Typological explanations in synchrony and diachrony: On the origins of third person zeroes in bound person paradigms

Sonia Cristofaro

Sonia Cristofaro[*]

Typological explanations in synchrony and diachrony: On the origins of third person zeroes in bound person paradigms

Abstract: A well-known cross-linguistic pattern in bound person paradigms is for zero marking to be restricted to third person forms, particularly third person singular subjects. This has been explained in terms of general language preferences for third person zeroes, possibly determined by principles of iconicity or economy. This paper discusses several diachronic phenomena that result in third person zeroes cross-linguistically, and argues that these phenomena pose two challenges for this type of explanations. First, several phenomena do not appear to be related to general language preferences for third person zeroes. Second, different phenomena are a result of different causal factors, meaning that third person zeroes are not amenable to a unified explanation. This calls for a source-oriented approach to third person zeroes and recurrent cross-linguistic patterns in general, one where these patterns are accounted for in terms of the properties of different diachronic phenomena that shape the pattern over time, rather than synchronic properties of the pattern in itself.

1 Introduction

A well-known cross-linguistic pattern in bound person paradigms is for zero marking to be restricted to third, rather than first or second person forms, particularly third person singular subjects (Benveniste 1946; Croft 2003; Greenberg 1966; Koch 1995; Siewierska 2004 among others; see Siewierska 2010 for an investigation of similar patterns for object and possessor person forms).[1] This is illustrated in Tables 1 and 2 for Seri (isolate) and Lakhota (Siouan).

Table 1: Subject person prefixes in Seri (Marlett 1990: 514).

	ɪSG	PL
1	ʔ-, ʔp-	ʔa-
2	m-	ma-
3	*Ø*	*Ø*

---

[*]Corresponding author: Sonia Cristofaro, Faculté de Lettres, Sorbonne Université, Paris, France, E-mail: sonia.cristofaro@sorbonne-universite.fr

[1]This pattern has been interpreted in the literature either as a general tendency for languages to have zero marking for third person, or as an implicational pattern whereby, if zero marking is used at all in bound person paradigms, it will be used for third person (see, for example, Ariel 2000 or Bickel et al. 2015). The second interpretation is the standard one in language typology (Croft 2003; Greenberg 1966), and will be the one adopted here.

Table 2: Subject person prefixes in Lakhota (Buechel 1939: 35).

	SG	PL
1	*wa-*	*ų-*
2	*ya-*	*ya-*
3	*Ø*	*Ø*

This pattern has long been discussed in historical linguistics and language typology. A classical explanation, which goes back to Benveniste (1946), is in terms of a semantic contrast between third person, on the one hand, and first and second person, on the other. Third persons are regarded as semantically unmarked, or basic compared to first and second persons, in the sense that they are characterized by the absence of two meaning features associated with first and second persons, the notions of speaker and hearer (or, more generally, speech act participant). This contrast has been assumed to be mirrored by the absence of overt marking for third persons, as well as to drive a recurrent cross-linguistic process whereby third person singulars become zero marked as an overt marker is reinterpreted as part of the verb base (Haiman 1977; Koch 1995, among others). This process, originally described by Watkins (1962), leads to the creation of new first and second person forms through the addition of overt person markers to the original third person form.

Alternatively, the use of third person zeroes has been related to the higher discourse frequency of third person singulars as opposed to other person values. More frequent meanings are more easily identifiable, and hence less in need to be disambiguated through overt marking. This will lead either to the non-development of overt third person markers, or to the loss of these markers once they arise in a language. The idea of a connection between zero marking and the lower need to disambiguate particular grammatical meanings, originally proposed in Greenberg (1966), is, in fact, central to the notions of typological markedness and economy, and has been used to account for a number of other cross-linguistic patterns pertaining to the encoding of alienable and inalienable possession, NP-split ergativity, and differential object marking (Comrie 1989; Croft 2003; Haspelmath 2008).

A third explanation has been proposed in the context of hierarchical alignment in indexation, a pattern whereby first and second person are indexed on the verb in preference to third person irrespective of grammatical role. This encompasses both cases of third person zeroes proper, where third person participants are never indexed, and cases where third person participants are only indexed when no first or second person participant is present. Both of these cases have been explained in terms of the higher conceptual prominence of speech act participants (first and second person) as opposed to non-speech act participants (third person), which leads to the former being indexed in preference to the latter (Comrie 1980; DeLancey 1981, among others). While differing in their details, these explanations all assume that there is some general language preference for third person zeroes over first and second person ones. In some explanations, this preference originates from an iconicity principle whereby the organization of the paradigm should mirror the semantic unmarkedness or lower conceptual prominence of third person referents vis-a-vis first and second person ones. Alternatively, the preference for third person zeroes is

assumed to be a result of an economy principle whereby the organization of the paradigm should reflect the lower need to disambiguate third person verb forms due to their higher discourse frequency. In both cases, the explanation is result-oriented. Third person zeroes are assumed to be motivated by principles of optimization of grammatical structure pertaining to properties of the resulting verbal paradigms, in the sense that it is beneficial for speakers to use paradigms with third person zeroes.

In most cases, these explanations are based on the synchronic cross-linguistic distribution of zero vs. overt marking across different person values, not actual diachronic processes that give rise to this distribution from one language to another. Yet, in the functional-typological approach that originated from the work of Joseph Greenberg (particularly Greenberg 1966, Greenberg 1978), recurrent cross-linguistic patterns are regarded as a result of recurrent diachronic processes that give rise to the relevant constructions in individual languages. Speakers use these constructions because they are conventionalized in the language, and cross-linguistic patterns arise as the same constructions are recurrently created and retained across different generations of speakers in different languages. This is distinct from language universals in the generative sense of universal grammatical components that operate in a speaker's synchronic production of individual constructions. From this follows that explanations for individual patterns should refer to the diachronic processes that give rise to the pattern, rather than the resulting synchronic pattern in itself (Anderson 2005, Anderson 2016; Aristar 1991; Bybee 1988, Bybee 2006, Bybee 2008, Bybee 2009; Cristofaro 2011; Givón 1975, Givón 1979; Evans and Levinson 2009; Newmeyer 1998, Newmeyer 2005).

In this paper, a number of recurrent diachronic phenomena will be discussed which can give rise to third person zeroes in bound person paradigms cross-linguistically. These phenomena fall into two scenarios. First, languages may develop overt marking for first and second person, but not third. This may be due to the absence of particular source elements that can give rise to overt third person markers (particularly independent third person forms, but also other sources, such as affixes indexing indefinite third person referents), or the failure for these elements to actually give rise to overt third person markers (for example, independent third person forms may fail to evolve into bound ones, and affixes indexing indefinite third person referents may fail to be reinterpreted as general third person forms). In a second scenario, an overt marker for some third person referent is lost, so that the verb form ends up with zero marking for third person. This may be because the marker is incorporated into the verb base, as described earlier, or because it is reinterpreted as performing a different function (for example, affixes originally indexing indefinite third person referents can be reinterpreted as inverse markers).[2]

---

[2]Several other phenomena have been described in the literature that gave rise to third person zeroes in particular languages, and the reader is referred to Siewierska (2004) and Seržant (this volume) for data and discussion. These are not taken into account here, however, because they are a result of accidental, language-specific factors, rather than instantiating some recurrent cross-linguistic pattern. An often cited such case, is the development of third person singular and plural zeroes in the past tense in Polish. This was a result of person markers arising through the cliticization of the inflected forms of the auxiliary 'to be', whose third person forms had been lost (Andersen 1987; Siewierska 2004: 257–258). Another case in point is the emergence of third person zeroes as a result of language-specific phonological conditions leading to the loss of overt third person

Cross-linguistic evidence about these scenarios is unsystematic and quantitatively limited. Some of the relevant phenomena, for example the absence of independent third person forms, are well attested cross-linguistically, but their role in the development of third person zeroes may be difficult to demonstrate for particular languages. In other cases, there is evidence that particular phenomena were responsible for third person zeroes in particular languages, but these phenomena have only been documented for a limited number of languages. The available evidence, however, will be argued to pose two general challenges for result-oriented, synchronically based explanations of third person zeroes. First, several phenomena that ultimately result in third person zeroes do not appear to be motivated by general language preferences for third person zeroes in themselves. Instead, they are related to the properties of specific source constructions and processes involved in the development and loss of overt person forms cross-linguistically. Second, third person zeroes do not represent a unified pattern that should be accounted for in its own right. Instead, they emerge as a combined result of the cross-linguistic distribution of several distinct diachronic phenomena, which are not amenable to a unified explanation.

These facts call for a source-oriented approach to third person zeroes and recurrent cross-linguistic patterns in general, one where individual patterns are investigated and accounted for in terms of the properties and cross-linguistic distribution of multiple diachronic phenomena that shape the pattern over time, rather than particular synchronic properties of the pattern in itself.

2   Non-development of overt third person marking

A first possible scenario leading to third person zeroes in bound person paradigms is the development of overt marking for first and second person, but not third.

The major diachronic process known to give rise to overt marking in bound person paradigms is the cliticization and subsequent affixation of independent person forms.[3] Many languages have been shown to have independent forms for first and second person referents, but not for third person ones (Mithun 1986, 1991; Ariel 2000; Siewierska 2004, among several others). As is sometimes observed in the literature (Bybee 1985; Mithun 1991), this means that one possible reason for the non-development of overt third person marking in bound person paradigms is the absence of corresponding independent forms that could become affixed.

In practice, direct cross-linguistic evidence about this hypothesis is hard to come by, for several reasons. As overt markers in bound person paradigms are often of considerable antiquity, they may bear no obvious synchronic similarity to independent person forms, either because they underwent phonological changes, or because

---

markers, as described by Idiatov (2010) for Mande.

[3]In some cases, independent person forms are directly affixed to particular verb forms, while in other cases the relevant verb forms are originally periphrastic ones involving auxiliaries marked for person, so that the person markers are actually those of the auxiliary (Siewierska 2004: 257–260 and references therein; Seržant, this volume). As long as the person markers are derived from corresponding independent forms, the difference between these two cases is irrelevant to the arguments developed here.

they are derived from independent forms that are no longer available in the language (Siewierska 2004). Also, the fact that a language lacks independent third person forms does not rule out that such forms originally existed in the language, but failed to become affixed and were subsequently lost, or that they became affixed but the affix was subsequently lost.

For a number of languages, however, a direct diachronic relationship has been postulated between independent and bound person forms, so that third person zeroes in bound person paradigms can be assumed to be a result of the absence of corresponding independent third person forms. For example, as illustrated in Table 3, Hinton and Langdon (1976) show that, in several Yuman languages, first and second person subject prefixes are derived from corresponding independent forms in Proto-Yuman. The latter evolved into bound forms at a time when Proto-Yuman did not have independent third person forms, which gave rise to third person zeroes in bound person paradigms.

Table 3: Subject person forms in Yuman (Hinton and Langdon 1976: 119–20).

	Proto-Yuman		Several Yuman languages
	Independent	Bound	Bound
1	* *ʔa*	* *ʔ-*	*ʔ-*
2	* *ma*	* *m-*	*m-*
3	* **Ø**	* **Ø**	**Ø**

Similarly, in a number of Tibeto-Burman languages, bound person affixes are similar to independent person forms, and have been assumed to be derived from the independent forms, though there is some controversy about the details of the process (Jacques 2016; Kepping 1994, among others). Thus, when the language has third person zeroes in bound person paradigms and no independent third person forms, the former property can be assumed to be a result of the latter. This is illustrated in Tables 4 and 5 for Tangut and the Delugong dialect of Sgaw Karen.[4]

Table 4: Person forms in Tangut (Sino-Tibetan) (Kepping 1994; Jacques 2016).

	Independent	Bound
1SG	ηa^2	$-\eta a^2$
2SG	$\eta j a^2$	$-\eta j a^2$
3SG	**Ø**	**Ø**
1PL	$\eta a^2 m i^2$ (INCL)	$-n j i^2$
	$n j i^2 m i^2$ (EXCL)	
2PL	$n j i^2$	$-n j i^2$
3PL	**Ø**	**Ø**

---

[4]For Tangut, Jacques (2009) submits that the language originally had a third person object marker, which was incorporated in the stem of some verbs, giving rise to a specialized stem used when a first or second person singular acts upon a third person. Under this hypothesis, the fact that the language did not develop overt bound forms for third person objects cannot be explained in terms of the absence of corresponding independent forms, but this explanation still applies to the absence of other types of overt third person forms.

Table 5: Person forms in Sgaw Karen, Delugong dialect (Sino-Tibetan) (LaPolla 1994: 74).

	Independent	Bound
1SG	ja^{33}	$j\breve{a}^{33}$-
2SG	na^{33}	$n\breve{a}^{33}$-
3SG	Ø	Ø
1PL	$pu^{33}\ w\varepsilon^{55}\ \theta e^{31}$	$p\breve{u}^{33}\ ka^{31}$-
2PL	$\theta u^{55}\ w\varepsilon^{55}\ \theta e^{31}$	$\theta\breve{u}^{55}\ ka^{31}$-
3PL	Ø	Ø

An alternative scenario for the non-development of overt third person markers out of corresponding independent forms is proposed in Ariel (2000). While many languages have independent forms for all three persons, third person ones may fail to become affixed. This provides a viable explanation for the emergence of third person zeroes in languages that do have independent third person forms (though in principle this could also result from bound person paradigms developing before the development of independent third person forms, or these paradigms losing an existing overt third person marker).

Overt marking in bound person paradigms can also be a result of the reinterpretation of affixes originally used for other functions. In several North American languages, for example, portmanteau affixes indicating combinations of first and second person agents and patients evolved from cislocative elements meaning 'hither', 'here', or 'this way' (Mithun 1996). This is the origin of the Mohawk prefix *takw-*, illustrated in (1), which denotes combinations of second person agent and first person patient.

(1)   Mohawk (Iroquian)

        Takw-*atkáhtho-s*
        1/2-look-HAB
        'You (always) look at me.'

        (Mithun 1996: 433)

Affixes indexing indefinite third person referents ('somebody', 'someone') are also a source for various types of overt person markers. This too has been documented for several North American languages. In Athapascan languages, for example, the indefinite prefix *ch-* gave rise to both first and general third person markers, as illustrated in (2). In Caddo, the indefinite affix *-yi-* evolved into a general third person marker, as shown in (3).

(2)   The evolution of the indefinite prefix *ch-* in Athapascan languages:

        (a) Tanaina **ch'**uyu '**Someone** is walking.'

        (b) Tanaina **ch'**ghiʔan '**We** saw him.'

        (c) Hupa **ch'**itehsyay '**He** saw me.'

(Mithun 1986: 336–8)

(3)   Caddo (Caddoan)

kúyt-**yi**-*ʔa=dih-hah*
where-one-go-HAB
'Where is he going?' (Literally 'where is one going?')
(Chafe 1990: 65)

While this issue has not been investigated in the literature, the fact that overt marking in bound person paradigms can develop through the reinterpretation of pre-existing affixes points to other possible factors in the non-development of overt marking for third, as opposed to first or second person cross-linguistically. The affixes that give rise to overt third person markers (for example, affixes indexing indefinite third person referents) may be rarer than the sources that give rise to overt first and second person markers (including both affixes that can evolve into first or second person forms, for example cislocative affixes, and independent first and second person forms). Also, the processes whereby particular affixes evolve into third person markers may be rarer than the processes that give rise to first and second person markers, including both the reinterpretation of particular affixes as first or second person forms, and the affixation of independent first and second person forms.

3   Explaining non-development

The facts just described point to two general factors in the non-development of overt third person marking in bound person paradigms cross-linguistically: the absence of particular source elements (independent third person forms that could become affixed, affixes that could be reinterpreted as third person markers) and the failure for these elements to actually evolve into bound third person markers. Ultimately, then, explanations for the non-development of overt third person marking should be explanations for these phenomena. Challenging the result-oriented explanations of third person zeroes, however, most of these phenomena cannot be related to principles whereby it is beneficial for speakers to use paradigms with third person zeroes, either because of properties of third person verb forms or because of properties of third person referents.

3.1   Absence of possible sources

Cross-linguistically, the absence of particular elements that could give rise to third person bound forms will plausibly be related to properties of those elements, rather than the fact that this phenomenon can ultimately result into verbal paradigms with third person zeroes. For example, the absence of affixes indexing indefinite third person references will be related to properties of these affixes, rather than the fact

that absence of these affixes ultimately results in the language not developing general third person bound forms.

In the case of independent third person forms, as these forms encode third person referents, their absence could in principle be due to some property of third person referents that make it beneficial for speakers not to encode them through independent forms. In this case, the ultimate explanation for third person zeroes in bound person paradigms would be similar to some of the existing result-oriented explanations: the use of third person zeroes ultimately reflects properties of third person referents, even if these properties play a role in the absence of independent, rather than bound third person forms. This hypothesis, however, is not supported by the available evidence about the origins of independent person forms.

In general, independent person forms are known to develop through the reinterpretation of pre-existing elements originally used for other functions. First and second person forms have been shown to arise from expressions that are used to refer to first and second person participants without overtly mentioning them, including lexical items indicating social relations (e.g. 'lord', 'servant', 'master'), indefinite and generic pronouns ('someone', 'somebody', 'one'), expressions of multitude ('people'), spatial deictics ('here', 'there', 'this side'), and reflexives or intensifiers ('self', 'the same'). Independent third person forms often originate from demonstratives, and sometimes from lexical expressions such as 'man', 'person', 'people', 'others', or from reflexives and intensifiers. Some independent person forms can also originate from the reintepretation of others, for example, second plural and third person forms can be used in honorific expressions referring to second person singular participants, so that they become a standard way to address these participants (Heine and Song 2011; Helmbrecht 2004; Lehmann 1995; Siewierska 2004). These facts mean that cross-linguistic regularities in the presence or absence of different independent person forms will be a result of different source elements evolving or failing to evolve into these forms.[5]

In all of these cases, the processes whereby individual source elements evolve into person forms are plausibly ones of context-driven reinterpretation, of the type described for grammaticalization and semantic change in general (see, for example, Bybee et al. 1994; Heine 2003; Traugott and Dasher 2005). Expressions not originally used to encode person evolve into first, second, or third person forms because they are recurrently used to encode first, second, or third person referents, so that the person meaning can be inferred to be the central meaning of the expression as other meaning components are bleached over time. This provides a natural explanation for the non-development of third, as opposed to first and second person independent forms. Speakers will recurrently use the same expressions to refer to themselves and their hearers, usually ones that make it possible to avoid direct reference to either speaker or hearer. This will lead to the conventionalization of

---

[5]In principle, these regularities could also be due to processes recurrently leading to the loss of particular independent person forms as opposed to others. For example, it could be the case that, while third person forms are recurrently lost cross-linguistically, first and second person ones are usually retained. This scenario, however, is usually not considered in the literature on the diachonic evolution of independent person forms (see, for example Helmbrecht 2004), and there is no cross-linguistic evidence for it.

individual expressions as a standard way to refer to speaker or hearer. Over time, then, these expressions can be reinterpreted as specifically encoding the notions of speaker and hearer, also because these notions are usually more central in the discourse context than the original meaning of the expression. Third person referents, on the other hand, are usually encoded through different expressions depending on the referent (e.g. 'John', 'book', 'cat', 'that one'), and the meaning of these expressions is usually more relevant to the discourse referent than the notion of third person in itself. As also observed by Mithun (1991) and Helmbrecht (2004), this means that there is no specific form that will consistently be associated with third person referents, and the original meaning of the various forms will be less likely to undergo semantic bleaching. This makes it less likely for individual forms to be reinterpreted as third person forms.

In this scenario, there is no principle whereby it is beneficial for speakers not to use independent third person forms, as opposed to first and second person ones. Instead, the development or non-development of different person forms is an epiphenomenal result of the relative likelihood for different elements to be reinterpreted as person forms in particular contexts, as determined by how often those elements are used in those contexts, and how central the person meaning is in each case.

This makes it possible, in fact, to account for another phenomenon pertaining to third person marking which has been described for several North American languages (Mithun 1991). In these languages, as shown by the Chipewyan data in (4), bound person paradigms have overt markers for indefinite third person referents ('somebody', 'someone'), usually derived from corresponding independent forms, whereas general third person is zero marked.

(4)   Singular subject prefixes in Chipewyan (Athabascan: Li 1946: 411, 416;
        Mithun 1991):

- 1st person: *-s-*;
- 2nd person: *-nɛ-* (disjunct), *-(n)-* (conjunct);
- general 3rd person: **Ø**;
- indefinite 3rd person: ***ho-*** (originally an independent indefinite form)

A possible explanation for this pattern is that the relevant paradigms develop through the affixation of independent forms at a time when the language has independent forms for indefinite third person referents, but not for general third person. This is, in fact, a common situation cross-linguistically, in that, while languages often lack general third person independent forms, most languages have indefinite such forms (Haspelmath 1997). In this case, one should ultimately account for why languages evolve indefinite independent forms as opposed to general third person ones. This too can be explained in terms of the relative likelihood for these two types of forms to evolve from pre-existing elements. Indefinite referents will recurrently be encoded through the same expressions, namely ones that convey their indefinite character ('some body', 'one', 'I don't know who': see Haspelmath 1997: Chs. 6–8 for exhaustive discussion). In contrast to what happens with other types of third person referents, then, there will be some form that is recurrently associated with

indefinite referents, which can pave the way for the evolution of that form into a specialized indefinite form.

3.2 Failure for possible sources to evolve into bound third person forms

The fact that particular elements fail to give rise to bound third person forms may also not be related to properties of third person verb forms, or third person referents. In some cases, bound person forms develop through the reinterpretation of pre-existing elements. In grammaticalization studies and historical linguistics, reinterpretation processes are generally assumed to be a result of context-driven inferences. Therefore, whether or not particular elements are reinterpreted as person forms will depend on the contextual uses of these elements. For example, the evolution of cislocatives and affixes indexing indefinite third person referents into first, second, and third person markers has been related to the fact that these elements are used in order to refer to first, second, or third person participants without overtly mentioning them, so they can be reinterpreted as dedicated markers for these participants (Heath 1996; Mithun 1986, Mithun 1996; Siewierska 2004). Whether or not these elements are actually reinterpreted as markers for particular persons will then likely depend on whether and how often they are used to refer to the relevant participants, rather than general properties of these participants or properties of the resulting verb forms.

In other cases, bound person forms develop through the affixation of corresponding independent ones, so that third person zeroes may be a result of the failure for independent third person forms to become affixed. Ariel (2000) accounts for this phenomenon in terms of the relative cognitive accessibility of first, second, and third person referents, as determined by their inherent or discourse saliency. The latter depend, for example, on topicality, subjecthood, animacy, and number of previous mentions in the discourse context. Speakers will use attenuated encoding for more accessible referents, both because this involves less energy, and because this is a means to alert the addressee to search for an already available entity. As first and second person referents will be consistently more accessible than third person ones, they will be encoded by more attenuated devices, namely unstressed or cliticized person forms, which will eventually become affixed (for similar remarks, see also Mithun 1991). Third person referents, on the other hand, will be encoded by stressed forms, which are less likely to become affixed.[6]

---

[6]A possible counterexample to explanations of third person zeroes in terms of accessibility is provided by some languages of the type illustrated in (4), that is, bound person paradigms where general third person is zero marked, but indefinite third person referents are indexed by overt markers derived from corresponding independent forms. These languages, for example Sarcee (Cook 1984; Mithun 1991), have independent third person forms, meaning that independent indefinite forms became affixed at some point in the evolution of the language, whereas third person ones failed to do so. This cannot be explained in terms of accessibility, because indefinite referents will arguably have lower inherent and discourse saliency than the referents of third person forms, so they will consistently be of lower accessibility. This problem, however, can be addressed by assuming that bound person forms evolved at a time when the language had developed independent indefinite forms, but not yet third person ones.

Similarly to some result-oriented explanations, this explanation implies that third person zeroes are ultimately motivated by a specific property of third person referents, in this case their reduced cognitive accessibility. This property makes it beneficial for speakers to use stressed, rather than cliticized or affixed forms for third person referents, leading to third person zeroes in bound person paradigms. Result-oriented explanations, however, are generally meant to account for all instances of third person zeroes. Cognitive accessibility, on the other hand, provides an explanation for a particular phenomenon leading to third person zeroes, the non-affixation of independent third person forms. This explanation will not apply to cases where third person zeroes are a result of other phenomena, such as the absence of independent third person forms or the failure for particular affixes to be reinterpreted as third person markers.[7]

3.3 Non-development and result-oriented explanations

The various phenomena involved in the non-development of overt third person bound forms are still in need of cross-linguistic investigation. In many cases, we still miss a detailed understanding of particular phenomena, for example what factors lead to particular source elements not being reinterpreted as person forms (either independent or bound). Also, more data are needed about the cross-linguistic distribution of individual phenomena. For example, while the absence of independent third person forms has been shown to be cross-linguistically widespread, no systematic data are yet available about the relative cross-linguistic frequency of affixation processes involving different types of independent person forms, that is, first and second person forms as opposed to third person ones.[8] In the absence of diachronic evidence, when a language has independent third person forms, this makes it impossible to assess how likely it is that third person zeroes in bound person paradigms are a result of the non-affixation of the independent forms, rather than the fact that bound person paradigms developed at a time when these forms

---

[7]Ariel (2000: 235) suggests that accessibility could also explain the non-development of independent third person forms, as opposed to first and second person ones. Due to their higher accessibility, first and second person referents will be encoded by specialized person forms, whereas third person referents will be encoded by less attenuated devices, such as demonstratives or full NPs. As first and second person forms originate through the reinterpretation of pre-existing elements, however, an accessibility account implies that either the source elements are more attenuated devices than the elements used to denote third person referents, or the reinterpretation process is driven by the higher accessibility of first and second person referents vis-a-vis third person ones. Yet, third person referents are often denoted by the same types of elements that give rise to first and second person forms (full NPs), and sometimes they are encoded through more, rather than less attenuated devices (for example, demonstratives as opposed to full NPs). As for reinterpretation processes, these usually involve mechanisms of context-driven inference, which are not generally assumed to be related to accessibility.

[8]Mithun (1991) points to some possible ways to use synchronic data to make inferences about these processes. In particular, she discusses several North American languages where third person affixes do not occur in the same position as first and second person ones, and argues that this is evidence that they developed at a later stage. This, however, is not necessarily evidence that independent first and second person forms have a greater tendency to become affixed than third person ones. Instead, such cases could be a result of the fact that independent third person forms develop after the affixation of first and second person ones, so they also become affixed at a later stage.

did not yet exist in the language. Also, processes whereby particular affixes evolve into person markers have only been documented for a limited number of languages, so the cross-linguistic frequency of these processes and the relevant affixes are not known either.

In spite of these limitations, the available data about the non-development of overt third person marking pose two challenges for result-oriented explanations of third person zeroes. First, in many cases non-development will be a result of particular elements failing to give rise to overt third person forms, either in the sense that these elements fail to be reinterpreted as independent third person forms that could become affixed, or in the sense that these elements are already affixed but fail to be reinterpreted as third person markers. These phenomena are plausibly independent of whether the resulting grammatical patterns (absence of independent third person forms, ultimately leading to third person zeroes, or third person zeroes proper) comply with particular principles of optimization of grammatical structure. Instead, the failure for some element to evolve into a third person form will be a result of the fact that the contextual uses of that element disfavor the relevant reinterpretation processes. Second, different phenomena resulting in the non-development of overt third person marking plausibly reflect different factors. For example, different processes of reinterpretation may fail to take place for different reasons pertaining to the properties and the contexts of use of different elements that could be involved in the process. The non-affixation of independent third person forms may be motivated by yet other factors, such as the lower cognitive accessibility of third person referents. In these various cases, then, third person zeroes will not be amenable to a unified explanation.

4   Loss of overt marking for third person referents

Another scenario leading to third person zeroes in bound person paradigms is the loss of an existing overt marker indexing a third person referent. While there is cross-linguistic evidence that bound person markers can be lost as a result of phonological attrition or decay (Siewierska 2004: 277–281), there is no evidence that these processes recurrently target third, rather than first and second person forms. Two cross-linguistic processes have, however, been described whereby an overt marker indexing a third person referent is lost as a result of reinterpretation, so that the relevant verb forms end up with zero marking for third person. The first process was mentioned in Section 1. In several languages, an overt third person marker was reinterpreted as part of the verb base, so that the verb form became zero marked for person. A well-known case in point is provided by Old and Modern Provençal, illustrated in Table 6. In Old Provençal, third person singular was encoded by the suffix *-et*. In Modern Provençal, this suffix was reinterpreted as part of the base, so that third person singular forms became zero marked for person and the other person forms were rebuilt by adding person suffixes to these forms.

Table 6: Provençal preterites (Bybee 1988: 55)

	Old Provençal	Modern Provençal, Charente dialect
1SG	*améi*	*cantí*
2SG	*amést*	**cantét***ei*
3SG	*am**ét***	**cantét**
1PL	*amém*	**cantét***em*
2PL	*amétz*	**cantét***ei*
3PL	*améren*	**cantét***en*

A second process involves forms originally encoding indefinite third person referents, of the type illustrated in examples (2)–(3) above. In some languages, when used to index agents in transitive configurations with an overtly marked first or second person patient, these forms were reinterpreted as markers indicating the direction of the action, that is, the fact that a third person referent acts upon a first or second person one (inverse markers). As a result, the verb form evolved into one with a zero marked third person agent. This is illustrated in (5) and (6) below.

(5)   Chukchi (Chukotkan)

ne-*lʔu*-gət
INV-see-2SG
'He saw you' (originally 'somebody saw you')
(Fortescue 1997: 382)

(6)   Japhug Rgyalrong (Sino-Tibetan)

pɯ́-**wɣ**-*mto-a*
AOR-INV-see-1SG
'He/she/it saw me.' (originally, in one possible analysis, 'somebody saw me')
(Jacques 2010: 129, 136)[9]

These processes raise the same type of issues as the non-development of overt third person marking. Third person zeroes are usually explained in terms of the properties of the resulting verbal paradigms, in that these paradigms comply with some assumed principle of optimization of grammatical structure whereby there should be general language preferences for third person zeroes, rather than first or second person ones. Yet, if third person zeroes result from the incorporation of an overt third person marker into the base, or the reinterpretation of an indefinite marker as an inverse one, explanations for this pattern should ultimately be explanations for these processes. Neither of these processes, however, is obviously related to properties of the resulting verbal paradigms.

---

[9]Jacques (2010: 136) argues that another possible origin of the Japhug Rgyalrong inverse marker is a cislocative element ultimately derived from a verb 'to come'.

As outlined in Section 1, for example, the incorporation of overt third person markers into the verb base has been related to an iconicity principle whereby verbal paradigms will be restructured so as to mirror the semantic unmarkedness of third person referents as opposed to first and second person ones. Bybee (1985, 2001, 2007), however, argues that this process is rather a manifestation of chunking, a general mechanism whereby frequently occurring sequences of linguistic items are automatized and processed as a single chunk, leading to the loss of the identity of the component units. Due to their high discourse frequency, third person verb forms will undergo chunking, leading to the third person marker being reinterpreted as part of the base.

In this analysis, third person zeroes ultimately result from a specific property of third person verb forms as opposed to first and second person ones, namely their higher discourse frequency. This means, however, that the loss of an overt person morpheme is a mechanic result of the repetition and automatization of more frequent items, not that there is a specific motivation for speakers to use zero, rather than overt marking for third person.

As for the evolution of indefinite markers into inverse markers, this can be explained in terms of the reduced communicative load of these markers. Due to their being semantically generic, indefinite participants ('somebody', 'someone') are usually unimportant in the discourse context. This may lead to the bleaching of the indefinite meaning over time, at which point the indefinite marker can be reinterpreted as something else, for example an inverse marker. This is basically a metonymization process whereby some formal component of a complex expression (the indefinite marker) becomes associated with a meaning originally associated with the expression as a whole (the direction of the action). This is independent of the fact that the result of the process, the emergence of third person zeroes, may comply with particular principles of optimization of grammatical structure.

These facts have similar implications as the phenomena leading to the non-development of overt third person marking. The emergence of third person zeroes as a result of the reinterpretation of some existing marker is no unified phenomenon related to properties of third person zeroes in themselves. Instead, this phenomenon will be a result of the properties of different source constructions, such as the fact that particular verb forms are more likely to undergo chunking, or the fact that indefinite markers in transitive constructions lend themselves to be reinterpreted as inverse markers. Third person zeroes will then reflect different factors depending on the source constructions involved, and these factors will not be the same as those involved in the emergence of third person zeroes through the non-development of overt third person marking.

5   Concluding remarks

The facts discussed in this paper have long been known from grammaticalization studies, historical linguistics, and specialized research on the relevant languages and language families. These facts, however, have a number of consequences for explanations of third person zeroes and recurrent cross-linguistic patterns in general,

which have generally failed to be considered both in these research domains and in language typology.

Current explanations of third person zeroes generally assume that this phenomenon reflects principles of optimization of grammatical structure pertaining to properties of third person verb forms or third person referents. These principles lead to general language preferences for third person zeroes and motivate different instances of third person zeroes cross-linguistically.

Yet, evidence about the origins of third person zeroes cross-linguistically points to a more complex scenario, one in which third person zeroes result from several recurrent and independently motivated, phenomena. Some of these, namely the non-affixation of independent third person forms, may be a result of principles that lead speakers to select specific forms for third person referents (non attenuated encoding) due to particular properties of these referents. In other cases, however, third person zeroes are ultimately a result of phenomena unrelated to the synchronic properties of third person referents or third person verb forms in themselves, such as the failure for particular elements to be reinterpreted as third person forms (either independent or bound), or processes of context-driven reinterpretation whereby markers which originally encode some type of third person referents lose this function.

These facts have general consequences for both theoretical assumptions and empirical research on third person zeroes. The theoretical debate on third person zeroes in typology and historical linguistics has concentrated on identifying factors that should lead to general preferences for this pattern, for example, the lower need to disambiguate third person verb forms, or particular properties of third person referents, such as semantic unmarkedness or lower conceptual prominence. At the empirical level, this is reflected by research aiming to test the statistical evidence for such preferences by obtaining more accurate data on the actual distribution of third person zeroes cross-linguistically (Bickel et al. 2015).

If third person zeroes result from a variety of independently motivated phenomena, however, this pattern should be explained in terms of the properties of these phenomena, rather than principles pertaining to third person zeroes in themselves. At the theoretical level, this means shifting the focus from explaining assumed language preferences for third person zeroes to explaining individual phenomena that shape this pattern from one language to another. These phenomena include the absence of independent third person forms, the failure for these forms to become affixed, the absence of particular affixes that could evolve into third person markers, the failure for overt third person markers to actually develop from these affixes, and different processes of reinterpretation whereby affixes originally indexing some third person referent lose this function. At the empirical level, the focus should shift from the cross-linguistic distribution of third person zeroes in themselves to the cross-linguistic distribution of these various phenomena, so so as to disentangle their respective contributions to the overall distribution of third person zeroes. For example, the frequency should be investigated of different source elements and developmental processes which can give rise to independent or bound person forms, or different processes of reinterpretation involving markers indexing third person referents.

This approach is in line with studies such as Siewierska (2010) or Bickel et al.

(2015), where third person zeroes are assumed to be a result of specific diachronic phenomena that should be investigated in their own right. The former study, for example, proposes some methods to disentangle third person zeroes resulting from the non-development of overt third person markers from ones resulting from the loss of some existing marker. In the latter study, third person zeroes are regarded as a result of the non-affixation of independent third person forms, or the reinterpretation or loss of existing third person affixes. These studies, however, do not address the issue of the possible motivations for individual diachronic phenomena, or the possible implications of these phenomena for classical explanations of third person zeroes.

From a logical point of view, the diachronic evidence about the origins of third person zeroes does not rule out that this pattern could ultimately also be motivated by general language preferences for the pattern in itself, irrespective of what phenomena give rise to the pattern in individual languages. This scenario would be similar to natural selection in biological evolution: the distribution of some genetic trait in a population depends on some preference for that trait (as determined by the trait's adaptiveness to the environment), irrespective of the processes that gave rise to the trait in the first place. This scenario is explicitly advocated by Haspelmath (2019) as an explanation for recurrent cross-linguistic patterns, and a detailed comparison between this scenario and the view defended here can be found, for example, in Schmidtke-Bode and Grossmann (2019).

A number of problems with this scenario are, however, detailed in Cristofaro (2017, 2019). In biological evolution, there is direct evidence for natural selection, in that particular genetic traits make it demonstrably more likely for the organisms carrying them to survive and pass them on to their descendants. The direct equivalent of natural selection in grammatical evolution would be if there were general preferences for some grammatical trait that lead to the retention of the trait from one language to another irrespective of its origins, for example general preferences for third person zeroes that lead to the retention of this pattern irrespective of how it developed in the first place. There is, however, no general evidence that the retention of grammatical traits is related to language preferences for that trait that make it more likely to survive cross-linguistically. An alternative would be for the assumed preferences to play a role in the development, rather than the retention of particular grammatical traits. For third person zeroes, this would mean that these preferences should favor different developmental processes leading to this pattern (such as the reinterpretation of different affixes indexing third person referents), and disfavor developmental processes or source constructons that could give rise to overt third person forms. In grammaticalization studies and studies of language change in general, however, the development of grammatical patterns is usually assumed to be related to the properties of particular source constructions and their contexts of use, not properties of the resulting patterns in themselves (see, for example, Bybee et al. 1994: 298–300 and Slobin 2002: 381 for an explicit rejection of this view in regard to grammaticalization). The cross-linguistic distribution of source constructions is also plausibly independent of the properties of the grammatical patterns that could arise from those constructions, as argued in Section 3.1.

In general, data about the diachronic phenomena that shape particular syn-

chronic patterns cross-linguistically are bound to be scantier and less systematic than data about the synchronic patterns in themselves. The facts discussed in this paper, however, call for a novel, source-oriented approach to the relationship between synchronic cross-linguistic patterns, as described in language typology, and the evolution of grammatical structure (Cristofaro 2013, 2014, 2017, 2019). This approach is similar to Evolutionary Phonology and other diachronically oriented approaches to cross-linguistic phonological patterns (see, for example, Blevins 2004; Ohala 1993, 2003). While typologists generally assume that synchronic cross-linguistic patterns are a result of diachronic phenomena, explanations for individual patterns are usually based on the synchronic properties of the pattern. This implies that the diachronic phenomena that shape the pattern over time will ultimately be motivated by its synchronic properties, in the sense that the development or retention of the pattern cross-linguistically will be motivated by principles pertaining to these properties. The available diachronic evidence shows, however, that synchronic cross-linguistic patterns can be a result of multiple diachronic phenomena, often unrelated to the synchronic properties of the pattern. Explanations of individual patterns, then, should be informed by a qualitative understanding of these phenomena and a quantitative understanding of their respective contribution to the pattern, rather than being based on the synchronic properties of the pattern in itself.

List of abbrevations

1 first person

2 second person

3 third person

AOR aorist

EXCL exclusive

HAB habitual

INCL inclusive

INV inverse

PL plural

SG singular

References

Andersen, Henning. 1987. From auxiliary to desinence. In Martin Harris and Paolo Ramat (Eds.), *Historical development of auxiliaries*, 21–52. Berlin & New York: Mouton de Gruyter.

Anderson, Stephen R. 2005. Morphological universals and diachrony. In Geert Booij & Jap van Marle (eds.), *Yearbook of morphology 2004*, 1–17. Dordrecht: Springer.

Anderson, Stephen R. 2016. Synchronic vs. diachronic explanations and the nature of the Language Faculty. *Annual Review of Linguistics* 2. 11–31.

Ariel, Mira. 2000. The development of person agreement markers: from pronouns to higher accessibility markers. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 197–260. Stanford: Center for the Study of Language and Information.

Aristar, Anthony R. 1991. On diachronic sources and synchronic patterns: an investigation into the origin of linguistic universals. *Language* 67. 1–33.

Benveniste, Émile. 1946. Structure des relations de personne dans le verbe. *Bulletin de la Societé de Linguistique de Paris* 126. 1–12.

Bickel, Balthasar, Alena Witzlack-Makarevich, Taras Zakharko & Giorgio Iemmolo. 2015. Exploring diachronic universals of agreement: alignment patterns and zero marking across person categories. In Jürg Fleischer, Elisabeth Rieken & P. Widmer (eds.), *Agreement from a diachronic perspective*, 29–51. Berlin and New York: Mouton de Gruyter.

Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Buechel, Eugene. 1939. *A grammar of Lakota, the language of the Teton Sioux Indians*. St. Louis: J. S. Swift.

Bybee, Joan. 1985. *Morphology: a study of the relation between meaning and form*. Amsterdam and Philadelphia: John Benjamins.

Bybee, Joan. 1988. The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining language universals*, 350–379. Oxford: Basil Blackwell.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2006. Language change and universals. In Ricardo Mairal & Juana Gil (eds.), *Linguistic universals*, 179–194. Cambridge: Cambridge University Press.

Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, Joan. 2008. Formal universals as emergent phenomena: the origins of structure preservation. In Jeff Good (ed.), *Linguistic universals and language Change*, 108–121. Oxford: Oxford University Press.

Bybee, Joan. 2009. Language universals and usage-based theory. In Morten H. Christiansen, Chris Collins & Shimon Edelman (eds.), *Language universals*, 17–40. Oxford: Oxford University Press.

Bybee, Joan, Revere Perkins, and Williamtonno Pagliuca. 1994. *The evolution of grammar*. Chicago and London: The University of Chicago Press.

Chafe, Wallace. 1990. Use of the defocusing pronominal prefixes in Caddo. *Anthropological Linguistics* 32. 57–68.

Comrie, Bernard. 1980. Inverse verb forms in Siberia: evidence from Chukchee, Koryak, and Kamchadal. *Folia Linguistica* 1. 61–74.

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

Cook, Eung-Do. 1984. *A Sarcee grammar*. Vancouver: University of British Columbia Press.

Cristofaro, Sonia. 2011. Language universals and linguistic knowledge. In Jae-Jung Song (ed.), *Handbook of linguistic typology*, 227–249. Oxford: Oxford University Press.

Cristofaro, Sonia. 2013. The referential hierarchy: reviewing the evidence in diachronic perspective. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: studies in the memory of Anna Siewierska*, 69–93. Berlin and New York: Mouton de Gruyter.

Cristofaro, Sonia. 2014. Competing motivations and diachrony: what evidence for what motivations? In Brian MacWhinney, Andrej Malchukov & Edith Moravcsik (eds.), *Competing motivations in grammar and usage*, 282–298. Oxford: Oxford University Press.

Cristofaro, Sonia. 2017. Implicational universals and dependencies between grammatical phenomena. In Nicholas Enfield (ed.), *Dependencies in Language: on the causal ontology of linguistic systems*, 9–24. Berlin: Language Science Press.

Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Suzanne M. Michaelis, and Ilja A. Seržant (eds.), *Explanation in typology: diachronic sources, functional motivations and the nature of the evidence*, 25–46. Berlin: Language Science Press.

Croft, William. 2003. *Typology and universals. 2nd edition*. Cambridge: Cambridge University Press.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.

Evans, Nicholas and Stephen C. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and brain sciences* 32. 429–448.

Fortescue, Michael. 1997. Eskimo influence on the formation of the Chukotkan ergative clause. *Studies in Language* 21. 369–409.

Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. In Charles N. Li (ed.), *Word order and word order change*, 47–112. Austin: University of Texas Press.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Greenberg, Joseph H. 1966. *Language universals, with particular reference to feature hierarchies*. The Hague: Mouton.

Greenberg, Joseph H. 1978. Diachrony, synchrony and language universals. In Joseph H. Greenberg, Charles H. Ferguson & Edith A. Moravcsick (eds.), *Universals of human language*. Volume 1. Method and theory, 62–91. Stanford: Stanford University Press.

Haiman, John. 1977. Reinterpretation. *Language 53*, 312–28.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2008. Creating economical morphosyntactic patterns in language change. In Jeff Good (Ed.), *Linguistic universals and language change*, 185–214. Oxford: Oxford University Press.

Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Karsten Schmidtke-Bode, Natalia Levshina, Susanne M. Michaelis, and Ilja A. Seržant (eds.), *Explanation in typology: diachronic sources, functional motivations and the nature of the evidence*, 1–23. Berlin: Language Science Press.

Heath, Jeffrey. 1996. Pragmatic skewing in 1<->2 pronominal combinations in native American languages. *International Journal of American Linguistics* 64, 83–104.

Heine, Bernd. 2003. Grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 576–601. Oxford: Blackwell.

Heine, Bernd & Kyung-An Song. 2011. On the grammaticalization of personal pronouns. *Journal of Linguistics* 47. 1–44.

Helmbrecht, Johannes. 2004. *Personal pronouns – form, function, grammaticalization*. Habilitationsschrifr, Universitaet Erfurt.

Idiatov, Dmitry. 2010. Person-number agreement on clause linking markers in Mande. *Studies in Language* 34. 832–868.

Jacques, Guillaume. 2009. The origin of vowel alternations in the Tangut verb. *Language and Linguistics* 10. 17–28.

Jacques, Guillaume. 2010. The inverse in Japhug Rgyalrong. *Language and Linguistics* 11. 127–157.

Jacques, Guillaume. 2016. Tangut, Gyalrongic, Kiranti and the nature of person indexation in Sino-Tibetan/Trans-Himalayan. *Linguistics Vanguard* 1, 20150033. doi: https://doi.org/10.1515/lingvan–2015–0033.

Kepping, Ksenia B. 1994. The conjugation of the Tangut verb. *Bulletin of the School of Oriental and African Studies* 57. 339–346.

Koch, Harold. 1995. The creation of morphological zeroes. In G. Booij & J. van Marle (eds.), *Yearbook of morphology 1994*, 31–71. Dordrecht: Kluwer.

Hinton, Leanne & Margaret Langdon. 1976. Object-subject pronominal prefixes in La Huerta Diegueño. In Margaret Langdon and Shirley Silver (eds.), *Hokan studies: papers from the first conference on Hokan languages*, 113–128. The Hague: Mouton.

LaPolla, Randy J. 1994. Parallel grammaticalizations in Tibeto-Burman languages: evidence of Sapir's 'drift'. *Linguistics of the Tibeto-Burman Area*, 61–80.

Lehmann, Christian. 1995. *Thoughts on grammaticalization*. Münich: Lincom Europa. First published 1982, as *Arbeiten des Kölner Universalien Projekt*. 48. Köln: Institut für Sprachwissenschaft.

Li, Fang-Kuei. 1946. Chipewyan. In Charles Osgood & Hans Hoijer (eds.), *Linguistic structures of native America*, 398–423. New York: The Viking Fund Publications in Anthropology.

Marlett, Stephen A. 1990. Person and number inflection in seri. *International Journal of American Linguistics* 56, 503–41.

Mithun, Marianne. 1986. Reconstructing the unidentified. In Henk Aertsen & Robert J. Jeffers (eds.), *Papers from the 9th international conference on historical linguistics*, 329–347. Amsterdam and Philadelphia: John Benjamins.

Mithun, Marianne. 1991. The develoment of bound pronominal paradigms. In Winfred P. Lehmann & Helen-Jo Jakusz Hewitt (eds.), *Language typology 1988: typological models in reconstruction*, 85–104. Amsterdam: John Benjamins.

Mithun, Marianne. 1996. New directions in referentiality. In Barbara Fox (ed.), *Studies in anaphora*, 413–35. Amsterdam and Philadelphia: John Benjamins.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, Mass.: The MIT Press.

Newmeyer, Frederick J. 2005. *Possible and probable languages*. Oxford: Oxford University Press.

Ohala, John J. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: problems and perspectives*, 237–78. London: Longman.

Ohala, John J. 2003. Phonetics and historical phonology. In Richard D. Janda & Brian D. Joseph (eds.), *The handbook of historical linguistics*, 669–686. Oxford: Blackwell.

Schmidtke-Bode, Karsten & Eitan Grossmann. 2019. Diachronic sources, functional motivations and the nature of the evidence: a synthesis. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne M. Michaelis, and Ilja A. Seržant (eds.), *Explanation in typology: diachronic sources, functional motivations and the nature of the evidence*, 223–241. Berlin: Language Science Press.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Siewierska, Anna. 2010. Person asymmetries in zero expression and grammatical function. In Franck Floricic (ed.), *Essais de typologie de linguistique générale. Mélanges offerts à Denis Creissels*, 471–485. Paris: ENS.

Slobin, Dan I. 2002. Language evolution, acquisition and diachrony: probing the parallels. In T. Givón & Bertram F. Malle (eds.), *The evolution of language out of pre-language*, 375–92. Amsterdam and Philadelphia: John Benjamins.

Traugott, Elisabeth C. and Richard B. Dasher. 2005. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Watkins, Calvert. 1962. *Indo-European origins of the Celtic verb. Volume I. The sigmatic aorist*. Dublin: Institute for Advanced Studies.