Introducing DEmA: the Pavia Diachronic Emergence of Alignment Database

Cristofaro Sonia, Inglese Guglielmo

Introducing DEmA: the Pavia Diachronic Emergence of Alignment database

Sonia Cristofaro (Sorbonne Université) & Guglielmo Inglese (KU Leuven – FWO)

Abstract

The Pavia *Diachronic Emergence of Alignment* (DEmA) database is a new resource for the study of the diachrony of alignment patterns cross-linguistically. In this paper, we offer a description of DEmA, its structure and the choices that have been made in its construction. The main goal of DEmA is to offer a platform that makes it possible to investigate the sources and processes out of which new alignment patterns come into being across languages. In order to do so, each instance of the emergence of a construction with a new alignment pattern is decomposed into a number of well-defined parameters pertaining to the initial situation in the language, the developmental mechanisms leading to the new alignment pattern, and the effects of the change. These various parameters are effectively implemented into a searchable format. This systematization enables users to easily retrieve and compare various type of information concerning the emergence of alignment patterns in the world's languages.

1. Introduction

Over the past decades, typologists have repeatedly stressed the importance of taking diachronic information into consideration when explaining cross-linguistic regularities (see recently Grossman

,

& Polis 2018, Cristofaro 2019, Haspelmath 2019). Unfortunately, resources providing information on how specific phenomena develop over time cross-linguistically are not numerous. Progress in grammaticalization studies and historical linguistics has brought to light an increasing body of evidence regarding the possible origins of different alignment patterns. Information about these processes is, however, scattered across specialized publications, and often not easily comparable from one language to another, nor accessible to non-specialists.

In this paper, we introduce the Pavia *Diachronic Emergence of Alignment* (DEmA) project. The project aims to build a comprehensive open access database on the emergence of alignment patterns cross-linguistically, so as to complement existing typological databases on alignment, for example the three WALS chapters devoted to this topic (Comrie 2013a; Comrie 2013b; Siewierska 2013), which only provide information about synchronic patterns.

The data in DEmA is systematized in such a way that one can readily search and compare various type of information pertaining to the role of the different components at play in the emergence of new alignment patterns. In particular, we propose to decompose the emergence of alignment patterns into three notionally distinct domains: the initial stage of the language, the developmental mechanism, and the results of the change.

The paper is organized as follows. In Section 2, we briefly outline current issues in the diachronic study of alignment patterns and discuss the possible research questions that DEmA will make it possible to explore. Section 3 focuses on the structure of DEmA: we first describe the parameters relevant to the initial situation of the language (Section 3.1) and developmental mechanisms (Section 3.2). W then move to the parameters describing the effects of the change on the global alignment of the language (Section 3.3). Section 4 deals with the practical aspects of how queries can be carried out in DEmA.

2. Alignment patterns in diachrony

,

By alignment pattern is meant here, in a maximally general sense, any possible grouping of the three argument roles A, S, and P (Comrie 1989, Dixon 1994), in terms of case marking (nominal inflection, adpositions, clitics), indexation, or other morphosyntactic phenomena.

Progress in grammaticalization studies and the study of language change cross-linguistically means that a comparatively large body of data is now available on the emergence of alignment patterns in a variety of languages across different families and geographical areas (see, for example, Gildea 1998 on Carib; König 2008 on African languages; Bubenik 1998, Haig 2008 and 2017, Verbeke 2013 on Indo-Aryan). This evidence, however, has not yet been integrated into a comprehensive overview of the possible sources and developmental mechanisms that can give rise to particular alignment patterns (for example, accusative, ergative, or active) from one language to another. An early study in this direction is Harris and Campbell (1995: chap. 9), which, however, concentrates on possible mechanisms of alignment change, rather than the specific alignment patterns emerging through each mechanism, or the source constructions that can give rise to individual patterns. Another strand of cross-linguistic research (e.g. Heine & Kuteva 2002 [now Kuteva et al. 2019]; Kulikov 2006) has focused on the etymology of particular case markers, irrespective of the contexts and developmental mechanisms that lead to these markers evolving from particular source elements, or the consequences of this process for the alignment patterns of the language.

In general, research on the emergence of alignment patterns in individual languages has shown that individual patterns typically emerge from pre-existing constructions, through various mechanisms of constructional reinterpretation or, sometimes, phonological change. The main goal of DEmA is to provide an expanding platform where the available evidence on these processes is integrated in a typologically informed framework that makes it possible to compare different processes from one language to another, so as to obtain data both on the emergence of alignment patterns in particular languages, and on the possible sources and developmental processes leading to

,

the emergence of particular alignment patterns cross-linguistically. This type of data can be used to address different research questions about the diachronic origins of alignment (Harris and Campbell 1995; Gildea 1998; Mithun 2005; Creissels 2008; Cristofaro 2012, 2013, 2014, Zúñiga 2018 among others):

- What source constructions give rise to particular alignment patterns cross-linguistically?
- What developmental mechanisms lead from particular source constructions to particular alignment patterns?
- What is the relationship between the properties of particular source constructions and developmental mechanisms and the properties of the resulting alignment pattern, in terms for example of what argument roles are or are not encoded in the same way, or the distribution of the pattern across different contexts (NP-based and TAM-based alignment splits, or other types of splits)?
- The same alignment patterns (for example, ergative or accusative alignment) originate from different source constructions and through different developmental mechanisms in different cases. Can individual patterns be explained in terms of some overarching principle that applies to all instances of the pattern, or should different instances of the pattern be explained in terms of different principles depending on the developmental processes involved?

3. The organization of DEmA

In DEmA, each entry is a process that has led to the development of a construction with a new alignment pattern in some language, as described in published sources. At present, we focus on monotransitive alignment (i.e. alignment of one- and two- place verbs) only.

,

In line with a number of cross-linguistically oriented accounts (see, for example, Harris and Campbell 1995: Chap. 9), the development of a new alignment pattern is conceived as a process that takes place within particular constructions, for example through the reinterpretation of the argument structure of these constructions, or through the development of a new marker for A, P, or S arguments as a result of grammaticalization. This process will lead to the development of a particular alignment pattern for the construction in question, and may have different effects depending on the original alignment pattern of the language in the relevant grammatical domain. For example, the development of a new perfective construction with ergative alignment may lead to a TAM based split if non-perfective constructions use a non-ergative pattern. If these constructions have ergative alignment, however, the language will remain consistently ergative.

The most innovative feature of DEmA is that it allows for a fine-grained research of the various components involved in the emergence of new alignment patterns. In particular, DEmA is structured so as to provide information about three different domains:

- The initial situation in the language, including both the original alignment pattern of the language and a detailed description of the source construction involved in the emergence of the new alignment pattern.
- Developmental mechanisms, that is, the nature and dynamics of the change that gives rise to the new alignment pattern.
- The effects of the process of change, including the alignment pattern that develops in the construction undergoing the change and the effects of this development on the global alignment pattern of the language.

For each of these domains, DEmA offers multiple searchable fields, which are described in detail in the reminder of this section.

,

3.1. The initial situation in the language

This domain pertains to the situation in the language before the emergence of the new alignment pattern. Two distinct fields are provided:

- **Original alignment pattern**: This refers to the alignment patterns originally attested in the language, along with any constraints in the distribution of these patterns, e.g. accusative, ergative, TAM or NP based splits, and the like.

Only the alignment pattern pertaining to the grammatical domain involved in the process of change is taken into account. For example, if a process of change involves alignment in indexation, only the alignment pattern originally found for indexation in the language (and not, for example, case marking alignment) is taken into account.

- **Source construction**: This refers to the construction that serves as the basis for the development of the new alignment pattern.

In this field, we focus on the specific elements that undergo change in the development of the new alignment pattern (for example, particular lexical items that that grammaticalize into case markers, particular adpositions or case affixes that undergo a change in their grammatical function). While we try to standardize the terminology used in the description of different source constructions cross-linguistically, this field contains highly heterogenous and language-specific descriptions. This is due to the fact that, for each language, different semantic, pragmatic or morphosyntactic properties of the source

,

construction must be taken into account that play a role in the development of the new alignment pattern.

As an example, consider the development of accusative case marking alignment through the reinterpretation of a construction involving the verb *bǎ* 'take' in Mandarin Chinese. The entry for this change in DEmA is shown in Figure 1.

The Pavia DEmA (Diachronic Emergence of Alignment) Database

Home   Languages   Search   How to use DEmA   Credits and contacts

Home › Languages › Mandarin Chinese

Mandarin Chinese

Glottocode: mand1415

Genalogical classification: Sino-Tibetan

Original alignment pattern: Neuter

Source construction: Constructions of the type 'take X (and) VERB (X)', where the verb bǎ 'take' and some other verb share a P argument, as in (1).

Developmental mechanism: The verb bǎ 'take' is reinterpreted as a marker for the shared P argument, and the original biclausal construction is reanalyzed as a monoclausal construction 'ACC X VERB'.

Resulting construction: Transitive construction with a P argument overtly marked by bǎ.

Type of change: Grammaticalization

Alignment in the resulting construction: Nominative-Accusative

Global alignment pattern following the change: Nominative-Accusative

Grammatical domains: Case marking

Symmetry: Asymmetric

Type of data: Historical data

References: Li & Thompson 1974

Figure 1. The emergence of accusative alignment in Mandarin Chinese in DEmA

The language originally had neuter case marking alignment, that is, A, S, and P arguments were not distinguished in terms of case marking. In constructions of the type 'take X (and) VERB (X)', where the 'take' verb and some other verb share a P argument, the 'take' meaning was lost, so that *bǎ* evolved into a marker for its former direct object, 'ACC X VERB'. This is shown by the contrast between the two sentences in (1) and (2), which illustrate, respectively, the use of *bǎ* as a lexical verb and its use as a direct object marker.

,

(1) Classical Chinese (Sino-Tibetan; Li & Thompson 1974: 202)[1]

Yù	*qīng*	***bǎ***	*tīan*	*zhǐ*	*ruì-lìng*	*yǐ*	*zhēn*

yǒu

Yu	himself	take	heaven	POSS	mandate	to	conquer

PTCL *Miáo*

Miao

'Yu himself took the mandate of heaven to conquer Miao.' (Mè-zǐ, 5th century BCE)

(2) Mandarin Chinese (Sino-Tibetan; Li & Thompson 1974: 203)

Tāmen	*bǎ*	*Zhāg-sān*	[...]	*jǐantao le*	*lǐan*	*xîaoshi*
They	ACC	Zhang-san		scrutinize	ASP	two	hours

'They scrutinized Zhang-san for two hours.'

In the DEmA entry for this process, the field 'original alignment pattern' has 'Neuter', whereas the source construction field provides a description of the construction that gave rise to the accusative pattern: "constructions of the type 'take X (and) VERB (X)', where the verb *bǎ* 'take' and some other verb share a P argument".

The need to distinguish between the source construction and the original alignment pattern attested in the language for the relevant grammatical domain is motivated by the fact that (i) the processes that give rise to a new alignment pattern take place within particular constructions, and may be independent of the alignment patterns previously attested in the language, but (ii) the global effects of individual processes in the language will depend on these patterns. For example, ergative patterns have been shown to develop as

---

[1] Glosses and translations of examples are generally taken from the sources. A list of all abbreviations can be found at the end of this paper.

,

intransitive resultative constructions with an oblique NP are reinterpreted as transitive ones, so that the S argument in the intransitive construction becomes a P argument, whereas the oblique NP becomes an A argument ('X is VERBed by Y > 'Y ERG VERBed Y': Gildea 1998, among others). This process will give rise to ergative alignment for resultative constructions, and is independent of the original alignment of S arguments, for example whether they are aligned with A (accusative alignment) or P (ergative alignment). The original alignment of S arguments, however, will determine the global effects of the process in the language. If S arguments were originally aligned with P arguments, the process will only lead to the development of an additional ergative pattern in the language, specialized for resultative constructions. By contrast, if S arguments were originally aligned with A arguments, this alignment will be retained for non-resultative constructions, leading to a split between accusative alignment in non-resultative constructions and ergative alignment in resultative ones.

A well-known example of this development comes from Indo-Aryan languages (see Dahl & Stroński 2016 with extensive references), where a tense-based split-ergative system arose through the reinterpretation of Old Indo-Aryan resultative participial constructions with nominatively marked S and instrumental A, as in (3), as transitive constructions with ergative marking on A, as in (4). Notably, while there is a general consensus that the participial construction with -*ta* in (3) served as the basis for the emergence of a new ergative pattern, whether the ergative postposition =*ne* of Modern Indo-Aryan languages, such as Hindi in (4), is a direct continuant of the Old Indo-Aryan instrumental case marking -*eṇa* remains a matter of dispute (see Verbecke & De Cuypere 2009).

(3) Vedic (Indo-European; Dahl & Stroński 2016: 18)

    ha-tā́   *índr-eṇa*   *paṇay-aḥ*   *śay-adhve*

,

kill-PPP.NOM.PL.M   Indra-INS   Pani-PPP.NOM.PL.M   lie_down-2PL.PRS.MID

'You Panis lie down smashed by Indra.'

(4) Hindi (Indo-European; Dahl & Stroński 2016: 12)

laṛke=ne   *kitāb*   *paṛhī*

boy=ERG   book(F):ABS   read:PST.PRF.F.SG

'The boy has read the book'

3.2. Developmental mechanisms

For this domain, we provide a number of fields pertaining to various aspects of the processes whereby the source construction gives rise to a new alignment pattern:

- **Developmental mechanism**: This field features a description of the mechanisms whereby the source construction gives rise to the new alignment pattern.

For example, the developmental mechanism whereby the Classical Chinese verb *bǎ* 'take' develops into an accusative marker in Mandarin Chinese is described in DEmA as follows "The verb *bǎ* 'take' is reinterpreted as a marker for the shared P argument, and the original biclausal construction is reanalyzed as a monoclausal construction 'ACC X VERB'".

- **Intermediate stages**: This is an optional field that is used in case the historical scenario can be described as unfolding in a number of distinct steps.

In some cases, for example, a new alignment pattern initially develops in particular constructions, and is subsequently extended to other constructions. A case in point is the

'

development of a new split intransitive system in Series II verbs in Georgian. As discussed by

Harris (2010: 213-216), these verbs originally had ergative alignment, but later developed a

split intransitive pattern. This process started from transitive constructions with light

(semantically generic) verbs such as `do, make' and an incorporated object. These

constructions were reinterpreted as intransitive ones, e.g. `gave a shout > shouted', as in (5).

In the resulting intransitive construction, the S arguments maintains the same marking of the

A argument from which it is derived, leading to an accusative pattern initially restricted to the

verbs that were derived in this way. A second step in the process was the extension of this

pattern to all active intransitive verbs in Series II. As other intransitive verbs in the series

maintained ergative alignment, this gave rise to a split intransitive pattern.

(5) Georgian (Kartvelian; Harris 2010: 215)

gagad-q'o   q'ovel-man   er-man

shout-make   all-ERG   people-ERG

'All the people shouted, gave a shout.'

- **Type of change**: This field provides a typological classification of different types of
  developmental mechanisms.

While this classification involves abstracting away from the details of individual processes of

change (for which the user is referred to the relevant sources), it aims to relate these processes

to the general mechanisms of change traditionally discussed in grammaticalization studies and

historical linguistics. We identify five main types of change (note that multiple such

mechanisms may be at play for individual types of change): grammaticalization,

reinterpretation of argument structure, extension, phonological change, loss.

,

– *Grammaticalization*: An element not originally used to encode grammatical relations (e.g. a verb form, a demonstrative, a topic marker) grammaticalizes into a marker for A, S, or P arguments (see Lehmann 2015).

An example of this change is the development of an accusative marker from a 'take' verb in Mandarin Chinese, as described above in (1) and (2). In this case, the grammaticalization of the 'take' verb into a direct object marker leads to the development of dedicated marking for P arguments, whereas A and S arguments remain undifferentiated, yielding an accusative pattern.

– *Reinterpretation of argument structure*: a new alignment pattern emerges through the reinterpretation of the argument structure of the source construction

This type of change, which has also been described as *reanalysis* (see Harris & Campbell 1995: Chap. 4; De Smet 2009), is illustrated by Hanis Coos. In this language, an ergative marker *x=* is derived from an instrumental marker. Mithun (2005) submits that this is a result of a reinterpretation processes that took place in two types of constructions: passive sentences with 1st/2nd person P and a 3rd person oblique A marked with *x=*, as in (6)a, and transitive sentences with an instrumental NP likewise marked with *x=* and no overt third person A, as in in (6)b. Passive constructions such as (6)a are the only possible strategy to encode combination of 1st/2nd person P and third person A in the language As a consequence, the distinction between active and passive is blurred in these contexts, so that the passive construction can be reinterpreted as a transitive construction with the oblique agent becoming an A argument. Similarly, given the lack of an overt A argument in (6)b, in this construction the originally instrumental NP can be reinterpreted as an A. In both cases, the reinterpretation of the source constructions leads to a new alignment pattern, in which the original

'

instrumental/oblique marker *x̣=* is reinterpreted as an ergative marker for A arguments, as in (6)c.

(6) Hanis Coos (Coosan; Mithun 2005: 87, 84)

a. ***x̣=lau***   *kwanł*   *tə=n=tsx̣ewé-i:ł*   *tə=x̣*   *hú:mɨs*

     OBL=that_one seems-will   that=1SG=kill-PASS   that=OBL   woman

     'I may be killed by that woman.'

b. *k'wɨn-t*   ***x̣=mɨl:aqətš***

     shoot-TRANS   OBL=arrow

     '(He) shot at him with an arrow.'

c. ***x̣=yɨqántštextbarime:x̣***   *mæ*   *hanƛ*   *eʔkwɨnai:ł*

     ERG=last   people shall   they_see_thee

     'The last generation shall see you.'

– *Extension*: the markers used for particular argument roles are extended to other roles (e.g from A to S) or the same roles in other contexts (e.g. from the S arguments of particular intransitive verbs to the S arguments of other intransitive verbs).

Consider the case of Bats (see Harris 2010: 210-213). In origin, Bats had distinct indexes for 1st/2nd person A and S roles, as in (7)a and (7)b, respectively. Later on, the index for A was also analogically extended to the S of intransitive verbs with A-like properties (possibly as a result of contact with Georgian), leading to the rise of a new accusative pattern for these verbs. However, this extension did not take place with other intransitive verbs, which retained P-like coding in an ergative pattern. As a result, Bats developed a system of split intransitivity, with S arguments of some verbs coded like A and others like P argument of transitive verbs, as comparison between (7)b and (7)c shows.

,

(7) Bats (Nakh-Daghestanian; Harris 2010: 212)

    a.   *p'ay*   *b-eył-n-**as***   *ħo*

         kiss:NOM   CM-give-AOR-1SG.ERG   2SG.DAT

         'I gave you a kiss.' 'I kissed you.'

    b.   *(so)*   *vož-en-**sǒ***

         1SG.ABS   fell-AOR-1SG.ABS

         'I fell down, by accident.'

    c.   *(as)*   *daħ*   *y-apx-yail-n-**as***

         1SG.ERG   PV   CM-undress-AUX-AOR-1SG.ERG

         'I took my clothes off.'

  – *Phonological change*: this are cases in which a new alignment pattern emerges as phonological changes lead either to the development of specialized forms for particular argument roles or to the loss of existing specialized forms.

The first scenario is illustrated by Louisiana Creole (Haspelmath and the APiCS Consortium 2013). In origin, pronouns for A, S and P roles were undifferentiated in this language. However, A/S pronouns underwent phonological reduction, possibly on account of their higher frequency. As a consequence, the form of A/S pronouns became different from that of P pronouns, yielding an accusative pattern, as shown in Table 1.

Table 1. Pronominal declension in Louisiana Creole French

Person	**A, S**	**P**
1SG	*mo*	*mwa*

,

2SG	*to*	*twa*

The development of a new alignment pattern through the loss of existing forms for particular argument roles is illustrated by English (see Blake 2001: 176-178). In Old English, some inflectional classes of nouns retained a distinction between nominative and accusative case in the singular, the former used for A and S and the latter for P. As shown in Table 2, the distinction was realized differently for distinct noun classes. The distinction between nominative and accusative cases was disrupted by two phonological changes. On the one hand, unstressed vowels were reduced to schwa, so that NOM *talu* and ACC *tale* both became /ˈtalə/. On the other hand, word final -*n* was lost, so that ACC *naman* became identical to NOM *nama*. The result of the loss of case distinction was the emergence of a new neuter alignment pattern for nouns.

Table 2. Core case marking in Old English

Case	**'name'**	**'tale'**
NOM	*nama*	*talu*
ACC	*naman*	*tale*

    – *Loss*: This refers to cases where an existing marker for some argument role was lost in the language, but there is no clear evidence that this was due to phonological change.

The emergence of a new alignment pattern as a consequence of loss has been discussed for Tākestāni, a Tāti dialect. Like many modern Indo-Iranian languages, Tāti dialects feature a TAM-based alignment split. In the past tense, argument roles are arranged ergatively: A arguments receive dedicated ergative marking, while S and P arguments are unmarked and are indexed on the verb, as in (8)a-b. In addition, A arguments may, under certain conditions, also trigger the occurrence of A-indexing clitics.

,

(8) Eshtehārdi (Tāti dialect) (Indo-European; Rasekh-Mahand & Izadifar 2016: 141;

    Yarshater 1969: 230)

    a.   *Maryam-ā*   *Hasan*   *beza(d)*

        Maryam(F)-ERG   Hasan(M)   hit:PST.3SG.M

        'Maryam hit Hasan.'

    b.   *bābā-š*   *bemárda*

        father(M)-3SG.POSS.M   die:PST.3SG.M

        'His father has died.'

In Tākestāni, past transitive constructions have undergone several changes that have led to the emergence of a new alignment pattern. These changes are partly due to loss. In particular, ergative case marking for A and verbal indexes for P were lost, as shown by the comparison between (8)a and (9)b. As a result, past tense transitive constructions show a new tripartite alignment pattern (Rasekhahand & Izdifar 2016 for discussion): S is the only argument that triggers agreement with the verb, P is the only available host for A-clitics, and A triggers the use of A-clitics. The pattern is shown in (9)a-b.

(9) Tākestāni (Indo-European; Rasekh-Mahand & Izadifar 2016: 148)

        a.   *ā*   *ketāb*   *xeyli* *sext* *ve*

            that:M   book(M)   very   hard be:PST.3SG.M

            'That book was very hard.'

        b.   *a*   *jā*   *ketāb=em*   *bo*

            1SG   that:OBL   book=1SG   bring:PST

            'I brought that book.'

,

3.3. The effects of the process of change

For this domain, a number of fields are provided that describe the effects of the process of change leading to the development of the new alignment pattern:

- **Resulting construction**: this field is similar to the 'Source construction' field in that it features a description of the construction resulting from the process of change.

For example, the reinterpretation of the 'take' verb construction in Mandarin Chinese illustrated in (2)above yields a transitive construction with a P argument overtly marked by *bǎ*.

- **Alignment in the resulting construction**: this field reports the alignment pattern in the construction resulting from the process of change.

For example, if an intransitive resultative construction of the type 'X is VERBed by Y' is reinterpreted as a transitive one 'Y VERBed X', as is the case of Hanis Coos in (6), this will give rise to ergative alignment, because X becomes a P argument and is encoded in the same way as the S argument from which is derived, whereas Y becomes an A argument with dedicated marking, because it retains the marking used for the oblique NP from which it is derived.

- **Global alignment pattern following the change**: this field describes the global alignment pattern resulting from the combination of (i) the new alignment pattern of the construction resulting from the change and (ii) the alignment pattern of other constructions within the same grammatical domain.

,

For example, some processes of change may give rise to new perfective constructions with ergative alignment. If non-perfective constructions have other alignment patterns, however, the language will end up with a TAM-based alignment split, rather than a global ergative alignment pattern, as discussed for Hindi in (4).

Another example showing why it is useful to distinguish between alignment in the resulting construction and global alignment pattern following the change comes from Galela (Holton 2008). This language originally had nominative alignment in indexation. A new alignment pattern as a result of the reinterpretation of in transitive constructions with third person non-human indefinite A arguments and experiencer P arguments indexes, as in (10)a. In these constructions, the indexes for A arguments were progressively lost, and the construction was reinterpreted as an intransitive one, e.g. 'something angers her' > 'she is angry'. As a result, the original P index was reinterpreted as an S index, as shown in (10)b.

(10)　　　Galela (North Halmahera; Holton 2008:272)

　　　a.　　*i-mi-tosa*

　　　　　3SG.A.NONHUM-3F.SG.P-angry

　　　　　'Something makes her angry'

　　　b.　　*mi-pereki*

　　　　　3F.SG.P-old

　　　　　'She is old'

This change led to the emergence of a new ergative alignment pattern for the relevant intransitive verbs. This is shown by examples (11)a-b, where the same index *ni-* is used for S and P argument as opposed to a distinct A index *wo-*. As the S arguments of

,

other intransitive verbs retains A-like marking, however, at a global level the process results into split-intransitivity.

(11) Galela (North Halmahera; Holton 2008:261)

  a. *ni-kiolo*

   2SG.P-asleep

   'You are asleep'

  b. *wo-ni-doto*

   3M.SG.A-2SG.P-teach

   'He teaches you'

- **Constraints**: This field is optional and provides further specification about possible distributional restrictions for the alignment splits resulting from the process of change.

If there is a TAM or NP based split, for example, the field will specify the exact properties of the split (e.g. perfective constructions vs. non-perfective ones, pronouns vs. nouns, inanimate nouns vs. other NP types).

- **Grammatical domain**: This refers to the grammatical domain involved in the process of change, for example case marking, indexation, or word order.

Particular processes of change may involve multiple grammatical domains, e.g. both case marking and indexation. An example is Tākestāni in (9) where the emergence of a new alignment pattern is the result of the loss of both ergative case marking and verbal agreement.

,

- **Symmetry**: This refers to the morphosyntactic encoding of argument roles in the construction resulting from the change

Symmetric encoding means that all roles are encoded though the same strategy (e.g. overt case marking, overt indexation, whereas asymmetric encoding means that different roles are encoded through different strategies (zero vs. overt case marking, zero vs. overt marking in indexation).

An example of asymmetric marking is accusative alignment in Mandarin Chinese in (2): A and S roles are unmarked whereas P receives overt marking by means of *bǎ*. Symmetric marking can be found in case marking in Modern English, in which A, S and P are all equally unmarked (see discussion of the data in Table 2), and in the indexing pattern of Tobelo in (12), where all roles are variously marked by indexation on the verb.

(12)   Tobelo (North Halmahera; Holton 2003: 22)

    a.   *to-ni-gohara*

        1SG.NOM-2SG.ACC-hit

        'I hit you'

    b.   *to-tagi*

        1SG.NOM-go

        'I go.'

4. How to use DEmA

'

DEmA allows for fine-grained searches of the various components involved in the emergence of alignment patterns. Users can browse data in DEmA in two ways.

- **By language:** the full list of languages included in DEmA is provided in the **Languages** section, as shown in Figure 1. By clicking on each entry, users can visualize all the fields with the relevant information on the emergence of a new alignment pattern in that specific language.
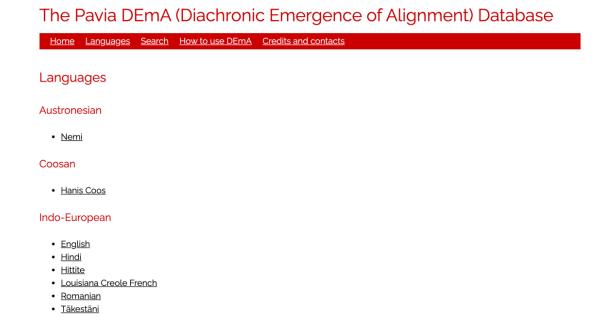

Figure 2. The DEmA Languages interface

- **By field:** our Search engine allows for queries on various fields, as shown in Figure 2. Users can simultaneously combine queries for multiple fields. Fields are divided into two categories based on the type of query parameter that they allow:

  - *Free text query*: users can freely enter their textual query in these fields (these are e.g. 'Language', 'Source Construction', 'Constraints').

,

An important free text query field is the **Keywords** field. Each Language is characterized by a number of keywords. These are intended as generic shortcuts for the various aspects of the historical process described in each entry and are meant to reflect the terminology most commonly used in the literature to refer to that specific process. Possible keywords include, for example, 'ergative', 'split ergativity', 'nominalization', 'passive', 'resultative construction'.

– *Selectable option query*: user can select one of the pre-existing options (e.g. 'Alignment in the resulting construction' features only a few options, such as Nominative-Accusative and Ergative-Absolutive).

The Pavia DEmA (Diachronic Emergence of Alignment) Database

Home   Languages   Search   How to use DEmA   Credits and contacts

Home › Search

Search language

Language:	
Glottocode:	
Genalogical classification:	
Original alignment pattern:	[Not selected]
Type of change:	[Not selected]
Alignment in the resulting construction:	[Not selected]
Global alignment pattern following the change:	[Not selected]
Constraints on the distribution of the resulting alignment:	
Grammatical domain:	[Not selected]
Symmetry:	[Not selected]
Keywords:	
	Invia

Figure 3. The DEmA Search interface

5. Conclusions

In this paper, we have offered an overview of the structure of the Pavia *Diachronic Emergence of Alignment* (DEmA) database. The database will be hosted by the University of Pavia, and will be available together with other linguistic resources developed at the Section of Theoretical and Applied Linguistics through the *The Pavia linguistic resources repository* (https://su-lab.unipv.it/tasf/).[2] Once released, the database will be fully searchable, allowing users to query the database for all parameters and combinations thereof. The database is also expandable, and we encourage scholars working on the diachrony of alignment to make their data available through DEmA.

At a more general level, the architecture of DEmA is unique in that it offers a theoretically well-grounded and explicit systematization of several parameters pertaining to language change (e.g. source constructions, type of change, type of data), so that these can be effectively implemented into a searchable format. In this respect, we hope that DEmA will also provide a suitable model for future typological resources dealing with the diachrony of other grammatical domains.

---

[2]   The DEmA interface has been built by Dr. Alessio Palmero Aprosio, whom we thank for his technical assistance.

,

Abbreviations

1 = first person; 2 = second person, 3 = third person, A = agent, ABS = absolutive, ACC = accusative, AOR = aorist, ASP = aspect, AUX = auxiliary, CM = (gender-)class marker, DAT = dative, ERG = ergative, F = feminine gender, INS = instrumental, M = masculine gender, MID = middle voice, NOM = nominative, NONHUM = non-human, OBL = oblique, P = patient, PASS = passive, PL = plural, POSS = possessive, PPP = perfect passive participle, PRF = perfect, PRS = present, PST = past, PTCL = particle, PV = preverb, SG = singular, TRANS = transitive

References

Blake, Barry J. 2001. *Case.* 2nd edition. Cambridge: Cambridge University Press.

Bubenik, Vit. 1998. *A historical syntax of late middle Indo-Aryan (Apabrahms´a)*. Amsterdam & Philadelphia: John Benjamins.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* 2nd edition. Oxford: Basil Blackwell.

Comrie, Bernard. 2013. Alignment of Case Marking of Full Noun Phrases. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/98> (19 May 2021)

Comrie, Bernard. 2013. Alignment of Case Marking of Pronouns. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/99> (19 May 2021)

Creissels, Denis. 2008. Direct and indirect explanations of typological regularities: the case of alignment variations. *Folia Linguistica* 42: 1–38.

Cristofaro, Sonia. 2012. Cognitive explanations, distributional evidence, and diachrony. *Studies in Language* 36: 645–670.

,

Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic

    perspective. In *Languages across Boundaries: Studies in the Memory of Anna Siewierska*,

    Dik Bakker & Martin Haspelmath (eds.), 69–93. Berlin & New York: Mouton de Gruyter.

Cristofaro, Sonia. 2014. Competing motivations and diachrony: what evidence for what

    motivations? In *Competing motivations in grammar and usage*, Brian MacWhinney,

    Andrej Malchukov, & Edith Moravcsik (eds.), 282–98. Oxford: Oxford University Press.

Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented

    explanations of typological universals. In *Explanation in typology*, Karsten Schmidtke-

    Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), 25–46. Berlin:

    Language Science Press.

Dahl, Eystein & Stroński, Krzysztof. 2016. Ergativity in Indo-Aryan and beyond. In *Indo-Aryan*

    Ergativity in Typological and Diachronic Perspective, Eystein Dahl & Krzysztof Stroński

    (eds.), 1-37. Amsterdam & Philadelphia: John Benjamins.

De Smet, Hendrik. 2009. Analysing reanalysis. *Lingua* 119(11): 1728-1755.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Gildea, Spike. 1998. *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford:

    Oxford University Press.

Grossman, Eitan & Stéphane Polis. 2018. Swimming against the typological tide or paddling along

    with language change? Dispreferred structures and diachronic biases in affix ordering.

    Journal of Historical Linguistics 8(3): 388–443.

Haig, Geoffrey. 2008. *Alignment Change in Iranian Languages; A Construction Grammar*

    Approach. Berlin & New York: Mouton de Gruyter.

Haig, Geoffrey. 2017. Deconstructing Iranian Ergativity. In *The Oxford Handbook of Ergativity*,

    Jessica Coon, Diane Massam & Lisa D. Travis (eds.), 465–500. Oxford: Oxford University

    Press.

,

Harris, Alice C. 2010. Origins of Differential Unaccusative/Unergative Case Marking: Implications

    for Innateness. <http://works.bepress.com/alice_harris/11/> (19 May 2021).

Harris, Alice C. & Campbell, Lyle. 1995. *Historical syntax in cross-linguistic perspective*.

    Cambridge: Cambridge University Press.

Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change?

    In *Explanation in typology*, Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria

    Michaelis & Ilja A. Seržant (eds.), 1–23. Berlin: Language Science Press.

Haspelmath, M. and the APiCS Consortium. 2013.. Alignment of case marking of personal

    pronouns. In *Atlas of Pidgin and Creole Language Structures Online*, Susanne Maria

    Michaelis, Philippe Maurer, Martin Haspelmath, and Magnus Huber (eds.). Oxford:

    Oxford University Press.

Heine, Bernd & Kuteva, Tania. 2002. *Word Lexicon of Grammaticalization*. Cambridge:

    Cambridge University Press.

Holton, Gary. 2003. *Tobelo*. Munich: Lincom Europa.

Holton, Gary. 2008. The rise and fall of semantic alignment in Northern Halmahera, Indonesia. In

    The typology of semantic alignment, Mark Donohue & Søren Wichmann (eds.), 252-276.

    Oxford: Oxford University Press.

König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.

Kulikov, Leonid. 2006. Case systems in a diachronic perspective. In *Case, valency and transitivity*,

    Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), 23–47. Amsterdam &

    Philadelphia: John Benjamins.

Kuteva, Tania, Heine, Bernd, Hong, Bo, Long, Haiping, Narrog, Heiko & Rhee, Seongha. 2019.

    World Lexicon of Grammaticalization. 2nd edition. Cambridge: Cambridge University

    Press.

Lehmann, Christian. 2015. *Thoughts on grammaticalization.* 3rd edition. Berlin: Language Science

    Press.

,

Li, Charles N. & Thompson, Sandra A. 1974. An explanation of word order change SVO→ SOV. *Foundations of Language* 12(2): 201-214.

Mithun, Marianne. 2005. Ergativity and language contact on the Oregon Coast: Alsea, Siuslaw, and Coosan. In *Proceedings of the Berkeley Linguistic Society*, 77–95.

Rasekh-Mahand, Mohammad & Izadifar, Rahele. 2016. Compensating ergative alignment loss in Takestani. In *Further topics in Iranian linguistics, Proceedings of the 5th international conference on Iranian linguistics, held in Babmberg on 24-26 August 2013*, Jila Ghomesh, Carina Jahani & Agnès Lenepveu-Hotz (eds.), 135-155. Leuven: Peeters.

Siewierska, Anna. 2013. Alignment of Verbal Person Marking. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/100> (19 May 2021)

Verbeke, Saartje. 2013. *Alignment and Ergativity in New Indo-Aryan Languages*. Berlin & New York: Mouton de Gruyter.

Verbeke, Saartje & De Cuypere, Ludovic. 2009. The rise of ergativity in Hindi: Assessing the role of grammaticalization. *Folia Linguistica Historica* 43: 367-389.

Yarshater, Ehsan. 1969. *A Grammar of Southern Tati Dialects*. The Hague & Paris: Mouton.

Zúñiga, Fernando. 2018. The diachrony of morphosyntactic alignment. *Language and Linguistic Compass* 12: e12300.

,