

From Proto-Indo-European to Old Latin *ipse* and *iste*: Two Sides of the Same Coin?^{*}

Verónica Orqueda

Pontificia Universidad Católica de Chile

vorqueda@uc.cl

ORCID iD: <https://orcid.org/0000-0003-0765-2928>

Demian Inostroza

Pontificia Universidad Católica de Chile

drinostroza@uc.cl

ORCID iD: <https://orcid.org/0000-0002-3651-3751>

Del proto-indoeuropeo a *ipse* e *iste* en latín arcaico: ¿Dos caras de la misma moneda?

This paper offers a comparative analysis of the pronouns *ipse* and *iste* in Old Latin. We aim to explain their differences and similarities as a consequence of partial common origin in Proto-Indo-European (PIE). We examine the distribution of these two pronouns according to various grammatical, semantic, and pragmatic features. The unequal distribution of *ipse* and *iste* suggests that these pronouns derive from items that ranked differently on a referential hierarchy. In consideration of cross-linguistic evidence, these results can be interpreted as remnants of an archaic system of marking proximate / obviative arguments.

Key words: Proto-Indo-European; Old Latin; referential hierarchy; demonstratives.

Este trabajo presenta un análisis comparativo de *ipse* e *iste* en latín arcaico. Nuestro objetivo es explicar las diferencias y similitudes entre *ipse* e *iste* a partir de un origen común parcial en PIE. Para esto, analizamos la distribución de estos pronombres según diversos rasgos gramaticales, semánticos y pragmáticos. La distribución desigual de *ipse* e *iste* sugiere que estos derivan de elementos ubicados en diferentes puntos de una jerarquía referencial. Considerando evidencia interlingüística, estos datos pueden interpretarse como la preservación residual de un sistema que marcaba argumentos proximativos y obviativos en PIE.

Palabras clave: proto-indoeuropeo; latín arcaico; jerarquía referencial; demostrativos.

^{*} This paper is part of the ANID-FONDECYT projects N° 11170045 and N° 1210321. We would like to thank Berta González Saavedra, Eugenio Luján Martínez, Roland Pooth, Cormac Anderson and two anonymous reviewers for valuable comments on earlier versions of this paper. All errors are our own.

Citation / Cómo citar este artículo: Orqueda, Verónica and Inostroza, Demian (2021): «From Proto-Indo-European to Old Latin *ipse* and *iste*: Two Sides of the Same Coin?», *Emerita* 89 (2), pp. 251-278.

I. INTRODUCTION, AIMS, AND METHODOLOGY

Iste and *ipse* belong to the complex pronominal and deictic system of Latin. The traditional grammars usually describe *iste* as a (spatial) demonstrative pronoun, probably with addressee (or near-hearer) deixis (Pinkster 2015, p. 972; Bassols 1992, p. 127), with a meaning akin to «that of yours» (de Vaan 2008, p. 310; Sihler 1995, p. 394). From a synchronic approach, Pieroni (2010, p. 478) claims that *iste* preserves its deictic nature even «when used to refer to entities which are found in the text», which is possible because «deictic and anaphoric strategies are not mutually exclusive». In turn, *ipse* is traditionally described as primarily an identificational, focal, or emphatic item («himself and not someone else», as in Bassols 1992, p. 130), although there is no doubt that *ipse* presents serious challenges to a uniform characterization. Thus, Álvarez Huerta (2009) and Pinkster (2015) allude to the need to differentiate the diverse uses of *ipse*¹, while Pieroni (2010) advocates for a general predicative function.

In the same way as other Latin pronouns, *ipse* and *iste* can be used as adnominals and as phrase heads. Both inflect for the three genders —masculine (m.), feminine (f.), and neuter (nt.) — and all grammatical cases except for vocative. Although most grammars do not deal with these pronouns together, a partially common etymology is usually assumed: for instance, de Vaan (2008) derives *ipse* from Proto-Indo-European (henceforth, PIE) **so-pe-so* and *iste* from **es-to*.

Most authors consider the PIE demonstrative elements **so* and **to* both part of a suppletive inflected deictic pronoun. This root, together with the anaphoric **h₁e/*i(s)-*, shaped the demonstrative system, according to Fortson (2004, p. 129), and Beekes & de Vaan (2011, p. 226)². In this view, suppletion is found only in the nominative: m. **so*, nt. **tod*, and the later f. **seh₂*, while the remainder of the inflected forms utilize the base in **t-*. If *ipse* and *iste* derive from this same root, one may wonder why and how they developed in such

¹ Pinkster (2015, pp. 1151-1152), for instance, differentiates between «pregnant» and «discretive» uses of *ipse*.

² Cf. Mendoza 1998, who posits that the stems **so-/to-* and **ei-/*i-* were anaphors in complementary distribution.

different ways. As we will show, *ipse* mostly takes animate predicates, while *iste* is not sensitive to animacy; *ipse* is frequently used in the nominative case, but *iste* shows no preference regarding case marking; *ipse* takes 1st- and 2nd-person referents, while *iste* only takes the 3rd person. The hypothesis underlying this paper is that a common etymology is responsible for these semantic and grammatical differences in Old Latin. Therefore, we provide a contrastive analysis to describe these forms together, and this analysis leads us to propose a different reconstruction of their PIE origins, as well as a different diachronic development from PIE towards the Latin pair *ipse* / *iste*: the pronouns **so* and **to* must not have been part of a suppletive paradigm but rather were related pronouns that ranked differently within a referential hierarchy.

As part of our methodology, we gathered all tokens of *ipse* (x39) and *iste* (x42), in their whole paradigms, registered until the 3rd century BC in the Library of Latin Texts database–Series A (LLT). We classified this sample according to morphological features (case, gender, and number), the referent’s grammatical and semantic information (person, and animacy), and the referent’s pragmatic topicality (given or new topics). This sample is limited to the data registered in the LLT database until the 3rd century BC (*peruētustum*, *terminus ad quem*, and *dubium*), which include works by Livius Andronicus, Naevius, Plautus and Ennius, because these data may represent the Latin system at its earliest stage. This selection led us to exclude some works by the same authors that were classified as 2nd century BC in the database. Nevertheless, the sample is still representative, and data are more comparable, because the same delimiting criterion was applied.

In the second stage of the analysis, we compared the data to another sample, that was limited to the same period and texts, also extracted from LLT, and classified according to the same criteria. This second sample comprises the totality of cases with the inflected forms of the reflexive pronoun *se* (x68), and *sibi* (x32). To this, we added the analysis of the sporadic anaphoric cases of *sam* (2 tokens), *sas* (1 token), *sos* (4 tokens), and *sis* (1 token), only attested in Ennius. We included these data under the hypothesis that these words also derive from PIE **so*³.

³ The sample does not include cases with *sum*, due to the similarity with the verb form (although we checked examples from the literature), nor reduplicated cases (e.g., *sese*), univentions with copular verbs (e.g., *ipsest*), or univented cases of *reapse* and of *secum*. We consider instances of *reapse* to be lexicalizations where the pronominal use of *reapse* is too

II. ANALYSIS OF *IPSE* AND *ISTE* IN AN OLD LATIN CORPUS

In this section, we provide the data obtained from the analysis. We show, individually, the distribution of morphological cases and gender, the referent's animacy and person, and the relationship with the referent's topical state.

1. *Distribution of morphological cases*

Table 1 and Fig. 1 show the distribution of *ipse* and *iste* according to case. We provide some examples of the various cases below in (1a-c) and (2a-c).

Table 1. Distribution of *ipse* and *iste* according to case / number categories

	<i>ipse</i>			<i>iste</i>		
	sg.	pl.	total	sg.	pl.	total
nom.	21	1	22 (56.4%)	7	1	8 (19.0%)
acc.	5	2	7	17	2	19
gen.	0	1	1	2	0	2
dat.	3	1	4	4	0	4
abl.	5	0	5	6	3	9
total			39			42

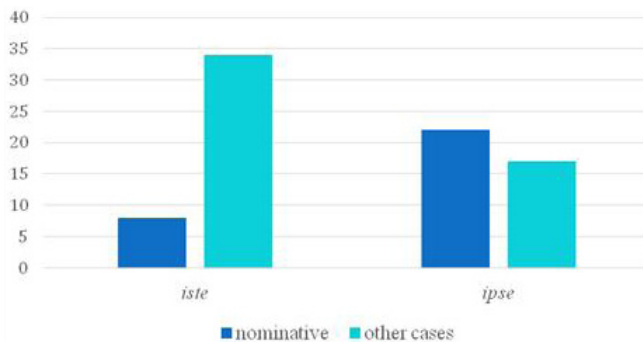


Fig.1. Nominative vs. other cases in *iste* and *ipse*.

opaque. In the cases of forms such as *sum*, *sese*, *ipset*, and *secum*, we expect that their results should not be different from those we show in this paper.

- (1a) *Nam ipse miles concubinam intro abiit oratum suam, ab se ut habeat cum sorore et matre Athenas* (Plaut., *Mil.* 1145)
«Because the soldier himself went inside to ask his concubine to go away from him to Athens together with her sister and mother»
- (1b) *Sed eccam ipsam, egreditur foras* (Plaut., *Mil.* 1215)
«But she's coming out herself»
- (1c) *Quoniam nuntiatumst ipsarum uenturos uiros, ibi festinamus omnis* (Plaut., *Stich.* 676)
«When it was announced that these women's husbands were going to come, we all hustled about» (play's character talking to the audience)
- (2a) *Ducite istum: si non sequitur, rapite sublimem foras* (Plaut., *Mil.* 1394)
«Bring him along; if he doesn't follow, lift him up and carry him out»
- (2b) *Quasi que istius causa amoris ex hoc matrimonio abierim cupiens istius nuptiarum* (Plaut., *Mil.* 1164)
«And that for the sake of my love for him I've left this marriage, keen on marrying him»
- (2c) *Quin tu iubes effferri omnia isti quae dedi?* (Plaut., *Mil.* 1313)
«Why won't you have all that I gave her taken out?»

Table 1 and Fig. 1 show that the nominative is far more frequent than other cases of *ipse*, while in contrast, *iste* in the nominative is attested to a lesser extent than in the accusative and even in the ablative. A quick look at cases with m. sg. *is* and *ille* in the same period, shows that nominatives are more frequent than accusatives but that the gap is narrower: 38x *is* (nom.) vs. 26x *eum* (acc.); 53x *ille* (nom.) vs. 26x *illum* (acc.)⁴. In fact, there is no reason to expect such a difference in the case of *ipse*, allowing us to put forward the hypothesis that the majority of cases of *ipse* in the nominative, especially in the singular, are explained through a diachronic development from an element that ranked high in the topicality hierarchy because the nominative case tends to be more highly topical than other cases⁵. These results are particu-

⁴ A similar tendency is found in the most archaic inscriptions (*CIL* 1): *ipse* is only attested in the nominative, *iste* has only adverbial forms (*istic*, *istoc*), while *is* is attested both in the nominative and the accusative, though the latter to a lesser extent. We thank Eugenio Luján for these observations.

⁵ The higher frequency of the singular in both pronouns could perhaps be explained by the deictic origin of both, in contexts such as stage plays, which represent communicative interactions.

larly interesting in the light of those by Puddu (2005, p. 219), who contrasts the distribution of *ipse* among different authors and demonstrates how the use of the nominative decreases with time: the nominative represents 63.06% of her Plautian corpus, but it decreases in frequency to 48.00% in Sallustius, to 41.94% in Caesar and 26.42% in Cicero. This waning tendency may suggest that these forms were originally connected to the marking of highly topical elements and that their use was slowly lost until *ipse* became wholly integrated into the pronominal system, with no remnants of topical marking. Moreover, both *ipse* and *iste* are finally wholly integrated into the demonstrative system of Romance languages.

Regarding the ablative case, there is a well-known lexicalization in Latin: *re eapse*. This lexicalization is curiously absent from our sample, indicating that the lexicalization is a secondary development. In a previous search of works dated to the 3rd and 2nd century BC in LLT, 13 out of 78 cases with *ipse* were lexicalized with *re*. Most of these cases are found in Terence and none in Plautus.

2. Grammatical gender

As for the distribution of gender, Table 2 shows differences regarding the distribution of neuter and feminine / masculine.

Table 2. Distribution of *ipse* and *iste* according to gender

	m.	f.	nt.	total
<i>ipse</i>	25	13	1 (2.56%)	39 (100%)
<i>iste</i>	13	23	6 (14.28%)	42 (100%)

There is only one neuter case for *ipse*, whereas *iste* appears in the neuter case more frequently, based on percentage. However, the neuter is also the least used gender for *iste* as well. These results contrast to the hypothesis that PIE **so* was masculine and **to* was neuter: although the attested evidence may be due to a secondary development within Latin or the Italic branch, if *ipse* preserves features from **so* and *iste* from **to*, we would not expect neuters for *ipse* and, more critically, we would not expect masculines / feminines for *iste*. Because *ipse* lacks the expected nom. nt. ending in a dental (***ipsud*; cf. Latin *quod*, *id*, *illud*), in contrast to the PIE nom. nt. **tod*, it may

be claimed that the differentiation between *ipsum* and *istud* represents a possible development from two different paradigms. Therefore, in our view, these data contradict the claim that the original demonstrative paradigm was distributed merely according to gender and that it was a suppletive deictic paradigm.

3. Referent's animacy

There is also an unequal distribution between *ipse* and *iste*, in terms of animacy, as shown in Table 3 and Fig. 2. In Table 3, we classify the pronouns according to their syntactic use, referentially (REF) as pronouns or adnominally (ADN) as complements in a (pro)nominal phrase.

Table 3. Distribution of *ipse* and *iste* according to the referent's animacy

	animate			inanimate		
	REF	ADN	total	REF	ADN	total
<i>ipse</i>	26	11	37	0	2	2
<i>iste</i>	11	6	17	6	19	25

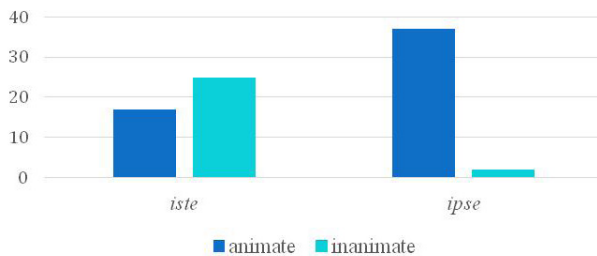


Fig. 2. Distribution of animate and inanimate referents for *iste* and *ipse*.

Table 3 shows that the number of passages containing an animate referent is striking for *ipse*, whereas *iste* does not present such a wide gap between animate and inanimate cases. This difference is notable because there are no reasons that explain it from a synchronic perspective. Moreover, one of the two inanimate cases with *ipse* contain an animized referent, as shown in (3).

- (3) *Non potuit reperire, si ipsi soli quaerendas dares, lepidioris duas ad hanc rem quam ego habe animum bonum* (Plaut., *Mil.* 803)
 «If you'd asked Sol himself to look for them, he couldn't have found for this business two lovelier girls than I have»

As the subject of the verbs *look for* and *find*, the sun (*sol*) is characterized with volition and agency. The second inanimate case has an uncountable noun (*tempus*) as its referent. These two cases together may indicate that uncountable nouns constitute the bridging context towards using *ipse* with inanimate referents in later layers of the language. Most interesting, in an expanded search in Plautus, *ipse* with an inanimate referent is more frequently used as adnominal with uncountable or abstract nouns (e. g., *res*, *tempus*, *ara*, and *iracundia*). In contrast, the two cases of *ipse* with an inanimate referent are limited to adnominal use and introduce a new, not topical referent, as in (3)⁶. Thus, we can expect that, already in Old Latin, the use of *re eapse* as an intensifier ('really') was gradually becoming limited to grammaticalization, as it lost the morphological properties of a demonstrative⁷.

Although all the possible combinations of *iste* are attested in Table 3, the cases of animate referent tend to be referential, consistent with their possible deictic origin, while the cases of inanimate referent tend to have an adnominal use. Interestingly, Pieroni (2010, p. 446) comments that *is*, unlike *ille*, privileges inanimate (and non-subject) referents. One may wonder how to connect this preference with the *ipse* / *iste* distinction, because these two pronouns are both partly diachronically connected to the same stem as *is*. Here, we think that, because this pair comprises unverbated former compounds, *ipse* and *iste* preserve the animate / inanimate distinction as an archaism.

⁶ Given topics «are situationally or textually evoked discourse referents that the speaker assumes the addressee is attending to» (de Vries 1995, p. 523). As this author notes about the Papuan languages, it is expected that the adnominal uses of an anaphoric pronoun will develop secondarily from a deictic pronoun.

⁷ Lexicalizations (*re eapse*) are only attested with inflection in the first part of the word, and there are no cases such as **re ipsa* with the same meaning. This is not isolated data, but the vast majority of inanimate referent cases present initial inflection (e. g., *eopse*, *eampse*, and *sapsa*).

4. Referent's person

Iste can only have 3rd-person referents, while *ipse* also accepts a few 1st-person sg. (x4) and 2nd-person sg. (x2) referents:

- (4a) *Tutin ipsus ipsum uidisti?* (Plaut., *Stich.* 373)
«Did you see him face to face?»
- (4b) *Ipse egomet quam ob rem auctionem praedicem* (Plaut., *Stich.* 207)
«The reason why I myself am announcing an auction»

Ipse's acceptance of local referents (Speech Act Participants, SAP hereafter) is usually associated with its emphatic or contrastive focus or with the in-person value of the discretive use in terms of Pinkster (2015, p. 1152). However, if this is the case, it is also necessary to recall that restrictions based on an animacy or empathy hierarchy can be associated with intensifiers cross-linguistically. As König & Gast (2006) show, languages manifest different cut-off points on such hierarchy (which may be simplified as SAP pronouns > non-SAP pronouns > human NPs > animate NPs > inanimate NPs) concerning the distribution of intensifiers. The sporadic acceptance of 1st- and 2nd-person pronouns by *ipse*, the animate feature of its antecedents, could derive precisely from its characterization as an element located at a high point in an animacy hierarchy. Contrarily, *iste* never has 1st and 2nd person referents. This is particularly significant if we consider that *iste* is frequently associated with local (1st or 2nd) deixis, as Dupraz (2012) claims⁸.

5. Introduction of a new referent

The distribution of *ipse* and *iste* is unequal here as well. While *iste* does not present significant differences regarding the introduction or not of a new referent, the percentage of cases in which *ipse* continues a previously mentioned referent is much higher, as seen in Fig. 3.

⁸ As Pieroni 2010 claims, deixis cannot be restricted to the spatial notion; there are pragmatic, discourse, and textual reasons that trigger the use of different pronouns. However, de Melo (2010, p. 336), states that the classification of two kinds of *ego* (non-correlative and correlative), one determining *hic* and *ille*, and the other one determining *iste* «is just a complicated reformulation of the traditional theory».

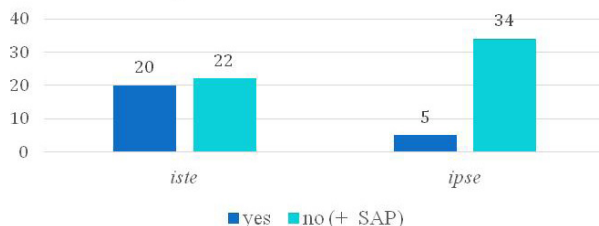


Fig. 3. Introduction of a new referent.

Fig. 3 demonstrates that *ipse* more regularly indicates the continuity of a referent already mentioned (given topics). Because local referents (1st- and 2nd-person) are easily retrievable by the communicative context, these SAP cases are also considered given topics. As in the previous analyses, the use of *ipse* as a topic continuity marker attracts this pronoun to the most topical pole of the hierarchy.

III. REFERENTIAL HIERARCHIES

The data presented in the previous section show that, in Old Latin, *ipse* and *iste* differ in a series of criteria that are related to what is generally known as the referentiality or empathy hierarchy⁹. This section first describes how these hierarchies work, especially concerning languages with a direct (DIR) / inverse (INV) voice system. Then, we offer a description of how *ipse* and *iste* work from the perspective of such hierarchies.

1. Referential hierarchies in languages with direct / inverse voice systems

According to Haude & Witzlack-Makarevich (2016), a referential hierarchy «is a scalar representation of types of referents or referring expressions that are ranked according to their deictic, semantic, and / or discourse-pragmatic properties» (p. 433). Although each language can establish more specific

⁹ Silverstein (1976) initially introduced the animacy hierarchy. Different hierarchies have since been proposed in the same vein, including referential and empathy.

intrinsic distinctions, the general pattern of a referential hierarchy is given in Fig. 4:

SAP > 3rd-person human > 3rd person > non-human animate > inanimate

Fig. 4. Referential hierarchy (Haude & Witzlack-Makarevich 2016, p. 434).

Some authors, such as Jacques & Antonov (2014), refer to this hierarchy as one of empathy: «speakers, being animate and humans, are more likely to empathize with (i. e., take the viewpoint of) human beings than animals, and of animals than inanimates» (p. 305). Thus, a local referent usually outranks a 3rd-person referent, either present in the context or absent but shared between the speech act participants. Further from the empathic pole are the animate referents who do not participate in the situation. Even more distant are the inanimate referents, which almost do not share characteristics (or generate the least degree of empathy) with the participants of the discursive situation. The connection between empathy and topicality is rather straightforward: the further to the left the referent is, the more likely it is to be topical, and a topical or known referent in the speech ranks higher (further to the left) than a non-topical one.

Referential or empathy hierarchies usually determine different morpho-syntactic features of a language and are definitional of direct / inverse systems (Jacques & Antonov 2014), as occurs in Algonquian and Sino-Tibetan languages¹⁰. The distribution of direct and inverse markers can be defined according to the following rule for a transitive construction: «If the patient is higher on the hierarchy than the agent, the verb receives inverse marking; conversely, if the agent is higher on the hierarchy than the patient (or if both are equal), the verb receives direct marking» (Jacques & Antonov 2014, p. 304). In other words, the inverse voice indicates that the clause must be interpreted in the opposite direction to the hierarchy.

¹⁰ An inverse system implies the existence of a differentiated type of alignment, the so-called hierarchical alignment. However, as Zúñiga (2007) states, it is unnecessary to admit a differentiated alignment because the distinction between direct and inverse voice can occur in languages with other types of alignment, such as ergative and semantic.

As Jacques & Antonov (2014) show, not all hierarchical languages display equally canonical direct / inverse systems, and specific hierarchies are required for some languages. However, regardless of the different reasons for a direct or inverse construction to be triggered in each language, Jacques & Antonov propose that the direction SAP > 3rd animate > inanimate is the most transversal in most languages.

Not only semantics but also pragmatics may determine the choice of a direct / inverse construction, as Golluscio & Hasler (2017) demonstrate for the case of Mapudungun. A referent that remains topical throughout the discourse, even if it belongs to the lowest of the hierarchy (for example, an inanimate referent), may not trigger an inverse construction due to its discursive prominence. This situation could reveal a tendency for the hierarchy's highest elements to remain topical in the speech, perhaps linked to their agency. Jacques & Antonov (2014, p. 304) make similar observations for Japhug, a Sino-Tibetan language: «in languages with direct-inverse contrast in the non-local domain, the choice of proximate or obviative status for a particular referent (and thus the use of an inverse or a direct form) is guided by both semantics (relative animacy of the agent and patient) and pragmatics (the relative topicality or saliency of the two referents)».

One could conclude that an inanimate referent, or one low in the hierarchy, requires a marked construction to remain topical. The data on *ipse* and *iste* could also be related to this phenomenon because the referents for *ipse* tend to be animate and highly topical. In contrast, *iste* usually has inanimate and non-topical referents because this pronoun precisely introduces new referents in the discourse.

In the case of transitive constructions, especially in cases with 3rd-person referents (i. e., non-local), languages with direct / inverse systems make a further classification of arguments as proximate (PROX) and obviative (OBV). While a proximate indicates the highest element on the referential scale, the obviative indicates the lowest or least accessible argument to the speech act participants. This is usually schematized as 3 > 3'. Although PROX and OBV arguments are typically indexed in the verbal morphology, languages can further mark such distinction on nouns and pronouns to reinforce the interpretation of the direction of the event.

- (5) Plain Cree (Wolfart 1973, p. 25, quoted in Jacques & Antonov 2014, p. 303)
- | | | |
|-----------------|--------------|----------------|
| <i>sēkih-ik</i> | <i>nāpēw</i> | <i>atimw-a</i> |
| scare-INV | man | dog-OBV |
- «The dog_(OBV) scares the man_(PROX)»

Example (5) shows that in Plains Cree, the highest element in the hierarchy scale (PROX) is unmarked, while the OBV argument is marked with *-a*. Because the lowest element in the hierarchy (‘dog’) is the one that acts on the highest (‘man’), an inverse construction is triggered.

In mixed 3 > SAP scenarios, it is expected that an inverse construction is triggered without marking the proximate; this occurs because SAP arguments are always higher than 3rd-person referents. However, even SAP arguments may receive proximate marking, as Bickel (2011) notes.

2. *The referential hierarchy in Old Latin and other ancient Indo-European languages*

From the perspective of the hierarchies introduced above, it is evident that the relationship between *ipse* and *iste* suggests the remnants of a referential scale, because this pair is sensitive to the animacy and topicality features, as represented in Fig. 5.

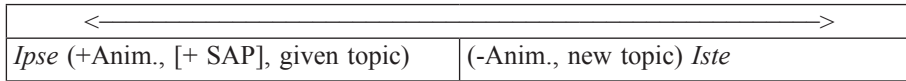


Fig. 5. *Ips*e and *ist*e according to an empathy scale.

On the one hand, *ipse* tends to accept more animate (and even some SAPs) than inanimate referents, to have a known / recoverable referent (given topic) by the speech act participants, and to appear most frequently in the nominative case, the most topical case. On the other hand, *iste* tends to have inanimate referents, to introduce new topics, and to appear in other cases. *Ips*e, thereby, tends to outrank *ist*e in a referential hierarchy.

In Old Latin, not only *ipse* and *iste* follow a referential hierarchy. Similar remnants can be found in other derivations from **so/*se*: the inflected forms of the reflexive pronoun, *se*, *sibi*, and those of the anaphor (cf. Mendoza 1998, p. 75), *sam*, *sas*, *sos* and *sis*:

- (6a) *Circum sos quae sunt magnae gentes opulentae* (Enn., *Ann.* III 156)
«Around them, who are great people and wealthy»

- (6b) *In foribus scribat occupatam esse se* (Plaut., *Asin.* 760)
«She shall write on the door that she is engaged»

- (6c) *Haec ita me orat, sibi qui caueat aliquem ut hominem reperiam* (Plaut., *Bacch.* 41)
 «This one asks me to find her someone who takes care»

Most uses of these forms also rank high in a referential hierarchy. Although *se* is usually identified as a reflexive pronoun, in Old Latin, it still functions much more frequently as a deictic or anaphoric pronoun that sets coreference with an NP beyond the limits of the minimal (subordinate) clause. Only some of the 60 passages containing *se* in our sample are reflexives in the strict sense. The use of *se* as the subject of a subordinate clause with infinitive can be seen as a topic continuity phenomenon; in these cases, its referent is most likely a given topic because it can be tracked in the main (superordinate) clause, and has usually been mentioned before. In fact, Cenamo (1993) and Puddu (2005) consider *se* to be a topic continuity marker, which has much in common with our claim for *ipse*¹¹. Personal pronouns such as *me* and *te* may also be accusative subjects of subordinate clauses with infinitives, but it is unnecessary to restrict such use to the so-called reflexive pronouns because, strictly speaking, in the case of coreference with the 1st and 2nd person, Latin uses personal pronouns (clearly animate, human, and individualized, high in the empathy hierarchy) and not exclusively reflexives.

Furthermore, *se* and *sibi* have animate referents in all passages from our Old Latin sample, while the number of inanimates slowly rises during the Classical period and reaches a 16% of the sample's total (x16 out of 120), as Orqueda & Toro (2020) show. Similarly, *sam*, *sas*, *sos* and *sis* tend to have animate and topical referents, known or shared by the participants. It is reasonable that all pronominal forms that continue the PIE root **so/(*)se* preserve remnants of the indexing of arguments high in the empathy / referential hierarchy, unlike *iste*.

The observations above have identified parallels in other ancient Indo-European languages. For example, the Sabellian demonstratives that contain a segment derived from **so* accept 1st- and 2nd-person referents, and they

¹¹ Multiple authors posit that derivatives from the **só* (or **se*) are connected to a topic continuity function. For example, Petit (1999) proposes that the derivatives of **s(w)e* refer to the topic and not necessarily the subject in Ancient Greek, Schmidt (1978) postulates that **se* was originally an anaphor / reflexive pronoun, and Puddu (2005) concludes that **se* was fundamentally an anaphor that could, eventually, refer to the subject.

are usually animate. It is worth noting that most Sabellian examples that, according to Dupraz 2012, derive from Common Italic (CIIt., hereafter) **eko/ekso* (<**ek-* + **so*), present more exophoric cases with animate referents in his book. Furthermore, when Dupraz presents the rare complex passages with *essuf / esuf* (equivalent to *ipse*), he also uses examples in which the referent is topical or has a 2nd-person referent (cf. Dupraz 2012, p. 242 ff.), and Dupraz (2012, p. 259) adds the possibility of interpreting the only attested case of South Picene *-sa* (<**so*) as an enclitic that may indicate «that the hearer or reader can identify the referent of the co-occurring noun». In turn, this author claims that forms derived from CIIt. **esto* (Sabellian *esto* and Latin *iste*) are proximal exophoric pronouns, but he does not mention that these examples have inanimate referents, as they do.

While the evidence provided above does not allow for reconstruction of demonstratives further than CIIt., similar claims can be made for other branches. For instance, only 2 out of 65 cases of the Vedic demonstrative *sá-* in the nom. m. sg. are inanimate in a sample from Book II in the *Rgveda*, according to our own research. In these cases, the demonstrative most frequently refers to given topics, which, in some instances, can even be 1st or 2nd person, as seen in (7), and as noted by several authors¹².

(7)	<i>sá</i>	<i>tvám</i>	<i>no</i>
	DEM.NOM.SG	2SG.NOM	1PL.GEN
	<i>agne</i>	<i>'vitéhá</i>	<i>bodhi</i> (<i>Rgveda</i> III 19.5)
	Agni-VOC	protector-NOM.SG	2SG.IMP
	«Be yourself our protector, o Agni!»		

Also, in a sample of 19 inflected forms of the reflexive pronoun, *sik* (<**se-ge*), *seina*, and *sis* in Gothic (cf. Ferraresi 1996; Puddu 2005), all attestations of this pronoun, which is not restricted to the reflexive function, have animate referents. The possibility of an analysis similar to that for other cognates would point to a reconstruction of a PIE pair sensitive to the features already described.

¹² Demonstratives with *t-* may also be part of phrases headed by a 1st or 2nd personal pronoun, such as *té vayám* literally, ‘these we’. More research on such cases is still needed.

IV. *IPSE* AND *ISTE* FROM AN ETYMOLOGICAL PERSPECTIVE

This section features an examination of the diverse etymologies claimed for *ipse* and *iste* from the perspective of our previous analysis. Most of these etymologies accept that these pronouns can be divided, at least, into two parts that can be reconstructed individually. Following the earlier research, we introduce in this section the proposals dealing with these segments: initial *is-*, final *-se/-te*, and the controversial middle *-p-* in *ipse*, the third segment of which lacks a clear origin. In the next lines, we offer an overview of the most relevant etymological approaches to each pronoun. At the end of this section, we argue in favor of a parallel origin for one of these pronouns' segments.

1. *Iipse*

Because the etymology for *ipse* changes depending on the interpretation of the *-p-* element, we must analyze this peculiar plosive before we may discuss the initial and final segments. There are three main approaches to *-p-*: an epenthetic, non-morphemic element; part of a **-pse* morpheme; or part of a **-pe-* morpheme.

According to the first of the above hypotheses (e. g., in Zamudio 1989), *-p-* emerges in the acc. sg. *eumpse*: **eum-se* > *eumpse*. Fortson (2011, p. 211) considers that *-p-* comes about from a secondary transitional consonant in the acc. sg. **eum-se* (or **eum-sum*), **eam-se* (or **eam-sam*) > **eumps-*, **eamps-*, and before spreading to the rest of the paradigm by analogy. To Meiser (1998, p. 164), it is an excrescent *-p-* that emerges between two consonants, in duplicated contexts, e. g., **sam-sam* > *sampsam*.

According to a second approach (e.g., Sihler 1995; Monteil 2003), *-p-* is part of an invariable morpheme *-pse*. Monteil claims that *-pse* would have an empathic value similar to *-dem* (in *isdem*) and an unclear origin. Initially, *-pse* would have stayed invariable with the inflection in the initial root *i-*.

Within the third approach, **-pe-* was an independent particle that lost its vowel via syncope, once it was added to the base (as can be inferred from the Sabellian evidence). According to de Vaan (2008, pp. 308, 452-453), this phenomenon is also potentially evident in the attestation of other words such as the Latin adverb *quippe* 'certainly' (< **kwid-pe*), Lithuanian *-p* (discourse particle) and *kaíp* 'how?', Hieroglyphic Luwian */kwipa/* 'indeed', and Lycian / Milyan *tibe=kibe* 'or'. From this perspective, *ipse* and *sapsa* would derive

from the three-parts compound form **so-pe-so*¹³, in parallel to other forms that would derive from the compound **so-so*¹⁴. In a similar vein, Berenguer (2000) concludes that the particles based on **p-* must have been used to emphasize or reinforce the preceding element. We agree with de Vaan (2013) in that the non-morphemic *-p-* is the weakest hypothesis within this context due to the existing comparative evidence. The individuation of a *-pe-* morpheme let us compare animacy, topicality, and grammatical features in other derivations without that morpheme or its cognates, such as Vedic *ayám sá* ‘this one’ < **(h₁)is-* plus **só-* (Dunkel 2014, p. 371). Therefore, in the same vein as these proposals, we believe that the clitic **=pe* could be added to diverse demonstrative and pronominal forms without altering their original syntactic behavior or the possibility of creating more complex compounds¹⁵.

Two etymologies have been proposed for the initial segment, **i-*. According to Weiss (2009, pp. 346), m. *ipse*, f. *ipsa*, nt. *ipsum* in Old Latin decline as if it were a compound of *is*, *ea*, *id*; thus, the first segment must derive from the same root from which *is* develops. Weiss (2009, pp. 346-347) hypothesizes that this paradigm originates from the combination of *is* and the archaic anaphor *sam*: **is-so* > **isse* > *ipse*, and **eam-sam* > *(*)eampsam* > *eampse*. Sihler (1995, p. 394-395) further claims that «once this **ispse* had become *ipse*, which obscured its morphological makeup, the way was open for the reinterpretation of the form as parallel to *ille* and *iste*». Then, a remodeling process would have occurred, and *ipsa* arose for *eapse*, *ipsum* for *eumpse*, and so forth. Therefore, forms with double inflection, like *eumpsum*, should have arisen during the transitional period. According to a second proposed etymology, the first segment derives from **so*. Thus, de Vaan (2008, p. 308) reconstructs **so-pe-so* for Clt., before the initial *i-* (<**e(s)*) spread to

¹³ If the original compound were **sopeso*, there would have been two different pathways to its creation: either through reduplication or by adding an enclitic particle to a base with the original proximate marker, as a way of re-characterization.

¹⁴ See de Vaan’s (2013) discussion of the analysis of **soso* as a compound or as a case of reduplication. We consider **so-so* and **so-pe-so* as compounds because there is plenty of evidence of expressions created through compounds in different stages of their grammaticalization and univerbation.

¹⁵ As a reviewer suggests, more research on **-pe-* is needed, but it goes beyond the limits of this paper, as the presence of *-p-* in Latin *ipse* has no direct incidence in our analysis. As Weiss (2009, p. 347) notes, *isse* might be preserved in *CIL* 4.148. This could suggest that forms with and without *-p(e)-* were in competition at the earliest stage.

the rest of the paradigm and was no longer recognized as a pronoun. This etymology would explain the sporadic attestations of *sapsa* and *sumpse* (cf. Weiss 2009, p. 339) and could suggest that different kinds of compounds must have overlapped or competed in the earliest stage. Note, however, that *sapsa*, *eumpsum* and *sumpsum* are not attested in our sample, although *sapsa* occurs twice in 2nd century BC; in turn, the only attestation of *sumpse* (Naev., *Com.* 96) is as a verb, not as part of the pronoun's paradigm.

For the final segment, *-*se*, de Vaan (2008) and Fortson (2011) argue in favor of a derivative from **so*. Per this hypothesis, forms such as the nom. f. sg. *sapsa* and the acc. m. sg. *sumpse* would indicate use of the same pronoun in both parts of the compound. In our view, although the exact relation between PIE **so* and Latin -*se* is still unclear, the claim that the last part of *ipse* derives from the PIE **so* appears logically sound as long as the proto-form is an element that ranks high in a referential hierarchy.

2. *Iste*

Similar to *ipse*, this pronoun is usually considered as composed of two parts. Sihler (1995) states that it is difficult to tell whether it developed from **es-* into *is*, or vice versa, due to the contradictory evidence in Latin and Sabellian, the latter of which preserves a cognate, *esto*. Dupraz (2012) considers that Sabellian would have retained the original form while Latin would have innovated, and he reconstructs **esto-* for the Clt. group. A further problem concerning the etymology of the first element in *iste* is evident in Sihler's (1995, p. 394), who doubts about whether the first element is a pronoun or a particle because of the lack of attestation of forms with inflection in the first segment. However, a clue for the inflection in the first element is found in the reconstruction of forms such as *eāste* (acc. f. pl.); this form is not preserved in any edition, but Norden (1939) restores it in a passage attributed to Varro (L. 7.8), which would reflect an ancient augural formula that would have been modernized to be better understood (cited in Bailey 1940, p. 44). Although this proposal could be related to non-Latin forms such as the Umbrian nom. / acc. sg. nt. *este* 'this', Weiss (2009, p. 345) states that «the passage is among the most vexed in all of Latinity». Furthermore, Sihler (1995) posits that this could also be a case of reanalysis among the speakers when considering *iste* as a compound of *is* (*ea*, *id*) plus -*te* (see also Dupraz 2012, p. 302). In our view, the current evidence does not allow us to reconstruct for PIE the exact morphology and function

of the first segment in Latin *iste*, although it is clear that it must have been a deictic or demonstrative pronoun that could eventually have been integrated as a part of a compound. Regardless, the etymology does not affect our observations on the second segment.

Sihler (1995) considers that the second element in *iste* may derive from an inflected deictic pronoun **to*, although it is also possible that this was originally an uninflected particle. Weiss (2009) similarly proposes two possible origins for *iste*: this pronoun was formed either as a compound of the pronoun *is* plus the particle *te* (and the pronoun *is* would have become invariable with inflection then only in the second part of the compound), or as a compound of *is* plus the PIE pronoun **so-/*to-*, from which it would have taken only the *t*-variant: **is-to*. In turn, de Vaan (2008) believes that Latin *iste* and Proto-Sabellian **esto* would have a common origin, **es-to-*, if Latin replaced **es-* for **is*. The second element would be identified with the PIE pronoun **so-/*to-*. Last, and although the investigation in Dupraz's (2012) is focused on Sabellian and not Latin, this author seeks an etymology that can jointly explain the attested forms in the different Italic branches. Thus, he proposes that *iste* derives from an exophoric proximal suppletive demonstrative paradigm, Clt. **i-/*ey-/*e-sy-/*e-sm-*, which would have later acquired an endophoric meaning. According to Dupraz, the root **e-sm-* would have split from this paradigm, creating a new suppletive paradigm, **e-sm-/*es-t-*, which would have been thematized in **es-mo-/*es-to-*. Thus, in Common Sabellian this would have been preserved as **esmo-/*esto-*, while in the Latino-Faliscan group the suppletive paradigm would have disappeared, with **esto-* as the only remnant. Then, **esto-* would have evolved into **isto-* by analogy with the root **i-*. The final form of the demonstrative would have a parallel with the distal *ollo* > *illo* > *ille*; **esto-* > **isto-* > *iste*. Under Dupraz's hypothesis, the archaic form *eāste* is understood as a meta-analysis in consideration of that *iste* as a compound of **is*, the nom. m. sg. from **i-/*eyo-/*ey-/*e-sy-*, and an invariable **-te*.

Regardless of the original form¹⁶, most approaches connect the second part of *iste* with **to* and not with **so*. This is significant if we consider the

¹⁶ Both the hypothesis of an original particle and that of a flecional pronoun for the final segment are complex. In the former case, it is not straightforward what the original function of this particle was in PIE; in the latter case, it is not convincingly explained either why the inflected form is crystallized as uninflected, as would be the case with *eāste*.

problematic connection of *iste* to nearer-hearer deixis. Several authors have already highlighted this inconsistency. As recalls Pieroni (2010, p. 406), there is a long-standing discussion with opposing arguments regarding the connection between *iste* and the second-person deixis. In contrast, Dupraz (2012) proposes that *iste* would develop from a Clt. exophoric pronoun, **esto*, that indicated proximity to the speaker. However, the nearer-speaker deixis is also frequently unlikely, as evident in the following passages:

- (7a) *Nempe ut adsimulem me amorem istius differri* (Plaut., *Mil.* 1162)
 «I should pretend that I'm torn apart through my love for him»
- (7b) *Edepol qui te de isto multi cupiunt non mentirier* (Plaut., *Mil.* 777)
 «Indeed, many people wish you were lying about that now»

While all these claims fail to explain the prevalence of **t-*, and not **s-*, in the second segment of *iste*, our analysis suggests that the preference for **t-* may be due to its low rank in a referential hierarchy, in comparison to **s-*.

3. *Is a common origin plausible?*

As shown above, several authors point to a partial etymological connection between *ipse* and *iste*, but the literature seldom suggests any clues for the preference for **s-* or **t-* in each case. The traditional distinction between **so-* and **to-* in PIE was according to the gender feature in the nominative case (e.g., > Ved. m. *sá*, f. *sā*, nt. *tad*), but the reconstruction of gender-based suppletive demonstratives has been extensively discussed from different perspectives, and one of the most complex problems is the assumption of morphological suppletion only for the nominative case, which, among other problems, has led scholars to seek alternative solutions. Thus, for example, Berenguer (2000) suggests that **so* and **to* were particles with a primitive deictic-anaphoric value and that their use depended on the introduction of a new / old topic. Our proposal concurs with Berenguer on different points of this problem. However, we consider that the results regarding the distribution of the animacy feature are closely connected to that of topicality. Furthermore, we have shown that all results from the earliest attested layer of the Latin language point toward a unified explanation of the use of **t-* to indicate values of a deictic pronoun low in a referential hierarchy. These values must have remained residually from a previous obviative marker.

V. DIACHRONIC DEVELOPMENT IN INDO-EUROPEAN ACCORDING TO A REFERENTIAL HIERARCHY

Referential or empathy hierarchies are not new in the field of PIE studies; various authors have linked it to diverse grammatical phenomena. Thus, for example, a referential scale would determine the formation of the female gender (cf. Luraghi 2011). Based on the Hittite data, Sturtevant (1939, pp. 16-17) proposes a distinction between the particles **so* and **to*, according to which the former would have indicated no subject change, and the second subject would have denoted a subject change in the new clause (cited in Berenguer 1998, p. 265). According to Berenguer (1998 and 2000), these initially deictic particles would have developed, among other ways, as inflectional pronouns in the post-Anatolian period. In this way, Berenguer (1998 and 2000) seeks a solution to the apparent problematic restriction of the alternation in the nominative case. The evidence for pronominal endings in the ancient Indo-European languages challenges the reconstruction of a case system indeed; however, we believe that the hypothesis of original non-inflectional particles that develop (jointly, in many cases) toward different case systems is also troublesome.

Early works also connected the **so/*to* alternation to an original animate / inanimate opposition (e.g., Szemerényi 1978, p. 265). This approach later received negative assessments, again due to the restriction of this alternation to the nominative case and to non-Anatolian languages. As notes Berenguer (2000, p. 145), this controversy led some authors to propose a primitive ergative or active (semantic) system, and therefore, the hierarchy of empathy is also fundamental in discussions about grammatical alignment. While the claims against a nominative-accusative type have progressively gained ground, an intense discussion persists about whether the alignment type was mainly ergative or active / semantic¹⁷. In the confrontation between these two options, the role played by the animacy hierarchy in PIE is crucial. Some authors believe that the differentiation between m. **so* and nt. **to* determine the existence of an ergative type. Based on this, Rumsey's (1987) idea that an ergative type of alignment in PIE would violate Silverstein's (1976) animacy scale is a fundamental milestone because, according to this scale,

¹⁷ For an extensive discussion of the different existing claims, see Pooth et al. (2019).

highly animate referents are likelier to be coded according to an accusative alignment type, while the lowest, as inanimates, are more likely to follow an ergative system (Pooth et al. 2019, pp. 247-248). This would mean that only inanimate nouns would require an agent marker (A) in transitive constructions. In proposals for a type of semantic or split alignment, such as that by Pooth et al. (2019), the nouns in argument slots would have accepted different markings according to their +/- agentive feature. Thus, in an active and transitive clause, a highly animate NP in subject position would receive an agentive marker, while an inanimate (or lower on the scale) would receive an absolutive or antiagentive marking.

Recently, Griffith (2018) has shown that Old Irish *notae augentes*, which are pronominal clitics added to verbs, are sensitive to a referential hierarchy because they have only animate referents. It is interesting that these clitics (*sa*, *se*) most probably derive from PIE **so/*se* and that Griffith connects their use to a hierarchical alignment, although he believes that this alignment must be an innovation within the Celtic branch¹⁸.

Evidently, there are important precedents that connect the **so/*to* root with a referential hierarchy in PIE that is partially preserved in some ancient cognates. In contrast, both the idea of animate / inanimate alternation and that of change / no change of subject support our hypothesis that the PIE root pair **so/*to* indicated a proximate / obviative distinction and that some of these roots' derivatives (such as *ipse*, *iste*, *se*, *sibi*, *sam*, *sas*, *sos* and *sis* in Old Latin) inherited specific features of this kind of indexing regardless of their natural diachronic grammatical and semantic development within each branch. Furthermore, the empathy hierarchy applied, at least, to the 3rd-person deictic system plausibly indicates the differentiation between proximate and obviative markers in a language with an inverse system, as Pooth (2018) proposed. Derivatives from **so* tend to take the nominative case, and therefore, to take the subject position, they tend to indicate highly animate (and even human) and topical referents. Derivatives from **to*, like *iste*, show the opposite tendency: *iste* has inanimate referents more frequently, it is not restricted to the subject position and introduces new topics. This would explain why *ipse* ranks high on the empathy hierarchy.

¹⁸ We thank Cormac Anderson for pointing to the discussion on animacy in Celtic.

Within this scenario, the most convincing morphophonological reconstruction for *ipse* is the compound form $*(h_1)i(s)=pe=se/so$, in which $*(h_1)i(s)$ represents an original demonstrative or deictic root (> Lat. *is*, *ea*, *id*, Ved. acc. m. sg. *ím*, Goth. *is*, and so forth) and $*=pe$ a contrastive focus or emphatic marker clitic (Dunkel 2014, pp. 622-624 and 363-366). Given that the last segment ($*=so$ or $*=se$) was added as clitic, it is not possible to confirm at this point if the added form was the original clitic or the full pronoun, which, in principle, should be the essential distinction between $*=se$ and $*so$ respectively. As a consequence, it is not possible to claim whether the compound received final inflection in PIE or if this was a later development. These and other issues should be left for further investigation.

After a process of univerbation and grammaticalization of the compound, the middle *-e-* would have been lost through syncope. This process must have been long and slow, which would suggest the preservation of some cases with inflection in the initial segment (such as *eapse*, *eaapse*, and *eopse*) in Old Latin¹⁹. Nonetheless, the initial inflection is abandoned at some point and externalized, in line with the rest of the pronominal system. The externalization of inflection in PIE has been noted by various authors; see especially Luján & Mendoza (2019) for a detailed analysis of the externalization of inflection in the case of pronominal stems plus the particle **sme*.

For cases like nom. f. sg. *sapsa*, and acc. m. sg. *sumpse*, as stated above, these are peculiar formations for two reasons: first, because they seem to be the only cases not derived from the previous reconstruction, and secondly, because of their sporadic use: *sumpse* does not occur and *sapsa* is only attested twice; in both cases, it is part of the lexicalization with *res* (*re sapsa* ‘evidently’). While de Vaan (2008) and Dupraz (2012) opt for an initial segment also based on the root $*s(o/e)$ - later replaced by another one, we cannot reject the possibility that there were two coexisting variants, of which one quickly lexicalized, letting the other one succeed.

In any case, the advantage of reconstructing a compound is the parallel with the reconstruction $*(h_1)i(s)=te/to$ > *iste*. These reconstructions imply

¹⁹ If the stem had taken the non-inflectional form initially, forms such as *eapse* or *eopse* would not be possible. In turn, the non-attestation of **ispse* could be explained as a case of early elision of *-s-* due to phonetic reasons. If this is the case, the parallel between *ipse* and *iste* becomes clearer. Note also that the phonetic context does not favor such elision in the latter pronoun, so *-s-* is preserved.

that *ipse* and *iste* are formed by a compound process in which the first element is a demonstrative item in both cases and the last one is a deictic that indexes an argument high or low on the empathy hierarchy. This connection offers a plausible fresh look at the distribution of the whole deictic and pronominal system in PIE. Once this system collapsed, these two deictic markers (**so* and **to*) that originally marked a proximate and an obviative argument would have merged or been reinterpreted in terms of spatial deixis as closer to the 1st and 2nd person, respectively. The hypothesis of the reinterpretation of **so/*to* as spatial deictics is not rare. As Mendoza (1998) already noted, this pair was originally not sensitive to the spatial parameter. In addition, the combination with inanimate referents would have been secondary to the use of these forms as deictics because the expected diachronic change is from deictics toward anaphorics through a resumptive function²⁰.

Our data also suggest that these deictics were used in compounds with other pronouns, and then, such compounds could have been reinterpreted as new exophoric demonstratives. This would explain the parallel of cases such as Latin *ipse* with Ved. *ayám sá* and Old Prussian *stas* <**kói-to*. If this hypothesis accurately represents this linguistic shift, we can expect that the first uses of these compounds were deictic and tracked a 3rd-person referent from the context of a speech act. Once grammaticalized, what could have happened before Clt. (if we agree with Dupraz 2012), the new forms, created through univerbation and semantic bleaching of the forming units, would have been extended to adnominal uses again, allowing for reinterpretation of *ipse* as a focal or identity pronoun.

VI. CONCLUSIONS

In this paper, we have jointly addressed the etymology of *ipse* and *iste*, based on the recurring references to the possible derivation of both pronouns from the suppletive root **so/*to*. We applied different tests to determine similarities and differences between these pronouns' referents in Old Latin, and all the results favor a distinction sensitive to an empathy scale and are congruent with other data, both within Italic and other IE branches. This allows us to conclude

²⁰ De Vries (1995) proposes that the resumptive function of demonstratives must have been an intermediate step toward other contexts in some Papuan languages.

that *ipse* and *iste* (regardless of their usual synchronic descriptions in Old and Classical Latin) preserve some features of a proximate (**so*) and an obviative index (**to*), respectively. Thus, *ipse* outranks *iste* in a referential or empathy hierarchy. Nevertheless, these pronouns show remnants but are not PROX and OBV markers themselves.

If **so* and **to* indexed proximate and obviative arguments, the language must have also differentiated between direct and inverse systems in the verb, as Pooth (2018) claims. Further, this claim has further consequences for the reconstruction of alignment types for PIE. Siewierska (1998) raises the possibility that a language with an inverse voice system diachronically develops a language with ergative alignment. In turn, Pooth et al. (2019) connect this feature to the existence of a type of semantic alignment in PIE. Although these connections transcend the limits of this paper, we believe that further research into these hypotheses is needed and that, even when no direct / inverse system is attested in an ancient IE language, a distinction between proximate and obviative is typologically valid and may be consistent with various other aspects of PIE grammar.

We hope that our research can contribute to a more exhaustive knowledge of the Latin language in its most archaic state. In addition, the objective of this paper was to increase openness in the reconstruction of the PIE grammatical system, a reconstruction not based solely on the attested comparative data but also on the cross-linguistic evidence offered by typological research.

BIBLIOGRAPHY

- Álvarez, O. (2009): «Pronombres», in Baños, J. (coord.), *Sintaxis del latín clásico*, Madrid, Liceus, pp. 273-297.
- Bailey, C. (1940): «Old Roman Religious Texts», *The Classical Review* 54, pp. 44-46.
- Bassols, M. (1992): *Sintaxis latina*, 10ª ed., Madrid, CSIC.
- Beekes, R. & de Vaan, M. (2011): *Comparative Indo-European Linguistics. An Introduction. Second Edition, Revised and Corrected by Michiel de Vaan*, Amsterdam, Benjamins.
- Berenguer, J. A. (1998): «La heteróclisis IE *SO/*TO», *Emerita* 66, pp. 257-290.
- Berenguer, J. A. (2000): *Estudio sobre las partículas indoeuropeas con base consonántica y laríngea*, Madrid, CSIC.
- Bickel, B. (2011): «Grammatical Relations Typology», in Song, J. (ed.), *The Oxford Handbook of Language Typology*, Oxford, Oxford University Press, pp. 399-444.

- Cennamo, M. (1993): *The Reanalysis of Reflexives: A Diachronic Perspective*, Naples, Liguori.
- de Melo, W. (2010): «Morfosintassi latina. Punti di vista by N. La Fauci; S. Pieroni». *Mnemosyne* 63 (2), pp. 336-340.
- de Vaan, M. (2008): *Etymological Dictionary of Latin and the Other Italic Languages*, Leiden-Boston, Brill.
- de Vaan, M. (2013): «Reduplicated Demonstratives in Ancient Indo-European Languages», *Transactions of the Philological Society* 113 (1), pp. 38-52.
- de Vries, L. (1995): «Demonstratives, Referent Identification and Topicality in Wambon and Some Other Papuan Languages», *Journal of Pragmatics* 24, pp. 513-533.
- Dunkel, G. (2014): *Lexikon der indogermanischen Partikeln und Pronominalstämme*, Heidelberg, Universitätsverlag Winter.
- Dupraz, E. (2012): *Sabellian Demonstratives: Forms and Functions*, Leiden-Boston, Brill.
- Ferraresi, G. (1996): «Middles, Reflexives, and Ergatives in Gothic», in Brandner, E. & Ferraresi, G. (eds.), *Language Change and Generative Grammar*, Opladen, Westdeutscher Verlag, pp. 273-291.
- Fortson, B. (2004): *Indo-European Language and Culture. An Introduction*, Oxford, Blackwell.
- Fortson, B. (2011): «The Historical Background to Latin Within the Indo-European Language Family», in Clackson, J. (ed.), *A Companion to Latin Language*, Chichester, Wiley-Blackwell, pp. 199-219.
- Golluscio, L. & Hasler, F. (2017): «Jerarquías referenciales y alineamiento inverso en mapudungun», *RASAL lingüística* 1, pp. 69-93.
- Griffith, A. (2018): «Diachrony and the Referential Hierarchy in Old Irish», in Cristofaro, S. & Zúñiga, F. (eds.) *Typological Hierarchies in Synchrony and Diachrony*, Amsterdam, Benjamins.
- Haude, K. & Witzlack-Makarevich, A. (2016): «Referential Hierarchies and Alignment: An Overview», *Linguistics* 54, pp. 433-441.
- Jacques, G. & Antonov A. (2014): «Direct / Inverse Systems», *Language and Linguistics Compass* 8, pp. 301-318.
- König, E. & Gast V. (2006): «Focused Expressions of Identity: A Typology of Intensifiers», *Linguistic Typology* 10 (2), pp. 223-276.
- Luján, E. & Mendoza, J. (2019): «La externalización de la flexión en los pronombres indoeuropeos y el alargamiento *-sm», *Revista Española de Lingüística* 49 (1), pp. 235-266.
- Luraghi, S. (2011): «The origin of the Proto-Indo-European Gender System: Typological Considerations», *Folia Linguistica* 45, pp. 435-464.
- Meiser, G. (1998): *Historische Laut- und Formenlehre der lateinischen Sprache*, Darmstadt, Wissenschaftliche Buchgesellschaft.

- Mendoza, J. (1998): «Pronombres, adverbios y partículas», in Rodríguez Adrados, F., Bernabé, A. & Mendoza, J. (eds.) *Manual de lingüística indoeuropea*, vol. III, Madrid, Ediciones Clásicas, pp. 3-124.
- Monteil, P. (2003): *Elementos de fonética y morfología del latín*, Sevilla, Universidad de Sevilla.
- Norden, E. (1939): *Aus altrömischen Priestbüchern*, Lund, Gleerup.
- Orqueda, V. & Toro, F. (2020): «Diachronic Development of Functions of *se* and *sibi* in Latin», *SLE Digital Conference*, 26 Aug - 1 Sept.
- Petit, D. (1999): **Sue- en grec ancien: La famille du pronom réfléchi: Linguistique grecque et comparaison indo-européenne*, Louvain, Peeters.
- Pieroni, S. (2010): «Deixis and Anaphora», in Baldi, P. & Cuzzolin, P.L. (eds.), *New Perspectives on Historical Latin Syntax*, vol. 3, Berlin-New York, De Gruyter, pp. 389-501.
- Pinkster, H. (2015): *The Oxford Latin Syntax: Volume 1: The Simple Clause*, Oxford, Oxford University Press.
- Pisani, V. (1945): *Lezioni sulla preistoria e storia più antica del latino*, Milán, Vita e Pensiero.
- Pooth, R. (2018): *Transitivity Direction in Proto-Indo-European: Direct, Inverse, and Undirected Verb Forms*. Unpublished manuscript <https://www.academia.edu/37429692/Transitivity_Direction_in_Prot Indo-European_Direct_Inverse_and_Undirected_Verb_Forms> (30/05/2020).
- Pooth, R., Kerkhof, P. A., Kulikov, L. & Barðdal, J. (2019): «The Origin of Non-Canonical Case Marking of Subjects in Proto-Indo-European: Accusative, Ergative, or Semantic Alignment», *Indogermanische Forschungen* 123, pp. 245-263.
- Puddu, N. (2005): *Riflessivi e intensificatori: greco, latino e le altre lingue indoeuropee*, Pisa, ETS.
- Rumsey, A. (1987): «The Chimera of Proto-Indo-European Ergativity», *Lingua* 71, pp. 297-318.
- Schmidt, G. (1978): *Stammbildung und Flexion der indogermanischen Personalnomina*, Wiesbaden, Harrassowitz.
- Siewierska, A. (1998): «Passive-to-ergative versus inverse-to-ergative», in Siewierska, A. & Song, J. (eds.), *Case, Typology and Grammar*, Amsterdam, John Benjamins, pp. 229-246.
- Sihler, A. L. (1995): *New Comparative Grammar of Greek and Latin*, Oxford, Oxford University Press.
- Silverstein, M. (1976): «Hierarchy of Features and Ergativity», in Dixon, R. (ed.), *Grammatical Categories in Australian Languages*, New Jersey, Humanities Press, pp. 112-171.
- Sturtevant, E. H. (1939): «The Pronoun **so*, **sa*, **tod* and the Indo-Hittite Hypothesis», *Language* 15, pp. 11-19.

- Szemerényi, O. (1978): *Introducción a la lingüística comparativa* (transl. Álvarez, A.), Madrid, Gredos.
- Weiss, M. L. (2009): *Outline of the Historical and Comparative Grammar of Latin*, Ann Arbor-New York, Beech Stave Press.
- Wolfart, H. (1973): «Plains Cree: A Grammatical Study», *Transactions of the American Philosophical Society* 63 (5), pp. 1-90.
- Zamudio, R. (1989): «La forma pronominal latina *ipse*: su origen», *Emerita*, 57, pp. 119-127.
- Zúñiga, F. (2007): «From the Typology of Inversion to the Typology of Alignment», in Miestamo, M. & Wälchli, B. (eds.), *New Challenges in Typology: Broadening the Horizons and Redefining the Foundations*, Berlin-New York, De Gruyter, pp. 199-221.

Fecha de recepción de la primera versión del artículo: 21/09/2020

Fecha de aceptación: 10/11/2020

Fecha de recepción de la versión definitiva: 09/12/2020