forest, wild animals, fishes, hunting, wild fruits in Fr. and Rom., sheep husbandry in Romanian, some agricultural terms in French)

Loanwords are what has been *added* to an existing language. A substrate is *what is left* from a language when its speakers have adopted another.

	French	Romanian
1	alouette <i>lark</i>	ciocîrlie <i>lark</i>
	amélanche <i>shadberry</i>	zmeură <i>raspberry</i>
	bernache <i>barnacle goose</i>	barză <i>stork</i>
	balai <i>broom</i>	mătură <i>broom</i>
	bauge <i>den</i>	vizuină <i>den</i>
	blaireau <i>badger</i>	viezure <i>badger</i>
	bouc, bouquetin <i>billy-goat</i>	țap <i>buck, billy-goat</i>
	mouton <i>ram</i>	
	boue, bourbe <i>mud</i>	baltă <i>bog, swamp</i>
	bec <i>beak</i>	cioc <i>beak</i> (cf. alb.)
10	bréhaigne <i>barren (animal)</i>	stearpă <i>barren (animal)</i>
	chamois <i>deer</i>	ciută <i>doe, hind</i>
	chêne <i>oak</i>	stejar <i>oak</i>
	daim <i>roe deer, buck</i>	ciută <i>doe, hind</i>
	douve <i>liver fluke</i>	gălbează <i>liver fluke</i>
	gaillard <i>lively, merry</i>	bucuros <i>merry, joyful</i>
	galet <i>pebble</i>	stîncă <i>rock</i>
	garenne <i>rabbit warren</i>	vizuină <i>den, warren</i>
	gosier <i>throat</i>	grumaz <i>neck</i>
	graisset <i>frog, toad</i>	broască <i>frog, toad</i>

20	grève <i>sandy shore</i>	mal <i>shore</i>
	jabot <i>gizzard, crop</i>	gușă <i>gizzard, crop</i>
	joue <i>cheek</i>	buze <i>lips</i>
	mêlèze <i>larch, Larix decidua</i>	zadă, <i>larch, Larix decidua</i> , molid
	mègue (petit-lait) <i>whey</i>	zer <i>whey</i>
	motte <i>hill</i>	măgură <i>hill</i>
	petit <i>little, small, tiny</i>	mic <i>little, small</i>
	roche <i>rock</i>	stîncă <i>rock</i>
	saie, sayon <i>mantle, cloak</i>	țundră <i>mantle</i>
	sapin <i>fir, Abies sp.</i>	brad <i>fir, Abies sp.</i>
30	talus <i>slope, embankment</i>	mal <i>slope, shore</i>
	tolon <i>hill (dialectal)</i>	măgură <i>hill</i>
	cheval <i>horse</i>	mînz <i>foal, colt</i>
	soc <i>plough-share</i>	bîrsă <i>share-beam, part of the plough</i>
34	berge <i>shore</i>	mal <i>shore</i>

There are around one hundred and fifty words in each language. Around one quarter of them belong to the substrate in both Romanian and French. We suggest the correlation is strong. We will formulate a short phrase in “substrate words”:

On a hill, a small badger lives in his den, on a slope, among the fir trees, larch and oaks. Chamois, bucks and kids run on the rocks and eat berries.

3. Loanwords

Germanic loanwords in French; Slavic loanwords in Romanian

Romanian and French share not only a Romance origin, but also the choice of loanwords.

Proto-Romanian borrowed from Slavic, and Proto-French borrowed from Frankish, the same words, from the same semantic classes. We will provide below a table arranged in semantic classes and an alphabetical list.

3.1 Loanwords arranged in semantic classes

Semantic class	Nr.	Meaning	French	Romanian
body	1	ugly pockmarked cute hip-bone spur back	laid grêlé joli hanche éperon échine	slut pistruiat drăguț șold pinten cîrcă

Semantic class	Nr.	Meaning	French	Romanian
sentiments	7	horrible sad, sadness merry hate haughty, -ness brazen, shameless bold sneer smile guilt	affreux chagrin gai haine, hargne hautain, orgueil éhonté, hardi ricaner	groaznic jale vesel ciudă mîndru obraznic dîrz a zîmbi vină
actions, work	18	to leave, abandon to hang to build to roast to look at to heal to float to take care of to plait	abandonner accrocher bâtir braiser rôtir regarder guérir flotter soigner tricoter	a părăsi a atîrna a zidi a prăji a privi a lecu a pluti a îngriji a împleti
abstract ideas	27	need end joke hurry werewolf to rot, spoil to lose, destroy row	besoin bout blague hâte (loup)-garou gâter rang	nevoie sfîrșit glumă grabă vîrcolac a prăpădi rînd
colors	35	white blonde grey	blanc gris	bălai sur
weather	37	breeze cold wind snowstorm cold weather, ice frost, snow hail, black ice mud	brise bise frimas givre grésil fange	viscol, vifor chiciură zăpadă polei noroi nămol
cereals	44	cereals gruel overripe, to ripen fallow sheaf	blé, épéautre gruau blet friche gerbe	ovăz a pîrgui miriște snop

Semantic class	Nr.	Meaning	French	Romanian
plants	49	wood scrub wicker reed	bocage, bois broussaile osier roseau	crîng mărăcini răchită trestie
food	53	dregs	marc	Drojdie
animals	54	beaver crayfish sturgeon marten, Mustelidae crane duck, gander rat herd, flock	bièvre écrevisse esturgeon martre, hermine héron canard jars rat troupeau	breb rac cegă dihor, jder bîțlan, stîrc gînsac șobolan cîrd
tools	63	brush chisel hook oven stall axe basket garden sieve wooden pole pickaxe hoe	brosse burin crochet étuve étal hache hotte jardin tamis gaule pioche houe	perie daltă cîrlig sobă poliță topor coș grădină, ogradă sită prăjină tîrnăcop sapă
trade	75	costs, payment to gain to waste pledge, guarantee guarantor	des frais gagner gaspiller gage garant	plată a dobîndi <i>to gain</i> , <i>interests</i> a risipi zălog
war	79	war troops look-out to watch to wound spy coat of mail	guerre troupe aguet garder blesser navrer espion haubert	război pîlc pîndă a păzi, a străjui a răni iscoadă zale
society, institutions	86	noble market town rich	baron bourg riche	Boier tîrg bogat

Semantic class	Nr.	Meaning	French	Romanian
	94	companion mob free to hire nickname	compagnon foule affranchi engager sobriquet	tovarăş gloată slobod a năimi poreclă

French has more Frankish loanwords than there are Slavic loanwords in Romanian. French colors have a heavy Frankish influence, Romanian does not., Romanian fish names are all of Slavic origin; French fish names are inherited or of mixed origin. But there is a clear symmetry between borrowings in the two languages. There are around four hundred words of Frankish origin in modern French. About 145 of them have an exact or apt equivalent in both languages. A third of the loanwords in both languages have the same meanings. From obsolete terms (*haubert/zale*) to central words (*riche/bogat*) the correlation is well-founded and strong.

3.2 Alphabetical List

A–B–C

aguet, guet / pîndă: *look-out*
 abandonner / a părăsi: *to leave, to quit*
 accrocher / a atârna: *to hang*
 affreux / groaznic: *horrible*
 agacer / a zădări: *to irritate, to tease*
 agrafe / copcie: *staple, hook*
 aigrette / bitlan: *egret, Ardea sp.*
 baron / boier: *noble*
 bâtir / a zidi: *to build*
 baudrier / tolbă: *baldric, quiver*
 besoin / nevoie: *need*
 bièvre / breb: *beaver*
 bise, brise / viscol, vifor: *North wind, breeze, snow storm*
 blague / glumă: *joke*
 blé, épeautre, gruau / ovăz: *wheat, spelt, gruel, oats*
 blessier, navrer / a răni: *to wound*
 blet / vs. a pîrgui: *overripe, to begin to ripen*
 blond / bălai: *blonde*
 bocage / dumbravă: *wooded countryside, young forest*
 bossu / gîrbov: *hunchback*
 bouquin / vs. buche: *book, letter*
 bourg / tîrg: *town*
 bout / sfîrşit: *end*
 brandon, braise / jar: *embers*
 braiser / a prăji: *to roast*
 brimer / a obidi: *to bully, to oppress*

broussaille / măracini: *scrub*
 burin / daltă: *chisel*
 brosse / perie: *brush*
 canif / cosor: *pocket knife*
 carpe / crap: *carp*
 chagrin / jale, a se posomorî, a se mîhni: *sadness, to be sad*
 chiffon, fripe, haillon / zdrențe: *rags*
 chope, cruche / cofă: *pitcher*
 choquer / a izbi: *to hit*
 cloche / clopot: *bell*
 coq / cocoș: *rooster*
 compagnon / tovarăș: *mate*
 crochet / cîrlig: *hook*
 croupe, hanche / șold, gleznă: *hip*

D-E-F-G

dégât / pagubă: *damage*
 digue, barrage / zăgaz: *dike*
 drogue / leac: *medicine*
 duvet / puf: *down, fluff*
 écrevisse / rac: *crayfish*
 échine / cîrcă: *back*
 écurie / grajd: *stable*
 esturgeon / nisetru, cegă: *sturgeon*
 étuve / sobă: *oven*
 épéron / pinten: *spurs*
 épier, espion / a iscodi, iscoadă: *spy, to spy*
 étalage / poliță: *stall*
 faute / greșeală: *mistake*
 flotter / a pluti: *to float*
 foule, racaille / gloată: *mob*
 affranchi / slobod: *free*
 frapper, heurter / a lovi, a izbi: *to hit*
 friche / miriște: *fallow*
 frais / plată: *expenses, payment*
 froncer / a zbîrci: *to frown, to wrinkle*
 frimas, givre / chiciură, promoroacă: *frost*
 gaspiller / a risipi: *to squander*
 garnir / a împodobi: *to decorate*
 gabelle / dajdie: *taxes*
 gage, garantie / zălog: *guarantee, to warrant*
 gager, engager / a năimi: *to pawn, to hire*
 gagner / a dobîndi: *to obtain*
 gai / vesel garance, guède: *Isatis tinctoria*; drobușor: *Isatis tinctoria*
 garçon / flăcău: *lad*
 guider / a cîrmi: *to guide*

garder / a păzi, a străjui: *to guard*
 regarder / a privi: *to look at*
 garnir / a împodobi: *to decorate*
 (loup)-garou / vîrcolac: *werewolf*
 gaspiller / a risipi: *to squander*
 gaule / prăjină: *pole*
 gerbe / snop: *sheaf*
 gêter / a prăpădi: *to spoil*
 grêle, grésil/ vs. lapoviță, polei: *hail, sleet; sleet, frost*
 gruau *gruel* / ovăz: *oats*
 guérir / a lecu: *to heal, to cure*
 guerre / război: *war*
 guet / pîndă: *look-out*
 gueux / biet, sărman: *poor, beggar*

H-I-J-L

haine, hargne / ciudă: *hate, envy*
 haveneau/vîrșă: *fish net*
 hardi / dîrz: *bold*
 hache / topor: *ax*
 halte / oprire: *stop*
 hangar / grajd: *stable*
 hâte / grabă: *hurry*
 haubert / zale: *coat of mail*
 hautain / mîndru: *haughty*
 heaume / chivără: *helmet*
 héberger/ a găzdui (through Hungarian): *to lodge, accommodate*
 hermine, martre / dihor, jder: *marten*
 héron / bîtlan, stîrc: *heron*
 heurter / a izbi: *to hit*
 éhonté / obraznic: *shameless, brazen*
 hotte / coș: *basket*
 houblon / hamei: *hop*
 houe / sapă: *hoe*
 houppelande / cojoc: *fur coat*
 hoyau, pioche / tîrnăcop: *pickaxe*
 jardin / grădină, ogradă: *garden*
 jars / gînsac: *gander*
 joli / drăguț: *cute*
 marc / drojdie: *dregs*
 laid / slut: *ugly*

M-N-O-P-R

maçon / zidar: *bricklayer*
 madré, grêlé / pistruiat: *pockmarked*
 marais, mare, fange / mlaștină, smîrc: *swamp*

émousser / timpul, toci: *to blunt*
 orgueil / mîndru: *haughtiness, haughty*
 osier / răchită: *wicker*
 pot / blid: *vessel, plate*
 rabot / rindea: *a plane*
 rang, arranger / rînd, a orîndui: *line, to arrange*
 rat / șobolan: *rat*
 regarder / a privi: *to look at*
 ricaner / a zîmbi: *to smile*
 riche / bogat: *rich*
 rigole / jgheab: *channel*
 rincer / a clăti: *to rinse*
 ripaille / ospăț: *feast*
 rosse / mîrtoagă: *a nag, a bad horse*
 rôtir / a prăji: *to roast*

S-T-U-V

sabre / sabie: *sword*
 sillon / brazdă: *furrow*
 sobriquet / poreclă: *nickname*
 soigner / a îngriji: *to care, to nurse*
 tamis / sită: *sieve*
 tarir / a slei: *to dry up*
 tas / grămadă: *a heap*
 touffe, toupet: *tuft of hair / cosițe, plete: long hairs, pleat*
 tricoter / a împleti: *to knit*
 trébucher / a se poticni: *to stumble*
 trop / prea: *too much*
 troupe / pîlc: *troops*
 troupeau / cîrd, cireadă: *a herd, a flock*
 tuyau / țevă: *a pipe*
 voguer / a vîsli: *to sail, to row*

There is, in Romanian and French, the expectable fare of loanwords concerning feudal institutions, as well as feudal war and weapons. In both languages, the majority of these words became archaisms: *baron/boier* “noble”; *bourg/tîrg* “town”; *haubert/zale* “coat of mail”. However, some words connected with the social hierarchy and with the social network were abstract enough to survive: *compagnon/țovarăș* “comrade”, *prieten*, *riche/bogat* “rich”, *affranchi/slobod* “free”, *sobriquet/poreclă* “nickname”, *rang/rînd* “rank, order”, *guerre/război* “war”, *garder/a păzi* “to guard”.

Trade intensified: one had *expenses to pay*, one earned *money* or *squandered* it, and had to *guarantee* loans: *frais/a plăți* “expenses, to pay”, *gagner/a dobîndi* “to earn”, *gaspiller/a risipi* “to squander”, *gage, garant/zălog* “pledge, security”.

Both languages borrowed, [French from Germanic (Frankish) and Romanian from Slavic], words related to sentiments. *Rire/a rîde* “laugh” or *pleurer/a plînge* “to cry” are Latin, but *gai/vesel* “merry” and *chagrin/jale* “sorrow” are Germanic/ Slavic

borrowings. *hargne/ciudă* “hostility, rancor”, *hautain, orgueilleux/mîndru* “haughty”, *éhonté/obraznic* “brazen” express negative feelings or negative judgements.

Animal names

Most animal names are Latin. Less conspicuous animals have Germanic/Slavic names.

The mouse, *la souris/șoarece*, is Latin. The rat, less frequent, is Germanic and Slavic: *rat/șobolan*. The hen, *poule/găină* is frequent; geese and gander *jars/gîscă* less so.

The bear and the wolf are conspicuous, *ours/urs, loup/lup*, but the marder *martre/dibor* is not. The first are inherited, the second ones are loanwords.

Plants and agriculture

The names of two cereals, *épéautre* “spelt” and *ovăz* “oats” were borrowed from peoples living in a northern climate. They are cold-tolerant, rain-resistant cereals. *Jardin/grădină, ogradă* “garden”, *hache/topor* “ax”, *houe/sapă* (hoe) and others are related to gardening, which has a very permeable, unstable vocabulary.

Abstract ideas

Regarder/a privi “to look” belong to another category. There are old pairs of perception verbs *see/look* and *hear/listen*. The first term in these pairs (see, hear) refers to an uncontrolled, unfocused perception; the second one (look, listen) to a controlled, focused action. From a historical point of view, the first are very stable, usually of Indo-European age, the second ones are easily replaced, despite their frequency. *voir/a vedea* “to see” are Latin, and, ultimately, Indo-European. *regarder/a privi* are recent loanwords. The same occurs in Germanic (*see* and *sehen* are old, while *to look/schauen* are later and divergent. In Slavic, *vidit’* is old, *gledati/smotriti* are recent, divergent in meaning, and easily replaced. Romanian has relatively few Slavic color names; French has many.

4. Conclusions

A clear clustering appears in the two semantic fields.

Some meanings are often, in many languages, associated with the substrate. They are geographically stable. Some meanings are easily replaced, either by synonyms, in the same language, or by loanwords. We hope to have abundantly illustrated both.

Many of the substrate etymologies we use to illustrate our thesis are the result of ongoing research. Some of them are the conjectures of individual researchers, and do not yet represent the consensus of the community of linguists. Some others are widely accepted. Some others may be old loanwords, and not substrate words. However, we considered that the research in the field has reached a critical mass, and that the substrata in different languages may be presented side-by-side, and some semantic clusters may be identified.

Bibliography

- AIKIO, Ante (2004): "An essay on substrate studies and the origin of Saami". In: Irma Hyvärinen, Petri Kallio, Jarmo Korhonen (eds.): *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag* (Mémoires de la Société Néophilologique de Helsinki 63). Helsinki: Publishing House of the Société Néophilologique de Helsinki. 5–34.
- ALTHEIM, Franz (1951): *Geschichte der lateinischen Sprache von den Anfängen bis zum Beginn der Literatur*. Frankfurt am Main: Vittorio Klosterman Verlag.
- BEEKES, Robert (2009): *Etymological Dictionary of Greek*. Leiden: Brill.
- BEEKES, Robert (2014): *Pre-Greek: Phonology, Morphology, Lexicon*. Leiden: Brill.
- BREZZINGER, Matthias (1992): "Lexical Retention in Language Shift". In Matthias Brenzinger (ed.): *Language Death: Factual and Theoretical Explorations*. Berlin: de Gruyter. 213–254.
- CORRIENTE, Federico (2008): "Coptic loanwords of Egyptian Arabic in comparison with the parallel case of Romance loanwords in Andalusí Arabic". *Collectanea Christiana Orientalia* 5. 59–123.
- GRZEGA, Joachim (2001): *Romania Gallica Cisalpina*. Tübingen: Max Niemeyer.
- HOLES, Clive (2001): *Dialect, Culture, and Society in Eastern Arabia: Glossary*. Leiden: Brill.
- HOPKINS, S. (1995): "Sara 'pebbles' – A Canaanite Substrate Word in Palestinian Arabic". *Zeitschrift für Arabische Linguistik* 18. 37–49.
- HUBSCHMID, Johannes (1982): "Vorindogermanische und Indogermanische Substratwörter in den Romanischen Sprachen". In: Sture P. Ureland (ed.): *Die Leistung der Strataforschung und der Kreolistik*. Tübingen: Max Niemeyer.
- KOSSMANN, Maarten (2013): *The Arabic Influence on Northern Berber*. Leiden: Brill.
- LUBOTSKY, Aleksandr (2001): "The Indo-Iranian Substratum". In: Christian Carpelan, A. Parpola, P. Koskikallio (eds.): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*. Helsinki: Suomalais-Ugrilainen Seura. 301–317.
- MATASOVIĆ, Ranko (2012): "The substratum in Insular Celtic". *Journal of Language Relationship* 8. 153–168.
- POLOMÉ, Edgar (1992): "Germanic, North-West-Indo-European and Pre-Indo-European Substrates". In Rosina Lipi-Green (ed.): *Recent Developments in Germanic Linguistics*. Amsterdam: John Benjamins.
- SCHMID, Wolfgang (1994): "Das Lateinische und die Alteuropa Theorie". In: Wolfgang P. Schmid: *Linguisticae Scientiae Collectanea*. Berlin: De Gruyter. 316–334.
- SCHMIDT, Uwe Friedrich (2008): *Praeromanica der Italo-Romania auf der Grundlage des LEI (A und B)*. Frankfurt am Main: Peter Lang.
- SCHRIJVER, Petrus Cornelis Hendrikus (1997): "Substrate words". In: Alexander Lubotsky (ed.): *Sound Law and Analogy: Papers in Honor of Robert S.P. Beekes*. Leiden: Rodopi.
- SCHRIJVER, Petrus Cornelis Hendrikus (2000): "Non-Indo-European Surviving in Ireland in the First Millennium AD". *Ériu* 51. 195–199.
- WILLMS, Lothar (2013): *Klassische Philologie und Sprachwissenschaft*. Göttingen: Vandenhoeck and Ruprecht.
- WITZEL, Michael (1999): "Substrate Languages in Old Indo-Aryan (Rgvedic, Middle and Late Vedic)". *Electronic Journal of Vedic Studies* 5, no. 1. 1–67.
- WITZEL, Michael (2006): "Early Loan Words in Western Central Asia, Indicators of Substrate Populations" In: Victor Mair (ed.): *Contact and Exchange in the Ancient World*. Honolulu: University of Hawai'i Press. 158–190.