

**From pre-Roman to Roman Latin, through “modern” dialects:  
the origins of Lat. *lumbricus* ‘earthworm’  
from Lat. *umbilicus* ‘navel’**

*Mario Alinei*  
Emeritus University of Utrecht

**Key words:** Palaeolithic Continuity Paradigm, Lat. *umbilicus* and *lumbricus*, Italian dialects, pre-Roman Latin, Fiorano and Linear Ware cultures

**Abstract**

Within the framework of the Palaeolithic Continuity Paradigm, Lat. *lumbricus* ‘earthworm’ is seen as a semantic development of an Emilian variant of Lat. *umbilicus* ‘navel’, *\*umbricum*, which was borrowed by the Latins with the new meaning of ‘earthworm’, along with the preceding article, wrongly interpreted as part of the word. The archaeological context of this sequence of events is found in the Emilian Neolithic culture of Fiorano and in its influence on the Linear Ware culture of the contiguous area to its South.

**Premise**

Lat. *lumbricus* ‘earthworm’ is without etymology: according to Ernout and Meillet (*DELL*), the only proposed solution (Welsh *llyngyr* ‘intestinal worm’), «est sans valeur».

In my opinion we are dealing with one of those Italic, dialect words, akin to Latin and originated outside of Latium, which were introduced into pre-Roman Latin in pre- or proto-historical times, and were adopted by the written norm during the Roman age. In previous works (Alinei 2000: 958-969; 2009a, 2010) I have studied similar cases – such as *caus(s)a* from dialect variants of *\*calcea*, *pratum* from dialect variants of *pilatūm*, *belua* of *bellula*, *obturare* of *\*taurare*, *ferrum* of *fabrum* (adjective), *umbra* and *mora* of *meridies*, *caseum* of *coagulum*.

*Scritti in onore di Eric Pratt Hamp per il suo 90. compleanno*, a cura di G. Belluscio e A. Mendicino, Rende, Università della Calabria, 2010, pp. 3-13.

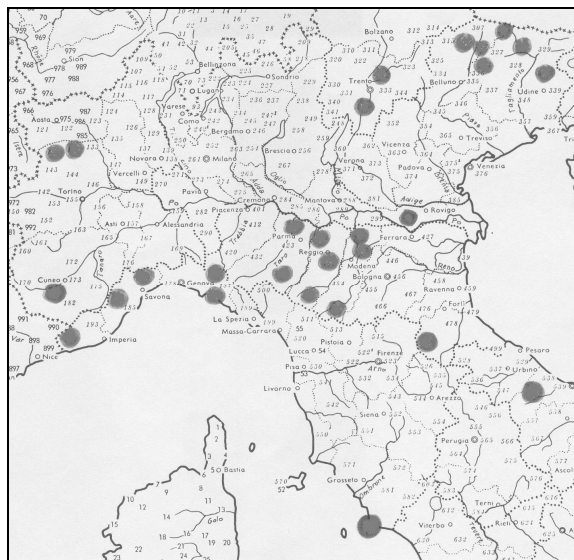
This new case seems to illustrate the development of an Emilian, Italic variant of Latin *umbilicus* ‘navel’ (of Indo-European origins: cfr. Greek ομφαλός, Sanskrit *nābhi*, Old Irish *imbliu*, Old High German *naba*, Old Prussian *nabis*: IEW 314-315), which was introduced into the spoken Latin of central Italy by a prehistoric culture which I will try to identify following the methodology of “Etymological archaeology” (illustrated in Alinei 2009b).

To begin with, I would like to recall that: (1) with regards to *umbilicus*, the great majority of its derivations in the relevant area continue a diminutive of the Latin term; (2) with regards to the forms that continue *lumbricus* the diminutive is often present, even if less frequent; (3) *lumbricus* is not the only form documented in Latin: the other forms, *lumbrix* and *lumbricis*, *lumbrica* (see DELL), prove that it is not necessary to presuppose an influence of the plural (as in REW) in order to explain the frequent variants of Italian *lombrico* with the palatalized velar (such as Marche *lumbrič*, Emilian *ombris* and *umbrizal* and Ligurian *lumbrižu*).

### From *umbilicus* to *\*umbricum*

In order to identify the Italian area where the name of the ‘navel’ apparently became that of the ‘earthworm’, we can use the phonetic change of the original group /mb<sup>vow</sup>/ of *umbilicus* into /mbr/, after the syncope of the atonic vowel, which produces the form *\*umbricum*. As is known (see also Alinei 2009a), this phenomenon is widely diffused in Italy, as well as in the whole Celto-Italic area (“Gallo-Romance” in the traditional terminology). Not by chance, it is in this area that we find French *nombril*, Old French and dialectal French *lombril*, Occitan *umbrilh* and Catalan *llombrigol*.

With regards to *umbilicus*, AIS map 130 (“l’ombelico [bottone]”) allows us to determine with sufficient precision the relevant lexical area (see figure 1).

Figure 1: Area where *umbilicus* changes to *\*umbricus* (from AIS 130)

As one can see, the type *\*umbricus* is concentrated in Southern Piedmont, Liguria, central-western Emilia and Friuli, and is also present in two isolated points (Marche and Southern Tuscany). In the North-West, apart from two points, the protonic vowel did not fall, and the group /mb<sup>vow</sup>l/ has remained unaltered, or has passed, with rotacism, to /mb<sup>vow</sup>r/: cfr. Piedmontese *ambuli/amburì*. In the central Alpine area, characterized by the syncope, the liquid consonant remained, giving way to the form /umblik/, and similarly in Eastern Emilia and Romagna, where the starting form is the diminutive *\*umbliculus*; in Lombardy and Veneto different words prevail, such as *bottone* or *bìgolo*, along with local variants of *umbilicus* (such as *bunigo*, *buligo*, *munigo*).

More important, the /mbr/-area can be subdivided in three separated areas, depending on the typology of the diminutive:

- 1) In Friuli the diminutive is of the type *\*umbilice-Ònem*: *ambritsòn*, *ambrithòn*, *umbričò* etc.;
- 2) in Southern Piedmont, Liguria, and western Emilia the diminutive is of the type *\*umbilic-Èllum*: *imbersà*, *ambarsà*, *umbritsàlu*, *umbrisàlu*, *umbarsàlu*, *obarsàl* etc.;
- 3) only in Central-Western Emilia and Romagna the diminutive is of the type *\*umbil-ìculum*: *umbrìgal*, *umbrigwal*, *umbrìgle*, *ombrìgle*, *umbrìgul*, *ombrìgol*, *umbrìgul* etc. We will concentrate on this variant, which is the only possible origin of *\*umbricum*, and thus the only one relevant for our thesis.

Figure 2 shows the relevant area more in details:

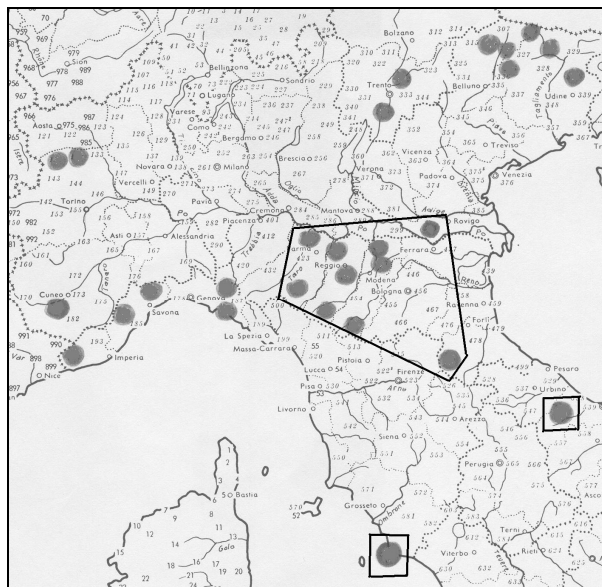


Figure 2 – Area of *\*umbricu(lu)s* ‘navel’

### *Lumbricus*

Let us now consider the name of the ‘earthworm’. First of all, the AIS map 457 (“il lombrico”) shows a development which is crucial for our thesis: in a vast northern-central area, derivatives of *lumbricus* appear with what in the traditional framework would appear as the ‘*deglutination*’ of the article – that is to say, the wrong interpretation of initial /l/ as an article, and its consequent elimination –: Italian *lombrico*, ‘earthworm’, is transformed in *l’ombrico*, ‘**the** earthworm’. Now, within the framework of our thesis (and, in general terms, of the Palaeolithic Continuity Paradigm (PCP: see <[www.continuitas.org](http://www.continuitas.org)>)), the same data can be interpreted in exactly the reverse way: it is the original Latin form *\*umbricus* ‘navel’, born as a dialectal development of *umbilicus*, which has been preserved, and has become, with the ‘*agglutination*’ of the article, the name of the earthworm *lumbricus*. Figure 3 shows the area where *lumbricus* appears without initial /l/, that is to say as *\*umbricus*:

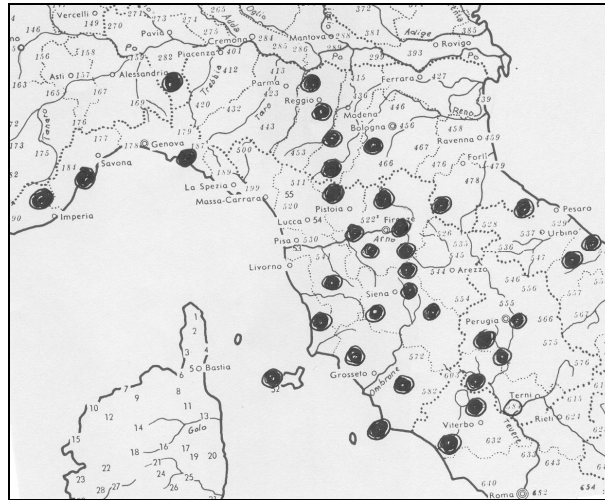


Figure 3 – Area of *\*umbricus* 'earthworm' (●)

As can be seen, the area is to the South of, and contiguous to, that of *\*umbricu(lu)s* 'navel' (see figure 2). Therefore, following our thesis, we can only conclude that *\*umbricu(lu)s* 'navel' has been introduced into Tuscany and Central Italy from Emilia. In fact, if we superimpose the two maps, this appears even clearer:

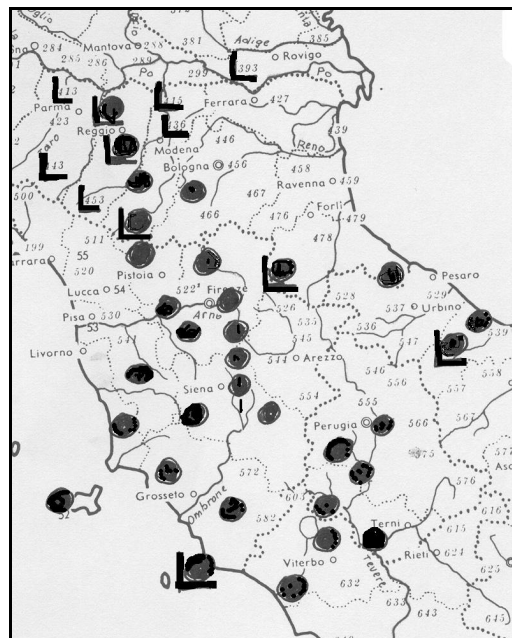


Figure 4 – Area of *\*umbricus* 'navel' (L) compared with the one of *\*umbricus* 'earthworm' (●)

As can be seen, the type *\*umbricu(lu)s* ‘navel’, concentrated in Central-Western Emilia, must have penetrated from there – as *illus umbricu(lu)s* ‘the navel’ – into Tuscany, Northern Latium, Romagna and Marche, where it became *lumbricu(lu)s* ‘earthworm’.

But the change of meaning from ‘navel’ to ‘earthworm’ must have already taken place in the focus area, that is Central-Western Emilia and Romagna, as can be shown by the absolute identity of the words in two AIS points:

AIS Point	‘navel’	‘earthworm’
464	<i>umbrìgul</i>	<i>umbrìgul</i>
490	<i>ombrìgle</i>	<i>ombrìgle</i>

It is quite likely that these correspondences would be much more numerous if the AIS points of inquiry for the Emilian area had been more numerous.

### The archaeological context

Passing now to the archaeological part of our study, I would first like to dwell on my theoretical notion of “lexical self dating” (*autodatazione lessicale*), before applying it to of our two main notions. Here is how I have defined it in my most recent work:

«Dating a word means [...] dating the moment in which the speaker has decided to lexicalize a new conceptual meaning, that is a new referent, and to this end has chosen an iconym (traditionally called *motivation*) to designate it: only this meaning is new, and as such is “born”. The date of birth of a word is the date of birth of its meaning, inextricably bound to the iconym which has been used for its designation. The moment in which an already existing word changes its meaning (that is becomes the iconym to designate a new notion), a new word is born, with its legitimate “date of birth”»<sup>1</sup>.

Now, in our case we are dealing with two “births”: that of the ‘navel’ name and that of the ‘earthworm’. The first can be easily connected to a precise prehistoric age: as it belongs, as we have seen, to the Proto-Indo-European lexicon, and this, in the PCP framework, is by definition Palaeolithic. The name of the earthworm, on the other hand,

---

<sup>1</sup> «Datare una parola significa [...] datare il momento in cui il parlante ha deciso di lessicalizzare un nuovo significato, cioè un nuovo referente, una nuova nozione, e ha quindi scelto un *iconimo* per designarlo: è *solo questo significato che è nuovo, e come tale “nasce”*. La data di nascita di una parola è la data di nascita del suo significato, indissolubilmente legato all’iconimo che è servito alla sua designazione, anche se questo può restare inaccessibile all’analisi. Il momento in cui una parola già esistente cambia di significato, cioè diventa un iconimo per designare un’altra nozione, nasce anche una nuova parola, con la sua brava “data di nascita”» (Alinei 2009b: 437-438)

can be dated with less certainty: its alternative names, in other languages and dialects, are motivated by {worm} and {earth} (cf. dialectal Italian such *verme della terra*, English *earthworm*, etc.), or by {worm} and {rain} (cf. dialectal Italian *verme della pioggia*, German *Regenwurm*), or by {intestine} and {earth} (Southern Italian *kasentaro*, from Greek γῆς ἔντερα [plural] 'worms [= bowels] of the earth'). While the generic worm – represented by Latin *vermis*, Old Icelandic *ormr*, Old English *wyrm*, Slavic *vъrmъje* 'insect', etc., and thus belonging to the PIE lexicon – is definitely a Palaeolithic notion, the name of the 'earthworm' can only be dated as a relatively more recent word.

This notwithstanding, we can arrive at a more precise dating thanks to our knowledge of the different prehistoric cultures of Italy, as well as to an anthropological consideration. To begin with, the configuration of the two areas – the primary one (Emilia) and the secondary one (Tuscany, Umbria, Northern Latium) – suggests that the innovation and its expansion belong to the Ancient Neolithic Culture of Fiorano. I have devoted an article to the study of this culture (Alinei in press), studying the etymology of the Italian verb *sdraiare* 'to lay down', whose geolinguistic area is similar to the one of *lombrico*. I only recall: that the focus of this culture is in the provinces of Modena and Reggio Emilia; that its area of expansion reaches Romagna, Northern Tuscany and Southern Veneto; and that in its most recent stages it reaches Central and Southern Tuscany, Umbria and Northern Latium, influencing the dominant culture of Southern Latium, the Linear Ware of Tuscany and Latium (5<sup>th</sup> and 4<sup>th</sup> millennia). Figure 5 shows the areas of the two adjacent cultures:

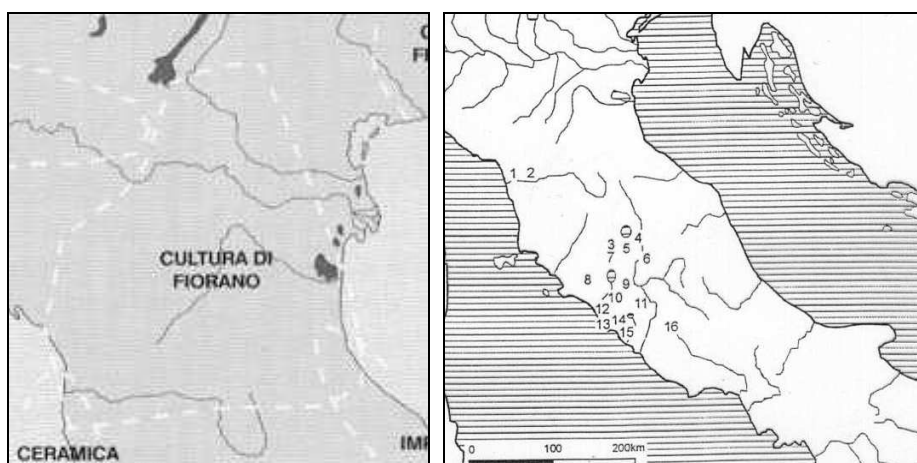


Figure 5 – Area of the Culture of Fiorano (on the left) and of the Linear Ware (on the right)

From an anthropological point of view, it cannot be surprising that the “discovery” of the earthworm took place in Neolithic. It seems in fact logical that agriculture, with the consequent attention paid to any aspect of the cultivated soil, would be the proper

occasion for discovering a particular ‘worm of the earth’ (or ‘of the rain’), which among other things resembled the umbilical cord.

### The problem of the agglutination of the article

Finally, it is important to consider another aspect of this etymology: if we assume the passage from a pre-Roman Latin *umbilicus* – through its Emilian dialectal variant *\*umbricus* – to a Classical (and more recent) Latin *lumbricus*, we have necessarily to postulate that also the agglutination of the article (in the development *\*umbricus* > *lumbricus*) took place in a proto-historic age, and not during the Middle Ages. In our research, the same assumption (but in the reverse sense: deglutination of the article) had already been postulated in order to explain the form *magnano* ‘itinerant tinker’ from *Lamagna* ‘Allemagna’, within the frame of the development of metallurgy in Southern Germany in the Bronze Age (Alinei 2000: 897-898). The date of the agglutination of the article in *lumbricus* would be the same, but in this case it would be more precise, as we could ascribe it to the period which precedes the formation of the written Roman *koiné*, that is to say in the Iron Age.

My antedating of the formation of the article from Middle Ages to the proto-historic or recent prehistoric period is usually refuted on the basis of the written record. So has done, for example, Loporcaro, in his recent (and excellent) book on Italian dialects (Loporcaro 2009: 45-46), where he quotes two texts, dated to the 9<sup>th</sup> and 3<sup>rd</sup> century AD, where the article is totally absent. I have already answered to this kind of objections, and in my last book I have devoted a whole chapter to this subject (Alinei 2009b: ch. X). My main criticism is directed at the importance usually credited to the written record as a method for dating words and things. As I have written in my above quoted book (p. 457):

«Written record never represents a *terminus a quo*, that is to say the real beginning of something. At the most, as Schuchardt had already understood (Schuchardt 1866-1868: I, 103, and see Varvaro 1968: 103), it represents a *terminus ante quem*, which is of no value whatsoever for the absolute dating of words, unless they are quite recent»<sup>2</sup>.

I have also maintained that dating words on the basis of written documents has more or less the same value as dating mountains on the basis of their first photographs (*idem*: 437)

To illustrate my point, I usually give the example of the current conjugation of the verb *avere*, ‘to have’, in standard spoken Italian, in which the particle *ci* is added to the

---

<sup>2</sup> «La documentazione scritta non rappresenta [...] mai un termine *a quo* (‘da cui’), cioè un vero e proprio “inizio” di qualcosa. Al massimo, come aveva già visto Schuchardt (Schuchardt 1866-1868: I, 103, e cf. Varvaro 1968: 103), essa rappresenta un termine *ante quem* (‘prima del quale’), che sul piano della datazione assoluta, però, ha ben poco valore, se non si tratta di parole e di datazioni recenti».

verb (*ci ho* 'I have', *ci hai* 'you have', *ci ha* 'he/she has' etc.). A construction which becomes almost obligatory when the verb is accompanied by *ne*, *lo*, *la*, *li* or *le*: *ce l'ho* 'I have it', *non ce l'ho* 'I do not have it'. Though this usage, certainly of dialectal origin, dates back (also in the traditional vision) at least to the Middle Ages, in written Italian it begins to appear only in the narrative of the Forties, after the Second World War, that is to say a thousand years after its (supposed) medieval origin. This should surprise no one. Who would *write* to a colleague, or even to a friend, *oggi non ci ho tempo* ('today I have no time')? Yet we say it often enough. It only takes a modicum of sociolinguistic sensibility to realize how huge is the distance between spoken and written language. And this distance can only be wider for an important innovation as that of the article, which did not exist in written Latin, but whose plausible existence in spoken Latin can be inferred on theoretical ground (I will come to this point), and confirmed – I would say demonstrated – by *Magnano* from *Lamagna* and *lumbricus* from *umbilicus*. The acceptance of such an important innovation in the written language may have required more time than other less significant innovations.

But there is yet another argument that seems to me decisive to confute the traditional position on the value of written record as a dating method: written records of a folk linguistic usage may indeed have a chronological value, but not so much with regards to the linguistic usage itself, as with regards to the society, within which a given measure of *distance/contrast* between written and spoken language has come to exist. For such written records of folk usages bear witness to a greater acceptance of subordinate groups and of their culture by the hegemonic elites represented by the written culture. Written documentation represents thus a crucial source for understanding the socio-political history, but not for interpreting the linguistic history.

In paradoxical terms, I would say that maintaining that written records of a folk linguistic usage can date the beginning of the folk usage itself, is more or less equivalent to maintaining that the date of the admission of women to the right of vote coincides with the date of birth of women themselves. If many features of our spoken languages appear during the Early Middle Ages, it is not because they were born in that period, but as a consequence of the disruption of the Imperial slave society and of the emergence of a new low-middle class, replacing the Roman plebs. In much the same way, pre-existing linguistic usages of the low classes started to influence the standard spoken and the written Italian language after the end of Fascism, with the introduction of democracy and of more modern social relationships. Changes in social relationships bring to light, in language use, what I usually refer to as "the hidden face of the moon".

More specifically, with regards to the determinate article, a final socio-linguistic consideration is in order: when considering its introduction in our languages, we should always remember that this is not an isolated phenomenon belonging to Romance (or, in the PCP terminology, "Neo-Italid": see Benozzo - Alinei 2011) languages, but a nearly pan-European phenomenon (a few scholars speak of a "late Indo-European" feature). Its earliest written manifestation is in classic Greek, but its geographical extension includes Germanic languages (neither Gothic nor Old English attest the article), Celtic (where it is absent in the beginning), and also Slavic (Macedonian and Bulgarian). In all these

languages, without exceptions and spontaneously, the article develops from the demonstrative adjective (in Neo-Italic dialects, in fact, from **two** demonstrative adjectives: *illum* and *ipsum*). Homeric Greek attests a transitional phase. Therefore, I would like to insist on two points: (1) the semantic difference between Latin *illum* ('that') and Italian *il* or *lo* ('the') is phonetically marked, and has thus become a lexical differentiation, but the semantic difference between Latin *illum* ('that') and Latin *\*illum* ('the'), before becoming phonetically marked, must have known a slow evolution; this means that we should assume, also for our languages, an earlier stage, equivalent to that of Homeric Greek. (2) A strong tendency of the article to expand from one language to the others is demonstrated by Macedonian and Bulgarian (the only two Slavic languages where it appears), which adopted it from the other languages belonging to the Balkan League. If then it is true, as Coseriu claimed, that the article in Neo-Italic languages is a borrowing from Greek, its introduction into the spoken language can be dated to a considerably earlier period than traditionally admitted (in addition, one should also consider that even in Greek the spoken usage of the article must have preceded the written one).

To conclude and summarize: *lumbricus* is a prehistoric Emilian word – originally without initial /l/ and with the meaning of 'navel', thus *umbricus* – which the ancient Emilians had adopted as iconym (motivation) for the name of the earthworm, and that the Latins borrowed, with the latter meaning, but hypercorrected it, assuming that the initial /l/ they heard was not the article, but part of the word.

## References

- AIS = Jaberg K., Jud J. (1928-1949). *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier
- Alinei M. (2000). *Origini delle lingue d'Europa*. Vol. 2 (*Continuità dal Mesolitico all'età del Ferro nelle principali aree etnolinguistiche*). Bologna: il Mulino.
- (2009a). Da lat. *meridies* 'meriggio delle pecore', a lat. *mora* e lat. *umbra*: origini italiche e sviluppo ligustico di un termine della pastorizia transumante. In *Quaderni di Semantica*, 30, 7-68.
- (2009b). *L'origine delle parole*. Roma: Aracne.
- (2010). Archeologia Etimologica: alle origini del formaggio. Da lat. *coagulum* 'caglio' a lat. *caseus/-m* 'formaggio'; *\*formaticum* e *\*toma*. In *Quaderni di Semantica*, 31, 73-112.
- (in press). Archeologia Etimologica: l'origine di *sdraiare* e dei suoi affini dialettali "agricoli". In *Quaderni di Semantica*, 31.
- Benozzo F., Alinei M. (2011). Dalla linguistica romanza alla linguistica neo-italide. In *Atti del VII Convegno Triennale della Società Italiana di Filologia Romanza (Bologna, 5-8 ottobre 2009)*, Bologna: Pàtron (in press).
- DELL = Ernout A., Meillet A. (1959-1960). *Dictionnaire étymologique de la langue latine. Histoire des mots*. Paris: Klincksieck (1<sup>st</sup> edition: *idem* 1932).

- IEW* = Pokorny J. (1959-1969). *Indogermanisches Etymologisches Wörterbuch*. Bern-München: Francke.
- Loporcaro M. (2009). *Profilo linguistico dei dialetti italiani*. Bari : Laterza.
- REW* = Meyer-Lübke W. (1935<sup>3</sup>). *Romanisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Schuchardt H. (1866-1868). *Das Vokalismus des Vulgärlateins*. Leipzig: Teubner.
- Vàrvaro A. (1968). *Storia, problemi e metodi della linguistica romanza*. Napoli: Li-guori.
- <[www.continuitas.com](http://www.continuitas.com)> = Palaeolithic Continuity Paradigm for the Origins of Indo-European Languages: PCP Workgroup's Official Website.