

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Evolutionary steps for linguistic signs: The place of indexicality

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1979712> since 2024-05-29T09:47:12Z

*Publisher:*

Mouton De Gruyter

*Published version:*

DOI:10.1515/9783110791433-003

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Livio Gaeta

### 3 Evolutionary steps for linguistic signs: The place of indexicality

**Abstract:** Peircean semiotics is notoriously based on triads, whose most debated triplet of index, icon and symbol has been taken to express the very essence of linguistic signs. Indexes refer to signified entities by means of physical contiguity, icons by means of structural complexity, while symbols reach the full-blown Saussurean arbitrariness interpreted in constructional terms as a conventionalized pair of form and meaning. In the paper, the place of indexicality will be discussed with reference to an emergentist perspective discussing one example in which it results from the refunctionalization of disruptive phonological change, as well as to an evolutionary perspective reconstructing the development of an indexical particle which displays peculiar “primitive” properties persisting for an impressively long timespan.

**Keywords:** index, icon, symbol, Umlaut, presentational, proto-sign, protolanguage

## 1 Introduction

As is well known, indexicality plays a central role in the semiotic model elaborated by Charles S. Peirce. While the term has been widely adopted in certain circles of linguistic pragmatics and sociolinguistics, especially with regard to the investigation of human interactions, social deixis, and the like (e.g., see Tagliamonte 2012: 30), its role as a causal factor in shaping natural languages and linguistic theorization appears to be still rather underestimated. In particular, the relevance of indexicality and indexes as specific means for organizing grammatical knowledge still needs to be couched in a systematic way within a comprehen-

---

**Note:** Parts of this paper were presented at the online workshop on “Indexicality” held during the 54<sup>th</sup> International Annual Meeting of the Societas Linguistica Europaea (30.8.-3.9.2021), and at the International Workshop on “Existential Constructions in Typological Perspective” (École Normale Supérieure, Paris 18-19.11.2016). I am very grateful to the people present on these occasions, to two anonymous reviewers, as well as to the editors of the volume and in particular to Peter Juul Nielsen for very insightful comments and remarks. The usual disclaimers apply. In this contribution, the apex <sup>+</sup> will be used for unattested reconstructed forms, while the asterisk \* marks ungrammatical expressions.

---

**Livio Gaeta**, Department of Humanities (University of Turin), e-mail: livio.gaeta@unito.it

sive semiotic theory (see Andersen 2020 and Nielsen 2016: chap. 3 for a detailed discussion). In this paper, after introducing the role of indexes for language from the perspective of Peircean theory, we will focus on their place adopting an emergentist and evolutionary perspective. We will see that indexicality emerges as a regulating factor in paradigmatic networks which have been rendered opaque by disruptive phonological change. On the other hand, we will see that indexicality is also at the heart of certain signs displaying an impressive amount of “primitive” properties and used for basic illocutionary operations like presentative utterances, which in spite of their very archaic character prove to have survived through millennia. In the paper, after discussing Peircean sign theory in details in §2, we will then focus on two specific examples of indexes illustrating respectively the emergentist development in §3 and the evolutionary perspective in §4, centering respectively on Umlaut phenomena attested in German and on the Latin particle *ecce* and on its offspring in the Romance languages. The final §5 draws a brief conclusion.

## 2 Sign triads in Peircean semiotics

Peircean semiotics is notoriously based on triads. One of those triads which attracted the scholars’ attention most is the triplet of index, icon and symbol, which has been taken to express the multi-faceted essence of linguistic signs:

One very important triad is this: it has been found that there are three kinds of signs which are all indispensable in all reasoning; the first is the diagrammatic sign or icon, which exhibits a similarity or analogy to the subject of discourse; the second is the index, which like a pronoun demonstrative or relative, forces the attention to the particular object intended without describing it; the third [or symbol] is the general name or description which signifies its object by means of an association of ideas or habitual connection between the name and the character signified (Peirce 1965: 1.370).

Especially Jakobson (1965) has drawn attention to the relevance of the three sign types – and foremost icons – for modern linguistic theory. Peirce’s categorization is based on the semiotic property of descriptiveness. Accordingly, signs are analyzed and classified on the basis of the relations they establish with the entities denoted. While symbols – i.e. symbolic coding – reach the full-blown “Saussurean” arbitrariness interpreted in constructional terms as a conventionalized pair of form and meaning, iconic coding or iconicity provides a diagrammatic description of the referent as suggested by Saussure’s idea of semi-motivation (cf. Gaeta 2002, 2022), and indexical coding or indexicality draws the speakers’ attention towards a significant semiotic event upcoming in the neighborhood. Peirce calls

symbols legisigns, while icons and indexes are labeled sinsigns because they are singular instances interpreted or semiotized as signs, displaying in other words a low degree of conventionalization (cf. Deacon 2012: 397).<sup>1</sup> In this light, conventionalized arbitrary signs are mature – i.e. fully evolved – tools for communication intended both as information exchange and mind reading (Scott-Phillips 2015). These latter two perspectives, which attempt to combine different research programs – an information-theoretic/functional and a mentalistic/cognitive one – developed through the last century to grasp the essence of language, are fully compatible with Peircean semiotics insofar as the latter crucially relies on a notion of sign involving an interpretant. The interpretant expresses at the same time the meaning – i.e. the “conventionalized” mental representation – of the sign and the concrete cognitive experience associated with it in the speaker’s mind. In this sense, the interpretant presupposes the concrete speaker intended as a full-blown semiotic animal, not merely as a computing machine for transmitting information in a classical “cybernetic” scenario. The distinction between the sign content and the interpretant’s role is important because the process of semiosis cannot simply be seen – in cybernetic terms – as the sum of different meanings contained in the speaker’s mind and arranged according to some combinatorics. Rather, semiosis – in cognitive terms – results from a continuous process of mediation between the speaker’s concrete individual experience and the social interaction in which the speaker is immersed and expresses his/her needs, feelings, intentions, etc. In other words, “reference is not intrinsic to a word, sound, gesture, or hieroglyph; it is created by the nature of some response to it” (Deacon 1997: 63), and for this reason “[n]o particular objects are intrinsically icons, indices, or symbols. They are interpreted to be so, depending on what is produced in response” (Deacon 1997: 71).

Clearly, the response and the interpretation also relate to the degree of conventionalization of the sign, i.e. to the “accumulation of instances” which “represent the relate to be a representation of the same correlate” (Peirce 1965: 1.553). To make a concrete example, we can speak of a smoke alarm’s sound as an indexical legisign because of its conventional creation and its physical linkage to smoke detection, whereas a particular smell of smoke can be considered an indexical sinsign. The interpretant’s perspective allows us to express the dynamic nature of the semiotic process, in which we constantly shift from one type of sign to the other. In this light, while we can generally consider written words as prototypical symbolic legisigns insofar as they are conventionally determined, we should also

---

<sup>1</sup> Clearly, with the increase of conventionalization also icons and indexes may reach the status of legisigns: we will come to this immediately below.

realise that a written word is first recognized as an iconic sinsign, i.e. an instance of a familiar form, then an indexical legisign because it is contiguous with other related types, and only finally as a symbolic legisign, i.e. a conventional type of sign referring to a conventional type of referent (see Deacon 2012: 398). Note that it is this potential interpretation as iconic and indexical sinsign that makes it possible to use written characters in iconic and indexical ways, as for instance the sequence :-) for the ‘smiley face’, which basically consists of three symbolic legisigns: colon, dash and bracket. This dynamic process of semiosis is essentially unlimited:

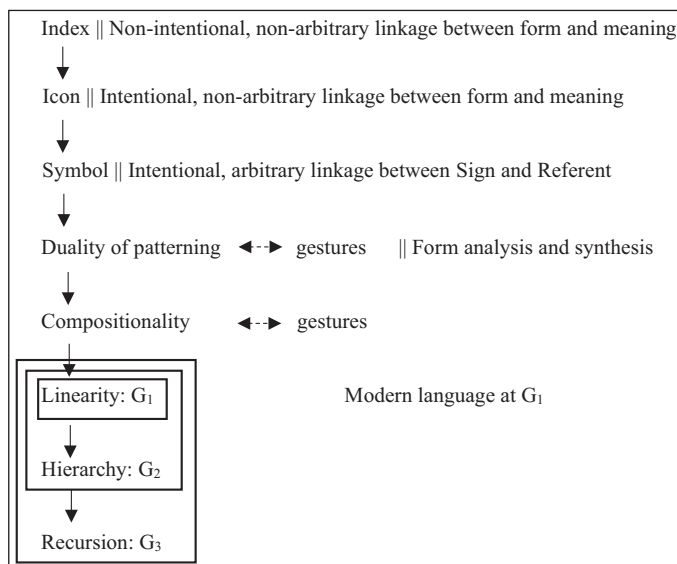
A **Sign** is anything which is related to a Second thing, its **Object**, in respect to a **Quality**, in such a way as to bring a Third thing, its **Interpretant**, into relation to the same Object, and that in such a way as to bring a Fourth into relation to that Object in the same form, **ad infinitum**. (Peirce 1965: 2.92, original emphasis)

As fully mature semiotic animals, speakers usually move back and forth along a scale of increasing conventionalization of the signs semiotizing – i.e. providing an interpretant to – objects, events and even symbols into indexes or icons. This means that indexes – and icons – are not merely natural or mechanic reflexes relating to objects or events such as for instance the knee-jerk effect: they clearly require a semiotic process which expresses “the common sense” deposited in the speakers’ community in terms of an interpretant (cf. Leone 2021).

On the other hand, this dynamic semiosis moving from descriptive sinsigns to highly conventionalized legisigns lends itself to an interpretation in evolutionary terms whereby indexes and icons can be seen as primitive signs or “proto-signs” whose meaning as a sign is not “finished” (see Sarbo 2006). Proto-signs are not characterized by the lack of a part, but rather by their being an incomplete representation of a signified entity insofar as they do not “include the meaning of all types of aspects [ . . . ] like the representation of the input qualia as independent entities, but also as a constituency, and a co-occurrence relation, etc.” (Sarbo 2006: 478). This idea of indexes and icons as primitive signs because of their “incompleteness” is an essential part of Everett’s (2017: 132) “semiotic progression”, which takes indexes and icons as true proto-signs on the evolutionary progression towards full-blown mature conventionalized Saussurean symbols, as represented in [Figure 1](#), slightly adapted from Everett (2017: 132):<sup>2</sup>

---

<sup>2</sup> For brevity we will cast aside further aspects of Everett’s (2017) model regarding the evolutionary progression of grammar like linearity, hierarchy, etc. where the step of modern language is concretely reached and further expanded.



**Figure 1:** Everett's (2017) semiotic progression.

As pointed out by Everett (2017: 132): “Indexes are ancient, far predating humans. Every animal species uses indexes, which are physical connections to what they represent such as smells, footprints, broken branches and scat. Indexes are non-arbitrary, largely non-intentional linkages between form and meaning”. Given their primitivity, indexes alone are not sufficient to give form to language in its mature shape because of the lack of two fundamental properties which are found with symbols, namely intentionality and conventionality, which are held to appear later in the evolutionary progression. Notice especially the connection of iconic sinsigns with intentionality which requires a further “semiotic” progression consisting not merely in the association of a sinsign with a referent, but in the conscious elaboration and interpretation of a sinsign “to look, sound, taste, feel, or characterise the things they represent” (Everett 2017: 134). On the other hand, “[l]anguages do have indexes where intentionality and arbitrariness have been added, going beyond the most primitive indexes shared by most species. These are words like ‘I’, ‘here’, ‘this’ and so on” (Everett 2017: 134). In Peircean terms, the latter behave like indexical legisigns, displaying a high degree of conventionalization. In other words, Saussurean legisigns are a prison from which languages cannot escape, with the marginal exception of traditional iconic signs like onomatopoeia, ideophones and the like.

Since it is a matter of fact that in languages even indexical signs appear to strongly rely on conventionalization, why are they nevertheless important and

worth researching in contrast to full-blown symbols? The answer is twofold. First, they reveal the proto-linguistic cognitive “scaffolding” on which symbols as mature signs as well as the rest of language is built, very much in the perspective of Everett’s semiotic progression. As pointed out by Jackendoff (2009: 212), “various ‘modern’ aspects of language can be added or lost piecemeal in different situations, revealing different amounts or aspects of the ‘scaffolding.’ The surprise is how robust this protolinguistic scaffolding is, emerging over and over again under different conditions”. In this sense, indexes are consubstantial with languages and we expect to observe the rise of indexes as important structural factors in historical languages. As discussed above, this “proto-linguistic” ability of elaborating indexes also underlies our capacity of creating indexical sinsigns which can subsequently be conventionalized. This can be grasped under the label of “emergent” indexicality and, as is typical of any “emergentist” approach to language (see Hopper 2015), it projects diachrony on synchrony to the extent that the synchronic grammar acquired and interactionally manipulated by concrete speakers is arranged in ordered strata of linguistic structures resulting from the broad spectrum of variation occurring within a speech community.

Second, we expect to find sparse examples of primitive indexes as remnants or witnesses of the proto-linguistic scaffolding of the human language. This idea descends directly from the assumption of such a proto-linguistic scaffolding, as suggested by Jackendoff (2009: 213):

This view of modern language as ‘laid over’ a protolinguistic substrate leads to the intriguing possibility that the coverage is incomplete: that relics of earlier stages of the language capacity remain as pockets within modern language. These relics would have only rudimentary grammatical structure, and such grammatical structure as there is would not do much to shape semantic interpretation. Rather, we would expect semantic interpretation to be highly dependent on the pragmatics of the words being combined and on the contextual specifics of use.

In spite of their actual conventionalized status, indexical legisigns are interesting per se because they are likely to display primitive properties which may tell us something about the proto-language. This can also be grasped in terms of a paleobiologist’s trip discovering fossilized relics of proto-linguistic structure which happen to occur in modern languages in spite of their archaic character, very much like the famous panda’s thumb. The latter goes notoriously back to the evolution of the sesamoid bone commonly found as a primary structure in mammals including humans, with the stabilizing function of providing the tendons a smooth surface to slide over and of maximizing in this way their ability to transmit muscular forces. In the giant panda this bone evolved in the direction of maximizing its primary stabilizing function to enhance the arboreal locomotion serving as a sort of thumb.

Thus, the panda's thumb represents a “fossilized” relic of a primary structure which has evolved getting integrated into a new functional space, i.e. arboreal locomotion.<sup>3</sup>

In the next sections, we will focus on indexicality in these two perspectives, starting with emergent indexicality as it can be observed in the course of the diachronic development of concrete linguistic systems, and then broadening our perspective to the evolutionary dimension, with the discussion of one example of an indexical legisign qualifying as a relic of protolanguage.

### 3 Emergent indexicality

Adopting an emergentist perspective, indexical legisigns can be seen to arise as the by-product of other factors affecting linguistic coding, as for instance disruptive phonological change. One paramount example of emergent indexicality is given by the Umlaut alternations consisting in a root-vowel fronting found in Modern Standard German (= MSG). As is well known, Umlaut alternations go back to vowel harmony triggered by a subsequent front high vowel in Old High German (= OHG), largely blurred in MSG:

(1)	OHG	<i>rōt</i> ‘red’	MSG	<i>rot</i> ‘red’
		<i>r[ø:]tir</i> ‘redder’		<i>röter</i> ‘redder’
		<i>r[ø:]tī</i> ‘redness’		<i>Röte</i> ‘redness’
		<i>r[ø:]tjan</i> ‘to redden’		<i>röten</i> ‘to redden’
		<i>r[ø:]tlīh</i> ‘reddish’		<i>rötlich</i> ‘reddish’
		<i>r[ø:]til</i> ‘red chalk’		<i>Rötel</i> ‘red chalk’

In MSG, the Umlaut alternation normally signals the morphologically complex, i.e. “derived”, status of the involved signs independently of their concrete meanings and accordingly invites the reconstruction of complex paradigmatic relations within a given lexical family (Wurzel 1984, Gaeta 1998, 2004, Andersen 2010). Notice that in MSG Umlaut largely works on a lexical family basis (Wurzel 1981: 936, Eisenberg 2020: 231). Thus, we will find the same Umlaut alternation in any member of the word family of the adjective *rot* ‘red’, far beyond the original OHG alternations as shown by *Röt* and *Rötling*:

---

<sup>3</sup> In biology – as well as in theoretical linguistics – the panda's thumb has been mentioned as a classical example of exaptation (cf. Norde & Van de Velde 2016). In this contribution, I will not adopt this point of view, and the reader is referred to Gaeta (2016) for a discussion of the issue.



- (2) a. *röter* ‘redder’      *rötest* ‘reddest’      b. *rote / rotes / roter / roten*  
*Röte* ‘redness’      *Röt* ‘redness’  
*röten* ‘to redden’      *Rötling* ‘orange agaric’  
*rötlich* ‘reddish’      *erröten* ‘to blush’  
*Rötel* ‘red chalk’      *Röteln* ‘measles’

Only the four inflectional forms of the adjective (2b) are excluded from the Umlaut alternation because their inflectional endings did not display any trigger for the rise of the alternation, while the comparative and the elative forms *röter* and *rötest* – where a palatal vowel used to occur, see (1) above – clearly side with the rest of the derivatives (2a). In neat contrast to *rot*, the adjective *roh* ‘raw’ never displays Umlaut alternations (3a), while in the lexical family of *zart* ‘tender’ Umlaut shows up only in some of the derivatives:

- (3) a. *roh* ‘raw’      b. *zart* ‘tender’  
*roher* ‘rawer’      *zarter* ‘tenderer’  
*rohest* ‘rawest’      *zartest* ‘tenderest’  
*Rohling* ‘brute’      *Zartheit* ‘tenderness’  
*Roheit* ‘brutality’      *Zärte* ‘tenderness’  
*verrohen* ‘to brutalize’      *zärtlich* ‘tender’  
      *Zärtling* ‘mollycoddle’  
      *zärteln* ‘to caress’  
      *Zärtelei* ‘caressing’  
      *verzärteln* ‘to mollycoddle’

Thus, we can contrast the family of *rot* which consistently shows Umlaut in any derivative (see Figure 2) with the lexical family of *roh* where no Umlaut is found (see Figure 3) and with the partial picture given by the lexical family of *zart* (see Figure 4):

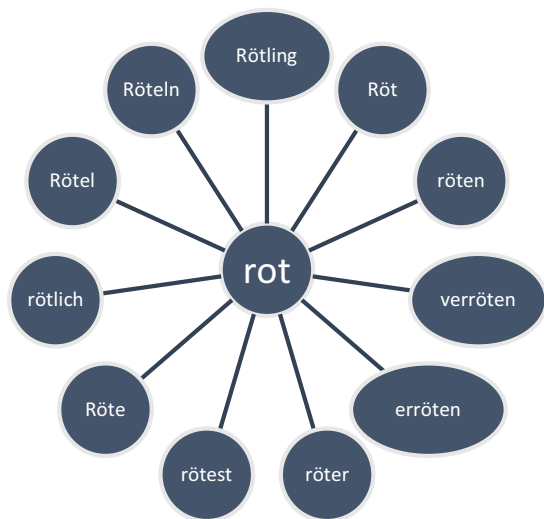


Figure 2: The family of *rot*.

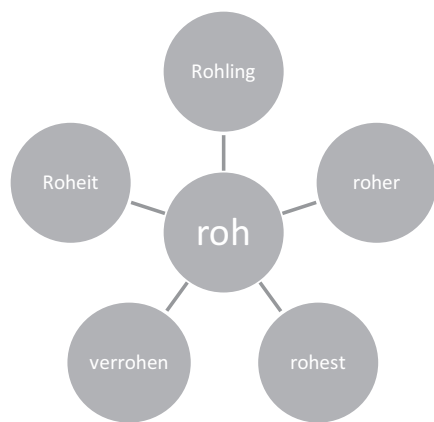


Figure 3: The family of *roh*.



**Figure 4:** The family of *zart*.

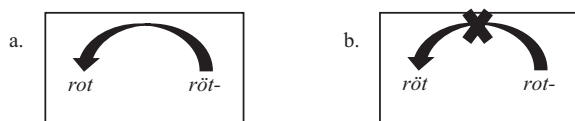
It has to be added that inflectional forms are not excluded a priori by the Umlaut alternation but their involvement depends on the word-class. For instance, in certain nominal lexical families the indexical legisign also extends to plural formation:

- |     |   |   |
|-----|---|---|
| (4) | a. <i>Gast</i> ‘guest’<br><i>Gastes</i> ‘guest.GEN’ <sup>4</sup><br><i>gastlich</i> ‘convivial’<br><i>gastieren</i> ‘to guest’            | b. <i>Gäste</i> ‘guests’<br><i>Gästen</i> ‘guests.DAT’  |
| (5) | a. <i>Haus</i> ‘house’<br><i>Hauses</i> ‘house.GEN’<br><i>Hause</i> ‘house.DAT’<br><i>hausen</i> ‘to dwell’<br><i>hausieren</i> ‘to hawk’ | b. <i>Häuser</i> ‘houses’<br><i>Häusern</i> ‘houses.DAT’<br><i>Häuschen</i> ‘house.DIM’<br><i>Gehäuse</i> ‘housing’<br><i>Häusler</i> ‘cottager’<br><i>häuslich</i> ‘homely’<br><i>außerhäusig</i> ‘off-site’ |

<sup>4</sup> In addition to the standard abbreviations of the Leipzig Glossing Rules, the following glossing abbreviations are used: CONJ = conjunction, DIM = diminutive, INTER = interrogative.

Note that while in the lexical family of *Gast* the Umlaut alternation is only limited to the inflectional plural forms (4), for *Haus* both inflectional and derivational forms are involved (5). Finally, the Umlaut alternation also concerns lexical families of verbs belonging to the so-called strong or irregular class. For instance, the verb *fahren* ‘to go, drive’ displays Umlaut alternations in the second and third person singular of the present indicative: *fahre* / *fährst* / *fährt* ‘go.PRS.IND.1/2/3SG’ and in the past subjunctive, which is based on the preterit: *fuhr* / *führe* ‘go.IND.PST.1SG/SBJV.PST.1SG’. Thus, it is not the case that the umlauted form pertains to word-formation while inflection remains unumlauted. Moreover, the mere occurrence of a rounded front vowel does not imply that such an indexical relation occurs. For instance, the word family of *schön* ‘beautiful’ regularly displays a rounded front root-vowel (see *schöner* ‘more beautiful’, *Schönheit* ‘beauty’, etc.) going etymologically back to Umlaut: see OHG *sc[ø:]ni* ‘handsome’, while a back vowel is found only in the word *schon* ‘already’, etymologically connected (see OHG *scōno* ‘already’) but nowadays completely unrelated with *schön*.

The occurrence of Umlaut across the MSG lexical families depicts a complex and partially idiosyncratic picture. However, one common property can be identified. While Umlaut in MSG cannot univocally be related to any particular morphological meaning, its indexical function enhances the paradigmatic tightness of the lexical family to the extent that the occurrence of a front rounded vowel invites the inference of a morphological relation with a basis containing a corresponding back vowel (see Figure 5a):



**Figure 5:** Indexicality of MSG Umlaut with adjectival bases.

Crucially, the opposite in which a back vowel invites the inference of a basis containing a front rounded vowel does not normally occur (see Figure 5b).<sup>5</sup> It has to be emphasized that this latter case is not excluded a priori, and in fact it is attested

<sup>5</sup> One reviewer asks whether the stem with the non-fronted vowel would also index the variant with the fronted vowel, given that one would know about the paradigmatic link between *rot-* in *rot/rote/roten* etc. and *röt-* in *röter; Rötel, erröten* etc. Clearly, the answer is yes because adjective inflection generally does not display Umlaut alternations, as observed above. But this does not affect the general point made here that the morphologically more complex (‘derived’) form displays a fronted vowel while the opposite does not normally hold.

through the history of German and marginally survives in MSG. It is known under the label of *Rückumlaut*, i.e. ‘backwards Umlaut’, in which a back vowel is found in derived contexts in correspondence of a front vowel in the base form as in OHG *h[ø:]ren* ‘to hear’ / *hörta* ‘hear.PST.1/3SG’, *stellen* ‘to put’ / *stalta* ‘put.PST.1/3SG’, *w[æ:]nen* ‘to mean’ / *wānta* ‘mean.PST.1/3SG’, *sc[y]tten* ‘to shake’ / *scutta* ‘shake.PST.1/3SG’ etc. (cf. Vennemann 1986, Fertig 2020: 209). The origin of this alternation is due to a prosodically-conditioned deletion rule occurring in West-Germanic. In particular, verbs displaying a heavy syllable (i.e., ending with a complex coda and/or containing a long vowel) in their root happened to lose the linking vowel *-i-* in the preterit form: *\*hōr-jan* / *\*hōr-i-da* > *\*hōr-jan* / *\*hōr-da*, *\*stall-jan* / *\*stall-i-da* > *\*stall-jan* / *\*stall-da*, etc. Subsequently, as a result of Umlaut, the back vowels of the infinitive turned out to be fronted while the root-vowel in the preterit did not: *\*hōr-jan* / *\*hōr-da* > *h[ø:]ren* / *hör-ta*, *\*stall-jan* / *\*stall-da* > *stellen* / *stalta*, etc.<sup>6</sup> On the other hand, verbs not matching the prosodic pattern retained the linking vowel *-i-* and displayed the effects of Umlaut through the whole paradigm, e.g. *\*nar-jan* ‘to save’ / *\*nar-i-da* ‘save.PST.1/3SG’ > OHG *nerien* / *nerita*, etc. In Middle High German (= MHG), these alternations were further developed to the point that they affected dozens of verbs of the weak inflectional class displaying a front vowel in its base form which was likely to be contrasted to a back vowel in the derived form, even in cases where the alternation was not etymologically justified like *liuhten* ‘to lighten’ / *lühte* ‘light.PST.1/3SG’, *lēren* ‘to teach’ / *lārte* ‘teach.PST.1/3SG’, etc. where no linking vowel *-i-* occurred (Paul et al. 2007: 259–263).

In spite of this impressive flourishing, in MSG only remnants of these alternations are found:

- (6) a. *brennen* ‘to burn’ / *brannte* ‘burn.PST.1/3SG’  
*kennen* ‘to know’ / *kannte* ‘know.PST.1/3SG’  
*nennen* ‘to name’ / *nannte* ‘name.PST.1/3SG’  
*rennen* ‘to run’ / *rannte* ‘run.PST.1/3SG’

<sup>6</sup> One reviewer asks whether this has something to do with Ablaut patterns which are normally found in the so-called strong verbs. Although the general morphological pattern involving non-concatenative morphology might well have played a role in favoring the expansion of a non-concatenative pattern, the type of alternation seems to be quite different. Besides the older status of Ablaut patterns, the crucial point is that Ablaut alternations are not connected with the type of vowel involved. Accordingly, we find still in MSG a front vowel in the infinitive alternating with a back vowel in the preterit: *bieten* ‘to offer’ / *bot* ‘offer.PST.1/3SG’, as well as its opposite: *rufen* ‘to call’ / *rief* ‘call.PST.1/3SG’. In other words, the morphological non-concatenative pattern displayed by *Rückumlaut* might well have been supported at a systemic level by the occurrence of Ablaut patterns. This has however nothing to do with the rise of indexical legesigns of the sort represented in Figure 5a and 5b.

- b. *senden* ‘to send’ / *sandte* ~ *sendete* ‘send.PST.1/3SG’  
*wenden* ‘to turn’ / *wandte* ~ *wendete* ‘turn.PST.1/3SG’

The *Rückumlaut* alternations have been strongly reduced to the effect that in MSG only the four verbs in (6a) survive, while the two verbs in (6b) also display forms in which the alternation is levelled out.<sup>7</sup> According to Ronneberger-Sibold (1990), the massive reduction of *Rückumlaut* alternations is due to two main factors, a general systemic factor and a more specific one. Starting with this latter, the decline of *Rückumlaut* verbs begun with the deletion of the schwa-vowel occurring in the competing class: cf. OHG *nerita* > MHG *nerete* > MSG *nährte*. This made the recovery of verb class assignment opaque, because the occurrence of the schwa-vowel before the dental suffix of the preterit used to provide a clear-cut cue to oppose this class (see *neren* / *nerete*) to the *Rückumlaut* verbs where a consonant was normally found in that environment like in *lëren* / *lärte*. Subsequently, the lexical set of *Rückumlaut* verbs was levelled in favor of the more widespread pattern of *neren* where no vowel alternation occurred in the preterit. This leads us to the second systemic factor. In this regard, Ronneberger-Sibold (1990: 127) observes that since in *Rückumlaut* verbs the Umlaut alternation was carried by the morphological simple, i.e. “underived”, form the alternation was levelled because the pattern did not correspond to the general indexical function of signaling a more complex, “derived”, form. Clearly, this leveling supports the indexical role of Umlaut which generally refers to the derivational status of a word with regard to its base form along the scheme depicted in Figure 5a above, while the opposite depicted in Figure 5b is levelled out in the long run.

It has to be stressed that there is nothing in *Rückumlaut* which makes it unfit for further elaboration into an indexical coding which runs into the opposition direction – i.e., back-to-front root-vowel alternation – with regard to the simultaneous occurrence of Umlaut. This scenario is actually found in a Walser German variety spoken in the linguistic island of Gressoney in northwestern Italy (cf. Hotzenköcherle 1956, Zürner 1982: 92, Gaeta ms.). In this variety the alternations reported in Table 1 are found in which the past participles of a certain verbal class systematically display *Rückumlaut* with regard to the respective infinitive or present forms giving rise to four alternation types:<sup>8</sup>

<sup>7</sup> Two further verbs can be held to belong to this class, namely *bringen* ‘to bring’ / *brachte* ‘bring.PST.1/3SG’ and *denken* ‘to think’ / *dachte* ‘think.PST.1/3SG’, but their additional allomorphy speaks in favor of a separate treatment in terms of weak suppletion (see Ronneberger-Sibold 1990 for a discussion).

<sup>8</sup> Note that in the Gressoney orthography the characters <é, ò> stand for [ɪ, u], while <ie, ue> are pronounced at their face value [iɛ, uɛ].

**Table 1:** *Rückumlaut* types of Walser German verbs.

<i>Rückumlaut</i> types	Examples (infinitive / past participle)
RÜ-1: <i>e / a</i>	<i>decke</i> 'to set' / <i>dackt</i> <i>henge</i> 'to hang' / <i>kangt</i> <i>verderpe</i> 'to corrupt' / <i>verdarpt</i>
RÜ-2: <i>ie / ue</i>	<i>ergriene</i> 'to green' / <i>ergruenet</i> <i>rieme</i> 'to praise' / <i>gruemt</i> <i>stiere</i> 'to lean' / <i>gstuert</i>
RÜ-3: <i>é / ò</i>	<i>chnéffe</i> 'to tie' / <i>knòpft</i> <i>drécke</i> 'to print' / <i>dròckt</i> <i>réschte</i> 'to roast' / <i>gròscht</i>
RÜ-4: <i>é / oa</i>	<i>bréme</i> 'to soot' / <i>broamt</i> <i>féerbe</i> 'to color' / <i>gfoarbt</i> <i>méche</i> 'to mow' / <i>gmoot</i>

Two factors are likely to have played a role in favor of the retention of *Rückumlaut* in Gressoney. First, the distribution of Umlaut is substantially limited with respect to MSG, insofar as for instance it is not found with diminutive forms like *bach* 'river' → *bachié* 'river.DIM' and *boum* 'tree' → *boumié* 'tree.DIM' (see respectively MSG *Bach* → *Bächlein* and *Baum* → *Bäumchen*) and with the corresponding plural forms *bacha* 'rivers' and *bouma* 'trees' (see again MSG *Bäche* und *Bäume*). Second, the Umlaut alternation is opaque due to a general process of unrounding of front rounded vowels as shown by *sò / sén* 'son(s)', *mus* 'mouse' / *misch* 'mice', etc. (see the MSG forms *Sohn* / *Söhne* and *Maus* / *Mäuse*). On this basis, it is straightforward to conclude that due to its limited distribution and to its opacity Umlaut could never be interpreted as a clear-cut index signaling a morphological complex or derived status in contrast to MSG, where it massively displaced the *Rückumlaut* alternations.

In sum, the general reductive phonological change accompanying the morphologization of the Umlaut alternation, which essentially reduces the number of segments as in OHG *rōti* > MSG *Röt* of (2) or their feature specification with the rise of central reduced vowels as in OHG *gesti* > MSG *Gäste* of (4), paves the way for the development of its indexical function. Although Umlaut alternations serve several different meanings, they share the same indexical function. Accordingly, the emergence of this indexical legisign can be motivated to provide the advantage of expressing in an explicit way the morphological complex or 'derived' status of a word. This advantage proves useful to the effect that (i) it has outdone the inverse alternation represented by *Rückumlaut* which is substantially reduced in the course of time, and (ii) it is further extended, also to cases where no Umlaut used to be etymologically present. Besides the extension of an additional non-etymological

Umlaut in OHG noun plurals like *biscof / biscofa* ‘bishop(s)’ > *Bischof / Bischöfe*, *naht / naht* ‘night(s)’ > *Nacht / Nächte*, *hano / hanon* ‘rooster(s)’ > *Hahn / Hähne*, etc., we also observe the extension of Umlaut as exclusive plural marker in OHG nouns like *muoter / muoter* ‘mother(s)’ > *Mutter / Mütter*, *fater / fatera* ‘father(s)’ > *Vater / Väter*, *garto / garton* ‘garden(s)’ > *Garten / Gärten*, *fogal / fog(a)la* ‘bird(s)’ > *Vogel / Vögel*, etc., which are especially widespread in southern varieties of MSG (see Wegener 2003: 121 for substandard forms like *Ärm* ‘arms’, *Täg* ‘days’ instead of the MSG forms *Arme* and *Tage*). In contrast to this, in Walser German where Umlaut did not develop any indexical function *Rückumlaut*, alternations flourished while etymological Umlaut alternations were levelled out as observed in plurals like *bach / bacha* ‘river(s)’, *wöre / wörma* ‘worm(s)’ which correspond to MSG *Bach / Bäche* ‘river(s)’, *Wurm / Würme* ‘worm(s)’, and the like.

## 4 Indexicality and the evolutionary perspective

Let us move to the other perspective, which is not emergentist but rather foreshadows a “paleontological” scenario, in which the quest for archaic, “proto-linguistic” indexical legisigns is in focus. In this regard, one straightforward example of an archaic feature surviving through time like the panda’s thumb comes from the Italian particle *ecco* ‘here’ which normally draws the attention towards an entity entering into a speech situation:

- (7) *Ecco il taxi!*  
*ecco* DEF taxi  
 ‘Here’s the taxi!’.

Besides this presentational function of referring to an entity entering into a speech situation, *ecco* can also appear in syntactically non-integrated contexts with a textual value focusing on peculiar articulations of a narrative (8a), or to emphasize particular moments of a discourse (8b) or, finally, as a true discourse marker signalling turn-taking (8c) (see Gaeta 2013 for the sources of the examples):

- (8) a. *E subito ecco m’empie la visione di*  
 and immediately *ecco* 1SG.OBL=fills DEF vision of  
*campagne prostrate nella luce*  
 countries[F] exhaust.PST.PTCP.F.PL in=DEF light  
 ‘And at once I’m filled by a vision of lands bowed in the light.’



- b. A: *Sai cosa facciamo?* B: *Cosa?* A: *Ecco: ti*  
 know.2SG what do.1PL what ecco 2SG.OBL  
*insegno i primi elementi di astrologia.*  
 teach.1SG DEF first.M.PL elements[M] of astrology  
 ‘A: Do you know what we can do? B: What? A: Here’s: I teach you the  
 basics of astrology.’
- c. *ecco, cioè: sono molto preoccupato*  
*ecco* namely am very worry.PST.PTCP.M.SG  
 ‘well, that is. . . I am very worried.’

The particular interest of *ecco* resides in its extraordinary longevity because it basically continues its Latin ancestor *ecce*. The latter is traditionally treated among the interjections, or alternatively as an adverb (see Cuzzolin 1998 for discussion). Its etymology is only partially understood. A common account assumes a complex form *\*ed-ke* in which the 3<sup>rd</sup> person neuter pronoun *\*ed* ‘it’ (see Lat. *id*) is followed by a deictic particle *\*-ke* ‘here’ also found in *ce-do* ‘give here’, *hic* ‘this’ from *\*hi-ce*, *cēterus* ‘the other’ from *\*ce-etero* ‘other here’, etc. (Pokorny 2002: 282). In partial contrast to this reconstruction, de Vaan (2008: 185) observes that “[t]he older etymology of *ecce* as *\*ed-ke* is unlikely, since Latin and Proto-Indo-European (= PIE) had no neuter pronoun *\*ed*, only *\*id*”, from a stem *\*(h<sub>1</sub>)i-* for Nom/Acc sg. vs. *\*h<sub>1</sub>e-sm-/-si-/-i-* elsewhere, and suggests instead a reconstructed form *\*ek-ke*, i.e. a combination of PIE *\*h<sub>1</sub>e-* ‘he’ found in Lat. *is*, *ea*, *id*, extended by means of *\*k*, and of the deictic *\*ke*. Further suggestions for the etymology of the first element of *ecce* involve the particle *ēn* ‘here’ which in imperial age can also occur in combination with *ecce* and is likely to contain a PIE particle *\*h<sub>1</sub>en* found in the Ancient Greek interjection *ēn* meaning ‘hey’ or, finally, the imperative form *em* of the verb *emere* ‘to purchase’ (see recently also Carvajal 2020: 33 for a discussion). In spite of their differences, all hypotheses share a deictic source morpheme, which fits well with its value of attention-driver.

Latin *ecce* is generally used, especially in dialogic texts like plays, to guide the attention of the listener conveying a deictic value referring to a concrete event coming on the stage (9a) or a textual-pragmatic value drawing attention to what the speaker is going to tell (9b), as shown by the following examples from Pre-Classical Latin (examples from Cuzzolin 1998):

- (9) a. *Ei, bene ambula. atque audin etiam? ecce*  
 go.IMP well walk.IMP and hear.IND.PRS.2SG.Q also ecce  
 ‘DEMAENETUS Go, walk good. Are you still listening? LIBANUS Look!’  
 (Plt. *Asin.* 108–109)

- b. *quid igitur faciam miser? quidue*  
 what then do.SBJV.PRS.1SG miserable.M.SG.NOM what.Q  
*incipiam? ecce autem video rure*  
 begin.SBJV.PRS.1SG ecce yet see.IND.PRS.1SG country.ABL  
*redeuntem senem*  
 return.PTCP.PRS old.man.ACC  
 ‘What then shall I do? But look, I see the old gentleman returning from  
 the country.’  
 (Ter. *Eun.* 966–967)

From this original usage not really integrated into the sentence, *ecce* develops a core usage with presentative constructions, i.e. a “thetic” value which is typical of these sentence-focus utterances consisting of monomial predications characterized by the absence of a topical subject (see Lambrecht 1994: 137–146, Sasse 2006). The latter is usually picked out as a predication base in canonical “categorical” utterances which are bipartite predications involving a predication base, the entity about which the predication is made, and a predicate, which says something about the predication base. Categorical utterances normally contain a topical subject and focus on the predicate. On the other hand, in thetic utterances the entire situation, including all of its participants, is asserted as a unitary whole and allows the speaker to focus on the whole sentence:

- (10) a. *Era, ecc-um praesto milit-em*  
 lady.VOC ecce-M.SG.ACC readily soldier[M]-ACC  
 ‘There’s the soldier all ready for you, ma’am!’  
 (Plt. *Mil.* 1216)
- b. *nam meu’ pater intus nunc est*  
 in.fact my.M.NOM father[M].NOM inside now is  
*ecc-um Iuppiter*  
 ecce-M.SG.ACC Jupiter[M].NOM  
 ‘The point is, my father Jupiter is now inside there, mark you.’  
 (Plt. *Am.* 120)
- c. *ecce tibi Sebosus*  
 ecce 2SG.DAT Sebosus.NOM  
 ‘Here’s Sebosus.’  
 (Cic. *Att.* 2, 15, 3)

The indexical value of *ecce* in these thetic sentences is clearly in connection with the introduction of a single focused argument, while the so-called “ethical” dative *tibi* in (10c) merely refers to a fictitious character implicitly present in the conversa-

tion, and cannot be replaced by other persons: *ecce tibi* / \**mihi* *Sebosus*. Because of its pragmatic function, the ethical dative has been labeled conversational dative by Masini (2012). The single occurring argument stands in accusative in Pre-Classical Latin (10a) and is therefore technically governed by *ecce*, which can also display accusative agreement with it. However, given its occurrence as a single argument the alternative nominative marking (10b) is also attested in spite of the accusative marking on *eccum*. Nominative marking becomes canonical in Classical Latin (10c). Since the Pre-Classical period, a significant syntactic elaboration of the dependent argument is observed insofar as modifiers of a different nature, e.g. participles (11a),<sup>9</sup> can be added to the focused noun, while this latter can also consist of a free relative clause (11b):

- (11) a. *Ecce nos tibi oboedientes*  
*ecce* 1PL.ACC 2SG.DAT obey.PART.PRS.PL.ACC  
 ‘Here, sir, ours to obey.’  
 (Pl. *Mil.* 663)
- b. *ecce tibi, qui rex populi*  
*ecce* 2SG.DAT REL.SG.NOM king.NOM people.GEN  
*dominusque omnium gentium esse concupiverit*  
 lord.NOM.CONJ all.PL.GEN people.PL.GEN be.INF desire.SBJV.PST.3SG  
 ‘here’s who wanted to be the king of the Romans and lord of all people.’  
 (Cic. *Off.* 3, 83)

The syntactic elaboration constantly proceeds further in the subsequent centuries. In Post-Classical Latin we observe the expansion to the sentential government of finite clauses of a different nature, firstly with a non-factive (12a) and then with a factive (12b-c) meaning in Late Latin:

- (12) a. *Ecce, ut doleat tibi, ut postea*  
*ecce* that regret.SBJV.PRS.3SG 2SG.DAT that afterwards  
*consideratius loquaris!*  
 appropriate.COMPAR speak.SBJV.PRS.2SG  
 ‘Here’s that you will regret and will talk in a more appropriate way.’  
 (Sen. *Ben.* 4.36)

<sup>9</sup> Notice that the dative *tibi* in (11a) depends on the verb *oboedio* ‘to obey’, and is therefore not an instance of the conversational dative observed in (10c).

- b. *ecce, quid faciunt Christiani*  
*ecce* what do.IND.PRS.3PL Christian.PL.NOM  
 ‘here’s what the Christians do.’  
 (Aug. *Psalm*. 34, 2, 10)
- c. *ecce, quomodo amabat eum*  
*ecce* how love.PST.IPFV.3SG 3SG.ACC  
 ‘here’s how he loved him.’  
 (*Vulg. Ioh.* 11,36)

Thereafter, we can trace the development of *ecce* in its Italian offspring *ecco*, while throughout the other Romance languages it is only sparsely attested in their older stages and fades subsequently out. Remarkably, the Italian offspring *ecco* basically inherits the constructional options of its forerunner, while at the same time it also displays a spectacular expansion of its constructional potential both with regard to the range of dependent constituents (i.e. its “internal” syntax, see Morin 1985 for this concept in connection with the French correspondent *voilà* to which we will come back below), and with regard to the possibility of being inserted into larger syntactic domains (i.e. its “external” syntax). In particular, we observe already in Old Italian since the 13<sup>th</sup> century the extension to the government of non-finite verbs, infinitives as well as past participles, giving rise to a complex verbal predicate:<sup>10</sup>

- (13) a. *ecco sonar un corno e i can baiare*  
*ecce* clang.INF INDF horn and DEF dogs bark.INF  
 ‘here’s a horn clanging – and the dogs barking.’
- b. *Ed ecco verso noi venir per nave un vecchio*  
 and *ecce* towards 1PL come.INF for boat INDEF old  
 ‘And here’s an old man coming toward us by boat.’
- c. *Ecco la potenza dell’amistade generare spregio*  
*ecce* DEF power of.DEF=friendship generate.INF disdain  
*di morte*  
 of death  
 ‘Here’s the power of friendship generating the disdain of death.’
- d. *muorsi costui, e ecco il legame rotto*  
 dies.REFL DEM and *ecce* DEF bond[M] broken.M.SG  
 ‘if this dies, here’s also the bond broken.’

<sup>10</sup> The Old Italian data presented here have been extracted from TLIO, a corpus of Old Italian texts from the 13<sup>th</sup>–14<sup>th</sup> centuries, to which I refer for the exact sources. It is interesting to observe that *ecce* occurs in about 4,000 examples, which testifies to its high frequency (see also Gaeta 2013).

- e. *Ecco giunta*                      *colei*    *che ne*    *pareggia*  
*ecco* arrive.PST.PTCP.F.SG DEM.F.SG REL 1PL.OBL level.3SG  
 ‘Here comes she who equalizes us.’

Notice in particular that the verbal complex can be highly selective. While *ecco* can combine with infinitives of transitive (13a), unergative (13b) and unaccusative (13c) verbs, only past participles of transitive (13d) and unaccusative (13e) verbs are compatible. In other words, *ecco* behaves like an unaccusative auxiliary, which also fits in Lambrecht’s (2000: 659) picture relating tothetic predications because the latter are generally held to be incompatible with unergative verbs typically displaying an agentive subject.

This is further profiled in Modern Italian where *ecco* replicates exactly the same possibilities of Old Italian, especially in the presentational constructions of (13) (cf. Gaeta 2023). In the following modern presentational sentences *ecco* can be held to behave like a sort of auxiliary of what can be called an “ostensive predication”. In fact, the following sentence displays two readings:

- (14) *Ecco arrivati*                      *gli amici*    *dagli*    *Stati Uniti*.  
*ecco* arrive.PST.PTCP.M.PL DEF friends[M] from.DEF USA  
 a. [[[*Ecco arrivati*] *gli amici*] [*dagli Stati Uniti*]].  
 ‘Here arrived the friends from the USA.’  
 b. [[*Ecco arrivati*] [*gli amici dagli Stati Uniti*]].  
 ‘Here arrived the friends who are/come from the USA / the American friends.’

In (14a) the PP *dagli Stati Uniti* directly depends on the whole VP *ecco arrivati* while in (14b) it only depends on the NP *gli amici*. This ambiguity is also found with the construction containing the true present perfect of the verb selecting the unaccusative auxiliary *essere* ‘to be’:

- (15) *Sono arrivati*                      *gli amici*    *dagli*    *Stati Uniti*.  
 are.3PL arrive.PST.PTCP.M.PL DEF friends[M] from.DEF USA  
 a. [[[*Sono arrivati*] *gli amici*] [*dagli Stati Uniti*]].  
 ‘There arrived the friends from the USA.’  
 b. \*[[*Sono arrivati*] [*gli amici dagli Stati Uniti*]].  
 ‘There arrived the friends who are/come from the USA / the American friends.’

Notice that the auxiliary value holding in (14a) disappears if *ecco* is separated by an NP from the past participle:

- (16) *Ecco gli amici arrivati dagli Stati Uniti.*  
 ecco DEF friends[M] arrive.PST.PTCP.M.PL from.DEF USA  
 a. [*Ecco [gli amici [arrivati dagli Stati Uniti]]*].  
 ‘Here are the friends arrived from the USA.’  
 b. \*[[*Ecco gli amici arrivati*] [*dagli Stati Uniti*]].  
 int. ‘Here arrived the friends from the USA.’

In (16a) the past participle can only be interpreted as an attribute of the NP, while the alternative interpretation corresponding to the auxiliary value of (14a) is odd as shown by (16b). Notice that Lambrecht (2000: 656) also discusses this effect of Adjacency as a typical feature characterizingthetic predications and actually underlying the distinction between unaccusative and unergative verbs. As expected, this behavior is paralleled by the construction containing *essere* insofar as it behaves like a copula in a categorical ascriptive predication containing a null subject (17a) while the past perfect interpretation is blocked (17b):

- (17) *Sono gli amici arrivati dagli Stati Uniti.*  
 are.3PL DEF friends[M] arrive.PST.PTCP.M.PL from.DEF USA  
 a. [*Sono [gli amici [arrivati dagli Stati Uniti]]*].  
 ‘They are the friends arrived from the USA.’  
 b. \*[[*Sono gli amici arrivati*] [*dagli Stati Uniti*]].  
 int. ‘There arrived the friends from the USA.’

In this light, *ecco* can be qualified as an “ostensive auxiliary” and serves as an indexical legisign, essentially distinct from the symbolic legisign *essere* because it entails an ostensive dimension pointing to the concrete presence of the referents on the speech stage in (14a) while their arrival might be only narrated by the construction containing *essere* in (15a). This process of auxiliatation testifies to the strong entrenchment of *ecco* into the morphosyntactic system of Modern Italian. This entrenchment is further supported by the capacity of hosting the reflexive marker both with obligatory reflexives (18a) and in the anticausative construction (18b) (see Gaeta 2013 for the sources of the examples and Gaeta 2017 for the anticausative construction):

- (18) a. *Se gli faceva i capelli in piedi eccosene andati i trenta cm avanzati.*  
 if 3SG.OBL do.IPFV.3SG DEF hairs in feet ecco.REFL.PART  
 gone.M.PL DEF thirty cm leave.PST.PTCP.M.PL  
 ‘If he trimmed her hair standing on his feet, here’s that the thirty cm of advantage were gone.’

- b. *se investono 1.000\$... quindi in un solo anno*  
 if invest.PRS.3PL 1.000\$[M.PL] thus in INDF only year  
*eccosi trasformati in circa 12.000\$.*  
 ecco.REFL transform.PST.PTCP.M.PL in about 12.000\$  
 ‘If they invest 1.000\$, thus within one single year here’s that they are transformed into 12.000\$.’

Generally, the combination with the reflexive marker is an exclusive property of the copula *essere*: see respectively *se ne sono andati* ‘they left, lit. REFL PART are.3PL gone.M.PL’ and *si sono trasformati* ‘they were transformed, lit. REFL are.3PL transform.PST.PTCP.M.PL’.

Furthermore, we find already in Old Italian the extension to the sentential government of several types of dependent finite clauses, namely declarative clauses (19a), temporal clauses (19b), WH-questions (19c), manner questions (19d), as well as the apodosis of a hypothetical sentence (19e):

- (19) a. *Ecco che bona fine venuta è*  
 ecco that good.F.SG end[F] come.PST.PTCP.F.SG is  
 ‘Here’s that a good end has come.’
- b. *Ecco quando è il tempo d’innestare l’uno arbore*  
 ecco when is DEF time of=graft.INF DEF=one tree  
*coll’altro*  
 with.DEF=other  
 ‘Here’s when it’s time to graft one tree with the other.’
- c. *ecco a cui dimanda dove sono li virtuosi*  
 ecco to INTERR.OBL asks where are.3PL DEF honest.M.PL  
*nominati di sopra.*  
 mention.PST.PTCP.M.PL of above  
 ‘Here’s who to ask where the honest persons mentioned above can be found.’
- d. *Ecco dunque come Idio s’era mostrato*  
 ecco thus how God REFL=was shown  
 ‘Here’s therefore how God had shown himself.’
- e. *se entrarò nella città, ecco che sono*  
 if enter.FUT.1SG in=DEF town ecco that are.3PL  
*dimagrati per la fame*  
 slim.PST.PTCP.M.PL for DEF hunger  
 ‘if I will enter into the town, here’s that they will be starving.’

One further remarkable extension observed in Old Italian with regard to Latin is the complexification of the external syntax of *ecco*, i.e. its employment within subordinate clauses of different types, for instance temporal (20a) and causal (20b) clauses:

- (20) a. *Di poco aveva così parlato, che ecco venire Seges*  
 of few had.3SG so spoken that *ecco* come.INF Seges  
*molto in fretta*  
 much in hurry  
 ‘He had just talked, that there was Seges coming very quickly.’
- b. *Io son morta, ché ecco il marito mio*  
 1SG am dead[F] ‘cause *ecco* DEF husband 1POSS  
 ‘I am dead, because here’s my husband.’

Finally, the extension of *ecco* to the occurrence in relative clauses is observed in Modern Italian as in the following example from the Internet:

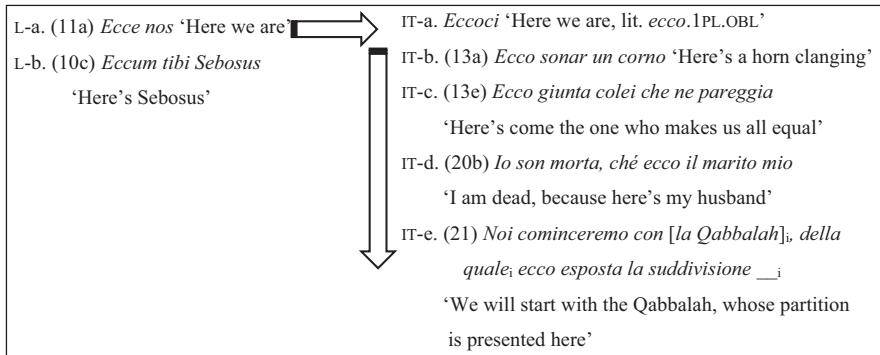
- (21) *cominceremo con [la Qabbalah], della quale<sub>i</sub> ecco*  
 start.FUT.1PL with DEF Qabbalah of=DEF which *ecco*  
*esposta la suddivisione<sub>i</sub>*  
 expose.PST.PTCP.F.SG DEF partition[F]  
 ‘We will start with the Qabbalah, whose partition is presented here.’

In this way, the formally subordinate clauses are presented as the main part of the message from an information-structural perspective. This effect is clearly due to the general indexical, ostensive value of *ecco*, which intrinsically links the clause to the speech context and gives rise to an interpretation whereby the clause containing *ecco* is communicatively primary or at least on a par with the formally superordinate clause (I thank a reviewer for this observation).

In sum, we observe from an evolutionary perspective that this “primitive” indexical legisign worms its way up into the Italian morphosyntax, expanding its use in conventionalized – i.e. highly symbolic – constructions (see Figure 6):

- First, it retains its core presentative usage through its expansion from Latin to Old Italian (see L-a-b and IT-a in Figure 6);
- Second, it can be further combined in Old Italian with a non-finite verb form (either an infinitive or a past participle) giving rise to a complex predicate behaving like an ostensive predicate (see IT-b-c in Figure 6);
- Finally, it can be embedded only in Modern Italian into a causal and a relative clause (see IT-d-e in Figure 6).





**Figure 6:** The expansion of *ecce/ecco* as a primitive particle.

Thus, its primitive indexical value does not prevent *ecco* from being wired to highly complex constructions as a true – though defective – verbal predicate, as already envisaged by Hall (1953: 279), who suggests “to classify ECCE and its descendants as verbs – defective, of course, since they are always monomorphemic”. On the other hand, the persistence of *ecce* > *ecco* is amazing: preserving its original nature of indexical legisign, it covers more than two millennia! It has to be stressed that since the early beginnings its primitive indexical value is strictly connected with:

- Initial clause position, independently of its embedding into compound sentences;
- Low degree of morphological integration, which is expressed by the lack of any morphological trait (possible agreement markers occurring in Latin were subsequently lost), which points to a scarce degree of word-class categorization;
- Persistent deictic and textual-pragmatic usages, already present in Latin (see (9) above), and still largely possible in Modern Italian (see (8) above).

At any rate, we record a constant development of a wide range of verbal characters relating to its internal and external syntax as a thetic predicate, tremendously expanded in Italian in which it gives rise even to complex predicates resembling the behavior of an unaccusative auxiliary, for instance hosting clitics as typically done by verbs. In other words, the indexical particle preserves its “primitive” nature, being partially non-integrated in syntactic contexts, while it is at the same time largely entrenched as an auxiliary into the grammatical core of Italian.

How can we account for this situation in which the persistence of primitive traits is accompanied by a lively blossoming of grammatical entrenchment? One suggestion which fits well with the paradoxical combination of retention and innovation displayed by *ecce* > *ecco* comes from the evolutionary perspective connected with a protolinguistic scenario. In fact, among the main properties characterizing

proto-language against full-blown language, Jackendoff (2011: 614–615) lists (i) the scarcity or absence of constituent structure; (ii) a pragmatically based word-order driven by default principles such as Agent First / Focus Last, and Modifier Adjacent to Modified, and (iii) a scarce or absent degree of word-class categorization because all combinatoriality is driven by phonology and semantics (see also Botha 2016: 86–92 and *passim* for further discussion). These properties outline a scenario in which the symbolic dimension is embryonal, no conventionalized constructs occur and the semantic dimension is highly profiled in terms of very simple instructions typical of indexical signs. In this perspective, *ecce* > *ecco* might represent a fossilized relic of protolanguage in Jackendoff's sense, retaining primitive properties and being particularly salient in the face-to-face communication. In fact, it is typically found in presentational constructions where (i) it occupies the first sentence position and displays a limited combinatorics; (ii) it indexes focused constituents which come later in the clause and is sensitive to Adjacency; and (iii) finally, it retains a scarce degree of word-class categorization. This does not deny its nature of indexical sign given that it displays a conventionalized form-meaning relation and has entered a certain number of highly conventionalized grammatical constructions. In a way, the persistence of *ecce* > *ecco* reminds us of the panda's thumb mentioned above, which has evolved as a fossilized relic within a highly complex organism serving a new function, namely arboreal locomotion. In the case of *ecce* > *ecco* the indexical "pointing" function serving as a primary cognitive scaffolding in Jackendoff's terms evolved towards the enhancement of its presentational value as deictic anchor of the sentence by developing a certain number of verbal properties.

In addition, this relic displays the traits of a true success story insofar as a similar development is observed in the French particles *voilà/voici*, which display a fully different substance but similar behavioral properties (Morin 1985, Bergen & Plauché 2005, Gaeta 2023). Remarkably enough, Italian is the only Romance language attesting the modern occurrence of Latin *ecce*, while the particle *ez* still attested in Old French as an offspring of *ecce* completely disappeared. Instead, the particles *voilà/voici* were subsequently grammaticalized – from an original indexical expression *voi là/ci!* 'see.IMP there/here!'. It has to be stressed that the development of *voilà/voici* covers exactly the same functional space displayed by *ecco*, which stands in continuity with its ancestor *ecce*. From our perspective, this might be treated as substantial evidence supporting the role of these indexical signs as remnants and witnesses of a primary / primitive cognitive scaffolding on which full-blown languages are based in evolutionary terms. While the primitive properties are retained, in French the indexical sign has been renewed recruiting a more transparent indexical expression like *voilà* 'look there', but involving again a deictic particle used in face-to-face visual communication.

## 5 Conclusion

Indexicality has a peculiar status as a coding technique insofar as it normally draws the attention towards a significant semiotic event upcoming in the neighborhood. Emergent indexicality results in the case of MSG Umlaut from disturbed symbolic coding because of the effect of reductive phonological change, which has rendered opaque the relation between base and derived forms. The alternation is reanalyzed as indexing the derived status of a word, i.e. the morphologically complex status of a word within a word family, and accordingly makes reference to a paradigmatic network of lexical relations.

From an evolutionary perspective, we discussed the case of the indexical particle *ecce* > *ecco*, typically characterizing presentational constructions, which can be held to display impressively persistent primitive properties referable to the primary cognitive scaffolding on which mature language is based. Accordingly, it qualifies as a relic of protolanguage, a sort of panda's thumb, partially assimilated to the morphosyntactic profile of the Italian complex verbal predicates. The primitivity of this relic and of its cognitive scaffolding has also left traces in French, where the indexical *leisign* has been renewed by a process of grammaticalization but entirely retains the primitive properties of such a primary cognitive scaffolding. It has to be hoped that future research will pursue this double goal of investigating the role of indexicality as a factor emerging among others from the refunctionalization of disturbed symbolic coding, as well as its persistence in fossilized relics which allow us to gain insight into the primitive scaffolding characterizing protolanguage.

## References

- Andersen, Henning. 2010. From morphologization to demorphologization. In Silvia Luraghi, & Vit Bubenik (eds.), *Continuum Companion to Historical Linguistics*, 117–146. London/New York: Continuum.
- Andersen, Henning. 2020. Synchrony, diachrony, and indexicality. In Bridget Drinka (ed.), *Historical Linguistics 2017. Selected papers from the 23<sup>rd</sup> International Conference on Historical Linguistics, San Antonio, Texas, 31 July – 4 August 2017*, 163–182. Amsterdam/Philadelphia: John Benjamins.
- Bergen, Benjamin K. & Madelaine C. Plauché. 2005. The convergent evolution of radial constructions: French and English deictics and existentials. *Cognitive Linguistics* 16(1). 1–42.
- Botha, Rudolf P. 2016. *Language evolution: the windows approach*. Cambridge: Cambridge University Press.
- Carvajal, Camilo Andrés Bonilla. 2020. The syntax of the Latin presentative adverb *ecce*: Relation to focus phrase. *Journal of Latin Linguistics* 19(1). 27–57.

- Cuzzolin, Pierluigi. 1998. Quelques remarques syntaxiques à propos de *ecce*. In Benjamín García Hernández (ed.), *Estudios de lingüística latina. Actas del IX Coloquio Internacional de Lingüística Latina*, 261–271. Madrid: Ediciones Clásicas.
- de Vaan, Michiel. 2008. *Etymological Dictionary of Latin (and the other Italic Languages)*. Leiden: Brill.
- Deacon, Terrence W. 1997. *The Symbolic Species*. New York: Norton & Company.
- Deacon, Terrence W. 2012. The symbol concept. In Maggie Tallerman & Kathleen R. Gibson (eds.), *The Oxford Handbook of Language Evolution*, 393–405. Oxford: Oxford University Press.
- Eisenberg, Peter. 2020. *Grundriss der deutschen Grammatik. Das Wort*. 5<sup>th</sup> ed. Stuttgart: Metzler.
- Everett, Daniel. 2017. *How Language Began: The Story of Humanity's Greatest Invention*. New York: Norton.
- Gaeta, Livio. 1998. Some remarks on analogy, reanalysis and grammaticalization. In Anna Giacalone Ramat & Paul J. Hopper (eds.), *The Limits of Grammaticalization*, 89–105. Amsterdam/Philadelphia: John Benjamins.
- Gaeta, Livio. 2002. Growth of symbols: The inexorable fate of diagrams. In Katarzyna Dziubalska-Kolaczyk & Jarek Weckwerth (eds.), *Future Challenges for Natural Linguistics*, 129–151. München: Lincom.
- Gaeta, Livio. 2004. Exploring grammaticalization from below. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What Makes Grammaticalization? A Look from its Fringes and its Components*, 45–75. Berlin/New York: Mouton de Gruyter.
- Gaeta, Livio. 2013. *Ecco, ecco, l'ho trovata*: La tenace persistenza di un'impalcatura cognitiva primaria. In Sabine De Knop, Fabio Mollica & Julia Kuhn (eds.), *Konstruktionsgrammatik in den romanischen Sprachen*, 45–74. Frankfurt/Main: Lang.
- Gaeta, Livio. 2016. Co-opting exaptation in a theory of language change. In Muriel Norde & Freek Van de Velde (eds.), *Exaptation in language change*, 57–92. Amsterdam/Philadelphia: John Benjamins.
- Gaeta, Livio. 2017. Valency alternations between inflection and derivation: a contrastive analysis of Italian and German. In Lars Hellan, Andrej Malchukov & Michela Cennamo (eds.), *Contrastive Studies in Verbal Valency*, 327–373. Amsterdam/Philadelphia: John Benjamins.
- Gaeta, Livio. 2022. Dangerous liaisons: An introduction to derivational paradigms. In Alba E. Ruz, Cristina Fernández-Alcaina & Cristina Lara-Clares (eds.), *Paradigms in Word Formation: Theory and Applications*. Amsterdam, 3–18. Philadelphia: John Benjamins.
- Gaeta, Livio. 2023. Pure existentials vs. pure presentationals: finding an existence out(side) of place. In Laure Sarda & Ludovica Lena (eds.), *Existential Constructions Across Languages: Forms, Meanings and Functions*, 102–138. Amsterdam/Philadelphia: John Benjamins.
- Gaeta, Livio. Ms. *Inflectional Classes and the Morphology-Syntax Interaction: Strong and Weak Verbs in Walser German*. University of Turin.
- Hall, Robert A. Jr. 1953. The classification of Italian *ecco* and its cognates. *Romance Philology* 6(4). 278–280.
- Hopper, Paul J. 2015. An emergentist approach to grammar. In Brian MacWhinney & William O'Grady (eds.), *The Handbook of Language Emergence*, 314–327. Oxford: Wiley-Blackwell.
- Jackendoff, Ray. 2009. Compounding in the parallel architecture and conceptual semantics. In: Rochelle Lieber & Pavel Štekauer (eds.), *The Oxford Handbook of Compounding*, 105–128. Oxford: Oxford University Press.
- Jackendoff, Ray. 2011. What is the human language faculty? Two views. *Language* 87(3). 586–624.
- Jakobson, Roman. 1965. Quest for the essence of language. *Diogenes* 51 [reprinted in *Selected Writings II*, 345–359. The Hague/Paris: Mouton 1971].
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

- Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence focus constructions across languages. *Studies in Language* 24(3). 611–682.
- Leone, Massimo. 2021. Indexes: Cultural nature and natural culture. *Rivista di estetica* 76. 112–129.
- Masini, Francesca. 2012. Costruzioni verbo-pronominali ‘intensive’ in italiano. In Valentina Bambini, Irene Ricci & Pier Marco Bertinetto (eds.), *Linguaggio e cervello – Semantica. Atti del XLII Congresso Internazionale della Società di Linguistica Italiana (Pisa, SNS, 2008)*, vol. 2, paper no. II.C.2. Roma: Bulzoni.
- Morin, Yves-Charles. 1985. On the two French subjectless verbs *voici* and *voilà*. *Language* 61(4). 777–820.
- Nielsen, Peter Juul. 2016. *Functional Structure in Morphology and the Case of Nonfinite Verbs: Theoretical Issues and the Description of the Danish Verb System*. Leiden: Brill.
- Norde, Muriel & Freek Van de Velde (eds.). 2016. *Exaptation in Language Change*. Amsterdam/Philadelphia: John Benjamins.
- Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. Revised edition by Thomas Klein, Hans-Joachim Solms & Klaus-Peter Wegera. Tübingen: Niemeyer.
- Peirce, Charles S. 1965. *Collected Papers*. Edited by Charles Hartshorne & Paul Weiss. Cambridge, MA: The Belknap Press of Harvard University Press.
- Pokorny, Julius. 2002. *Indogermanisches etymologisches Wörterbuch*. 5th edn. Bern: Francke.
- Ronneberger-Sibold, Elke. 1990. Zum Auf- und Abbau des Rückumlauts bei den schwachen Verben. In Armin Bassarak (ed.), *Wurzel(n) der Natürlichkeit*, 119–132. (Linguistische Studien, Reihe A: Arbeitsberichte 208). Berlin: Zentralinstitut für Sprachwissenschaft
- Sarbo, Janos J. 2006. Peircean proto-signs. *AIP Conference Proceedings* 839. 474–479.
- Sasse, Hans-Jürgen. 2006. Theticity. In Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe*, 255–308. Berlin/New York: Mouton de Gruyter.
- Scott-Phillips, Thom. 2015. *Speaking our minds*. New York: Palgrave Macmillan.
- Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.
- Wegener, Heide. 2003. Normprobleme bei der Pluralbildung fremder und nativer Substantive. *Linguistik online* 16(4). 119–157.
- Wurzel, Wolfgang U. 1981. Phonologie: Segmentale Sutrktur. In Karl E. Heidolph, Walter Flämig & Wolfgang Motsch (eds.), *Grundzüge einer deutschen Grammatik*, 898–990. Berlin: Akademie-Verlag.
- Wurzel, Wolfgang U. 1984. Was bezeichnet der Umlaut im Deutschen? *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 37(6). 647–663.