



Comparing intelligibility between Occitan, Gallo-“Italic”, and Tuscan: Classification of Gallo-“Italic” empirically tested

Lissander Brasca

Bangor University, Bangor, Gwynedd LL57 2DG, Wales, UK

Received 6 November 2025; revised 13 January 2026; accepted in revised form 14 January 2026;

Abstract

Scholars unanimously describe Gallo-“Italic” as genealogically Gallo-Romance. Nonetheless, some refer to it as part of Italo-Romance. Some of the latter advocate their stance by claiming that heteronomy towards Italian/Tuscan would have made Gallo-“Italic” more similar to the bordering Italo-Romance than to the bordering Gallo-Romance varieties (‘critical convergence claim’). However, this claim, which is quantitative in nature, has never been supported by quantitative evidence. I have now tested it and will be presenting the results in this contribution. Based on the link between intelligibility level and linguistic distance demonstrated by empirical research, I argue that the critical convergence claim predicts that intelligibility is currently lower and decreases more abruptly across the Occitan–Gallo-“Italic” border than across the Tuscan–Gallo-“Italic” border. Therefore, through the SPIN method, I measured intelligibility between pairs of twelve geolects spoken in the Occitan-speaking Alps of France and Italy, in the Po valley and Romagna (Gallo-“Italic”), and in the Tuscan Apennines, and tested for statistical difference. Results falsify the critical convergence claim. They demonstrate that intelligibility is significantly lower and decreases more abruptly between Tuscan and Gallo-“Italic” than between Occitan and Gallo-“Italic”. These findings support a synchronological classification of Gallo-“Italic” as being still closer to the rest of Gallo-Romance than to Italo-Romance.

© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Gallo-Italic; Language classification; Gallo-Romance; Italo-Romance; Intelligibility; Linguistic distance

1. INTRODUCTION

Scholars deem the Massa-Senigallia line to be an important bundle of isoglosses in the Romance domain. It divides Western from Eastern Romance areas along the Apennine mountains ridge in Italy¹ (Trager, 1934; Bartoli, 1936; Von Wartburg, 1936, 1967; Lausberg, 1965; Pellegrini, 1975; Loporcaro, 2009).² Scholars also consider most Romance vari-

E-mail addresses: L.Brasca@bangor.ac.uk, brasca888@gmail.com

¹ In the present study, ‘Italy’ is used with the currently common extra-linguistic meaning of ‘territory of the Italian Republic’.

² Scholars agree on the geographical position of most of this bundle of isoglosses, although there is hesitation/debate concerning its extremities: older literature proposes “La Spezia–Rimini line” (e.g. von Wartburg, 1967); some suggest “Carrara–Fano line” (Pellegrini, 1992: 285); Hull (1982/2017: 9) proposes “Carrara–Pesaro line”. For the present aims, these fine-grained geographical differences are irrelevant.

<https://doi.org/10.1016/j.lingua.2026.104117>

0016-7037/© 2026 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

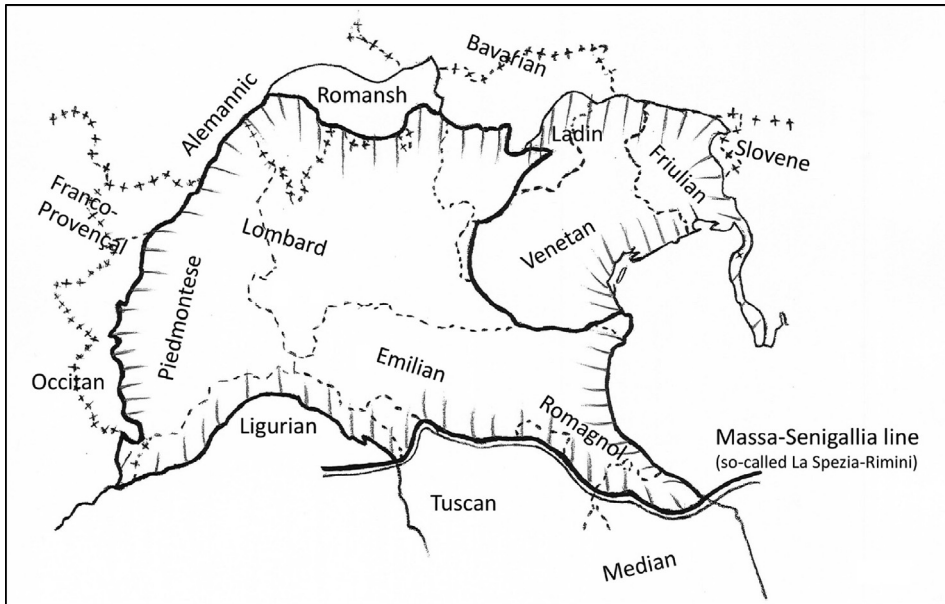


Fig. 1. The Gallo-“Italic” languages are spoken inside the thick line. The Cisalpine continuum (including Gallo-“Italic”) is inside the ‘fringed’ line. The Rhaeto-Cisalpine (or Padanian) continuum (Hull, 1982/2017) includes Cisalpine plus Romansh. Small crosses: state borders; dashed lines: regional administrative borders. The Töitschu, Mòcheno, Cimbrian, and Resian languages are not represented in this map. Find them in Moseley (2010), also online: <https://unesdoc.unesco.org/ark:/48223/pf0000187026> <https://unesdoc.unesco.org/ark:/48223/pf0000187026/PDF/187026eng.pdf.multi>.

eties spoken between the Massa-Senigallia line and the Alps to the north as a homogeneous group known as Gallo-“Italic”³ (Fig. 1; for the sub-grouping of Gallo-“Italic”, see e.g. Devoto and Giacomelli, 1972).

While scholars agree about the extent and the internal articulation of Gallo-“Italic”, an issue arises when dealing with its ascription to wider Romance sub-groups: some scholars group Gallo-“Italic” with Gallo-Romance (henceforth *pro-Gallo-* scholars), while others group it with Italo-Romance (henceforth *pro-Italo-* scholars). Hull (1982/2017), Schmid (1956), and Bec (1970–1971) hold the *pro-Gallo-* stance, which is also supported by hierarchical clusterings resulting from dialectometric studies (Goebel, 2008; Tamburelli and Brasca, 2018). Some representative works of the *pro-Italo-* stance include von Wartburg (1967), Lausberg (1965), Hall (1976), Pellegrini (1975), and Loporcaro (2009).

Brasca (2020, 2021, 2023, and 2026)⁴ points out a discrepancy in the *pro-Italo-* works: despite referring to Gallo-“Italic” as Italo-Romance, they describe it as structurally Gallo-Romance. This is inconsistent, since ‘Gallo-Romance’ and ‘Italo-Romance’ are universally used as mutually exclusive labels of the same nomenclature. Particularly, they normally name two cousin taxa in the family tree model: Gallo-Romance is a Western Romance daughter, while Italo-Romance is an Eastern Romance daughter.

Brasca (2020, 2021 and 2023) showed that such a discrepancy is the consequence of epistemological inconsistencies present in the *pro-Italo-* literature, the main two of which are as follows (a. and b.):

- a. *An inconsistent mix of classificatory criteria.* Most *pro-Italo-* scholars (the ‘mixed criterion tradition’) inconsistently mix the ontologically and conceptually distinct *Abstand* and *Ausbau* criteria (see Kloss, 1967) in a unidimensional (Brasca, 2020, 2021) “classification”. In their studies (1.) they describe, compare, and group the vast majority of the geolects considered, based on the *isoglosses that these geolects share* (i.e. *Abstand*); (2.) they describe Gallo-“Italic” as sharing all the isoglosses that they consider critical in distinguishing Western Romance, and particularly Gallo-Romance, from Italo-Romance (von Wartburg, 1967; Lausberg, 1965; Pellegrini, 1975; Hall, 1976); but (3.) some sociological (which is an aspect of *Ausbau*) or geopolitical arguments are eventually presented in these works, determining the ascription of Gallo-“Italic” to the Italo-Romance taxon instead of to the expected

³ All Gallo-“Italic” languages—Emilian, Ligurian, Lombard, Piedmontese, Romagnol—are spoken in the north of the Italian state and listed in the UNESCO Atlas of the World’s Languages in Danger (Moseley, 2010). I write Gallo-“Italic” with double inverted commas because the Italic linguistic profile of this group is questioned in this contribution.

⁴ More argumentations and examples concerning the present Introduction are in Brasca (2023: Ch. 1, 2, and 3).

Table 1
Bidimensional (partial) classification of Romance varieties.

	<i>Gallo-Romance</i>	<i>Italo-Romance</i>
<i>heteronomous towards Italian</i>	Gallo-“Italic” , Occitan of Piedmont, etc.	Neapolitan, Sicilian, etc.
<i>heteronomous towards French</i>	Occitan of France, etc.	Corsican

Table 2
Unidimensional (partial) classification of Romance varieties.

<i>Gallo-Romance</i>	<i>Italo-Romance</i>
Occitan of France and Piedmont, etc.	Gallo-“Italic” , Neapolitan, Sicilian, etc.

Gallo-Romance taxon (detail in Brasca, 2020, 2021; this is called the “taxon problem” in Brasca, 2023). Such a “classification” has to be rejected since the *Abstand* and the *Ausbau* criteria ground two independent classifications: an *Abstand*-based one, grouping Gallo-“Italic” with Gallo-Romance based on shared isoglosses, and an *Ausbau*-based one, grouping Gallo-“Italic” with the varieties currently heteronomous towards standard Italian. These classifications can be combined in a bidimensional classification (Table 1) but not mixed into a unidimensional one (Table 2).

In Table 1, the independence of the horizontal (*Abstand*) and vertical (*Ausbau*) dimensions is correctly respected and recognizable. On the contrary, Table 2 does not distinguish them. Therefore, it cannot fulfil the core scientific function of classifications, which is that of *informing* about the traits shown and shared by the classified items (see Simpson, 1945; also Devitt, 2008; also Bailey, 1994; more detail in Brasca, 2023, 2020, 2021).

Such a mix of *Abstand* and *Ausbau* criteria also explains what was called the “category problem”⁵ in Brasca (2023), which entails the inconsistent classification of the (not only) Gallo-“Italic” *languages* as “Italian *dialects*” (detail in Brasca, 2023; see also Brasca, 2026).

- b. A claimed ‘critical linguistic convergence’ not supported by empirical evidence. In Brasca (2023) it was also shown that some scholars presume to support their pro-Italo- stance by claiming that Gallo-“Italic” would have become more similar to its second-cousin Italo-Romance (also ‘Peninsular’) varieties than to its sibling Gallo-Romance varieties through contact with the Tuscan/Italian literary language (henceforth, ‘critical convergence’⁶ claim’, also ‘CCC’). It was argued, however, that even assuming this critical convergence scenario as a working hypothesis, these scholars propose a *synchronological*⁷ classification which, as such, should be expressed by a different nomenclature from a *genealogical* and an *Ausbau*-based one. Therefore, they should avoid the labels “Italo-Romance” (typical of genealogical classifications, like ‘Gallo-Romance’, ‘Ibero-Romance’, and ‘Daco-Romance’) and “Italian” (factually—although problematically—used in *Ausbau*-based analyses) (detail in Brasca, 2023). Besides this, and more importantly, although their claim is quantitative in nature, these scholars do not provide quantitative evidence in favour of it. In the present study, I will test the critical convergence claim⁸ through a quantitative method.

2. LITERATURE REVIEW

Two examples of pro-Italo- critical convergence claimers. Von Wartburg’s (1967) is the only pro-Italo- ‘classical’ study⁹ among those that I have looked at that classifies Gallo-“Italic” formally based on exclusively *Abstand* criteria,

⁵ The “taxon problem” and the “category problem” are the two forms of the “ontological problem” (Brasca, 2023).

⁶ On the model of physics and other disciplines’ concept of “critical mass” (see collinsdictionary.com), I intend ‘critical convergence’ as the minimum amount required for a linguistic convergence to make Gallo-“Italic” more similar to its second-cousin Italo-Romance geolects than to its sibling Gallo-Romance geolects.

⁷ That is, based on current linguistic similarity (Brasca, 2023). The synchronological criterion grounds a classification that is conceptually distinct and partly independent from a genealogical one. In fact, current similarity can be due not only to a shared genealogical profile (a common ancestor language), but to effects of contact after the split (and to possible unrelated but convergent changes), too (detail in Brasca, 2023). Therefore, a synchronological grouping could be different from a genealogical one.

⁸ What in this contribution I call ‘critical convergence claim’ is called ‘post-Wartburg hypothesis’ in Brasca (2023).

⁹ By ‘classical studies’ I mean those written in the 19th and 20th centuries, and whose proposals of classification for Romance languages are normally considered authoritative in more recent studies (see for instance Posner, 1996 and Loporcaro, 2009).

without apparent recourse to *Ausbau* arguments. In von Wartburg (1967), Gallo-“Italic” seems to be considered Western Romance and Gallo-Romance in genealogical terms, but would have then *become* “Italian” (1967: compare maps 9 and 10) due to sharing with Tuscan the Longobard superstrate¹⁰ and to a linguistic convergence towards Tuscan starting during the Middle Ages:

1. “... as Longobards have *avoided the constitution* of a real linguistic frontier between Florence [in Tuscany] and Bologna [in the Gallo-“Italic” Cisalpine], the forming linguistic domain had to set its limits elsewhere. These limits were set along the Alps, but in such a way that almost all the upper valleys of the rivers that go down the Western Alps towards the Po valley *have maintained* their Gallo-Romance nature. ... What happened on the western edge of the Po plain *also happened on its northern edge*” (1967: 130–131 [emphasis added]).
2. “...from the 13th century, when the linguistic supremacy of Tuscan became unopposed, the *waves of linguistic changes progress from South to North*. They have taken away some pieces of that old frontier [La Spezia-Rimini line] and *have carried them towards some more northern areas*. So, the whole of Northern Italy lost the -s; the line between the retention and the loss of this consonant passes today north of the plain, and between the old and the new isoglosses only a few traces of the old situation are still preserved.” (von Wartburg, 1967: 130 [emphasis added]).

A more recent instance of CCC can be found in Loporcaro (2009). Besides recognizing the intervention of a “socio-political” (*Ausbau*) criterion in his proposal of pro-Italo- classification (see also “storico-culturale” and “storico-politica”, Loporcaro, 2009: 11–13, 70), Loporcaro also refers to the “linguistic repercussions” of heteronomy (for “autonomy [vs.] heteronomy” see Trudgill, 1992),¹¹ which would constitute an additional *Abstand* reason for classifying Piedmontese (and supposedly all of Gallo-“Italic”) as an “Italian dialect” within Italo-Romance:

3. “Autonomy/heteronomy is a socio-political fact, but it has *linguistic repercussions*, both on the structural and on the classificatory level. *On the structural level*, a heteronomy relationship opens the way to the progressive *approach of the (heteronomous) dialect to [Italian]*. . .” (Loporcaro, 2009: 8 [emphasis added]).
4. “[The geolect] of Turin has Italian as its roof-language¹² since the sixteenth century, and the processes of standardization push it imperceptibly *towards a structural homologation to Italian*.” (Loporcaro, 2009: 12 [emphasis added]).

Previous pro-Gallo- synchronological quantitative studies. A synchronological classification of the geolects of Italy obtained by quantitative methods is proposed by the Dialectometric School of Salzburg University¹³ (DM-S) (Goebel and Winterleitner, 1971; Goebel, 2008; Goebel et al., 2019 dialektkarten.ch). DM-S analyses the data provided by the linguistic atlas AIS, *Sprach- und Sachatlases Italiens und der Südschweiz* (Jaberg and Jud, 1928–1940), which were collected a century ago, between 1919 and 1928. In particular, the data concern 380 of the Romance AIS data points plus standard French and standard Italian as “fictitious” data points added by the DM-S researchers. In Brasca (2023), some currently relevant comparisons were made between values obtained by DM-S. They are summarized as follows (1. to 3.).

1. A DM-S *isogloss map* represents the “Relative [linguistic] Distance Values” (RDV) between the AIS geolects. A comparison between *isogloss maps* values (taken from dialektkarten.ch, which concern all AIS data points) indicates that the linguistic distance is generally greater between Tuscan and Emilian/Romagnol (Gallo-“Italic”) than between Piedmontese (Gallo-“Italic”) and the uncontested¹⁴ Gallo-Romance (Franco-Provençal and Occitan) of Piedmont.

¹⁰ Brasca (2023: Section 3.2.6) contests the validity of this argument.

¹¹ “It is usual in sociolinguistics to say that. . . standard varieties, which are enshrined in grammar books and dictionaries and have bodies of literature written in them, are *autonomous*: they have, as it were, an independent existence. The nonstandard local dialects, on the other hand, are *heteronomous* or dependent. . . A reasonable definition of an *Ausbau* language is thus that it consists of an autonomous standard variety together with all the nonstandard varieties from the dialect continuum which are heteronomous with respect to it. . . . Autonomy can be lost or acquired. . . . Autonomy and heteronomy can also be disputed.” (Trudgill, 1992: 169 [italic original]).

¹² See in Brasca (2021, 2023) arguments rejecting this definition of Italian as the “roof-language” of Turinese.

¹³ In the Hans Goebel works that I am aware of, the pro-Gallo- genealogical and/or synchronological nature of Gallo-“Italic” is not explicitly stated. In my opinion, however, it surfaces in his analyses of data, as shown here. Moreover, it seems implicitly acknowledged in some statements of his (e.g. Goebel, 2008: 44, 58). However, the latter cohabit with at least as many statements in which Gallo-“Italic” is referred to as not being part of Gallo-Romance (e.g. Goebel, 2008: 26, 43, 45, 48).

¹⁴ They were considered ‘uncontested’ based on Allasino et al. (2007); see Brasca (2023: 120).

2. A DM-S similarity analysis compares pairs of data points by taking a single data point as a reference point. The results of similarity analyses with Milanese—a major Gallo-“Italic” “innovating” geolect with “Italianizing” tendencies (Sanga, 1997: 255)—as a reference point are far from indicating that Milanese got predominantly more similar to the Tuscan geolects bordering the Gallo-“Italic” domain than to the Gallo-Romance geolects bordering the Gallo-“Italic” domain.¹⁵ Therefore, this should be true for most Gallo-“Italic” geolects, which show fewer “Italianizing” tendencies than Milanese.
3. In DM-S *hierarchical clustering analyses* based on phonetic and lexical data, either using *Complete Linkage* and *Joe Ward Jr.’s* algorithms, the first clustering results as follows:
 - one cluster puts together the entire Cisalpine with standard French and with the uncontested Gallo-Romance and Rhaeto-Romance AIS geolects;
 - the other cluster puts together the rest of the geolects, namely those south of the Massa-Senigallia line, those of Sardinia, and standard Italian.

In summary, none of the comparisons made between the Salzburger values lend support to the CCC. As far as I am aware, no other dialectometric studies results contradict those of the Salzburger School.

3. THE PRESENT STUDY

3.1. Theoretical background

I argue that the critical convergence claim can be tested empirically through functional intelligibility testing, namely by measuring and comparing actual intelligibility between native speakers. Intelligibility is “the extent to which the native speaker understands the intended message” (Derwing and Munro, 1997: 2). Already in 1959, Hans Wolff wrote that “data on intelligibility between two linguistic media has been used (*a.*) as a convenient criterion for distinguishing the concepts of language and dialect; and (*b.*) for determining ‘dialect distance’, i.e., the degree of divergence between dialects of the same language or between closely related languages” (Wolff, 1959: 34). More recent quantitative research shows that differences in intelligibility scores are directly linked to differences in linguistic distance (*Abstand*) between the speaker’s and the listener’s geolects (Gooskens, 2007; Munro, Derwing and Morton, 2006; Tang and van Heuven, 2009):

1. “Intelligibility can to a large extent be predicted by phonetic distance. . .” (Gooskens, 2007: 445).
2. “We argue that mutual intelligibility testing is an adequate way to determine how different two languages or language varieties are. . . [T]he degree of mutual intelligibility. . . can be predicted from objective measures (lexical similarity and phonological correspondence). . .” (Tang and van Heuven, 2009: 709).

Crucially, in the quoted studies, the measures of objective linguistic distance—which were correlated to intelligibility measures—were obtained through *dialectometric* methods, which avoid the classificatory shortcomings deriving from arbitrary trait selection (see Tamburelli and Brasca, 2018). Particularly, in Gooskens (2007), Levenshtein distance was used, which “has proved [to be] a useful way of measuring [linguistic] distances between dialects and closely related languages” (Gooskens, 2007: 446; see Gooskens and Heeringa, 2004), while in Tang and van Heuven (2009) “two objective measures of linguistic distance, viz. a lexical similarity index (LSI. . .) and a phonological correspondence index (PCI). . .” were used (Tang and van Heuven, 2009: 724). These two indexes were elaborated on the basis of “an impressive body of digital resources” established by Chinese linguists (Tang and van Heuven, 2009: 711).

In any case, the correlation between intelligibility and linguistic distance resulting from the above studies provides empirical ground to what has been an axiomatic assumption at the basis of the entire comparative research since its beginning, and at the basis of the use of the language family tree as a descriptive model. Indeed, in the comparative literature, when two dialects of the same language develop reciprocally independent linguistic changes, due, for example, to emigration or geopolitical separation of the respective sub-communities, they are understood to lose mutual intelligibility. Moreover, “if the divergence gets so wide as to hinder the mutual *intelligibility*. . . we can then say that each sub-community has become an independent linguistic community endowed with its own language” (Shukla and Connor-Linton, 2008: 63–64 [my translation to English from Italian, emphasis added]). Scholars call these newly separated lan-

¹⁵ Besides, one should consider that, over the centuries, the bordering Tuscan geolects were in turn Cisalpinised, even more than the other Tuscan geolects (see e.g. Rohlf, 1949–1966; Hull, 1982/2017).

guages, stemming from the same immediate ancestor language, “sister” languages (Campbell, 1998/2006; Shukla and Connor-Linton, 2008). In historical linguistics, the link between linguistic distance and intelligibility is therefore assumed when defining the very concepts of language vs. dialect. We can see this summarised in some definitions in Campbell (1998/2006):

3. “*Mutual intelligibility*: when speakers of different linguistic entities can understand one another. This is the principal criterion for distinguishing dialects of a single language from distinct languages” (Campbell, 1998/2006: 217).
4. “*Dialect* means only a variety (regional or social) of a language, which is mutually intelligible with other dialects of the same language. ‘Dialect’ is not used in historical linguistics to mean a little-known (‘exotic’) or minority language,^[16] and it is no longer used to refer to a daughter language of a language family, though the word has sometimes been used in these senses.¹⁷
Language means any distinct linguistic entity (variety) which is mutually unintelligible with other such entities” (Campbell, 1998/2006: 186).

3.2. Aims of the present study

Based on the above considerations, I assume that, before the purported critical convergence, Gallo-“Italic” was more intelligible to its bordering Occitan (sibling) geolects than to its bordering Tuscan (second-cousin) geolects. Accordingly, I argue that a possibly critical linguistic convergence would have a parallel communicative counterpart. Particularly, it would predict a reverse of the original intelligibility situations across the two borders at issue, namely, it would predict a currently higher level of intelligibility across the Tuscan–Gallo-“Italic” border than across the Occitan–Gallo-“Italic” border. In fact, measures of intelligibility collected from *monolingual* speakers of, say, von Wartburg’s time (first half of the 20th century) would have allowed testing the critical convergence claim. However, in the past decades, and as far as I am aware, scholarly attempts were not made in this direction as regards the geolects under study. Pellegrini (1975) generically refers to the well-known lack of mutual intelligibility between the geolects of Italy:

1. “It is well known that the mutual understanding between a Northern Italian and a Southern Italian is almost always impossible when both express themselves in local and archaic linguistic means which are not influenced by the Italian koine. I could cite many testimonies of this.” (Pellegrini, 1975: 64–65; see also 1970: 222).

However, no systematic intelligibility studies followed this statement, nor comparisons with intelligibility across the Occitan–Gallo-“Italic” border. Some early attempts at measuring intelligibility are found in Voegelin and Harris (1951) and Wolff (1959), but they—like the subsequent studies—do not involve the geolects under study.

In the present contribution, I will compensate for this lack. I will (measure and) compare intelligibility between pairs of geolects spoken along two chains of localities running from the Occitan-speaking French Alps to the Tuscan-speaking Apennines, passing through the Gallo-“Italic”-speaking area (Po valley and Romagna), thus testing the critical convergence claim.

¹⁶ See in Brasca (2023) examples and criticisms of the polysemic use of ‘language’ and ‘dialect’ in the literature, specifically—but not only—concerning the classification of Gallo-“Italic”. In a nutshell, an *Abstand*-based meaning (i.e. based on isoglosses and consequent intelligibility) and an *Ausbau*-based (i.e. socio-political) meaning of the ‘language/dialect’ pair are present in the literature. In this latter case, ‘dialects’ are meant as the substandard varieties that are heteronomous towards a standardized dominant ‘language’. As presented in Section 1, the ontologically distinct *Abstand* and *Ausbau* criteria ground two mutually independent types of classification. Their conceptual distinction should be expressed through distinct terminologies. In this respect, two issues are found in the literature. The first one concerns polysemy per se. In fact, in contrast to what most scholars tend to do (see examples in Posner, 1996), it would be preferable that they do not refer to the items classified in these different types of classifications by the same terms *languages* and *dialects*, but rather by different terminologies. The second—and more serious—issue is found (at least?) when varieties of Italy are concerned. Indeed, *Abstand* and *Ausbau* criteria are often inconsistently *mixed* in (“unidimensional”, see Section 1) classifications. This inconsistent mix affects not only the grouping of the varieties of Italy (the “taxon problem”, briefly seen in Section 1), but also their categorization as ‘languages’ or ‘dialects’ (the “category problem”). As presented above, the “taxon problem” and the “category problem” are the two forms of the “ontological problem” (Brasca, 2023).

¹⁷ Note the consequent inconsistency of defining two “daughter *languages*”—i.e. two reciprocally sister *languages*—as “*dialects*”, too. Actually, they can consistently be referred to as *former* dialects of the same immediate ancestor language, but—by definition—not as *current* dialects of the same language.

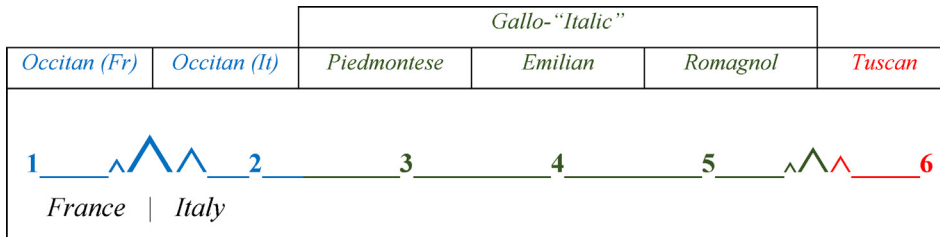


Fig. 2. Sequence of six localities and their respective linguistic areas included between the Occitan Alps of France and the Tuscan Apennines. Blue triangles: Alpine geographical boundary; brown/red triangles: Apennine geographical boundary. Their different graphic size just intends to mirror the difference of height between the two real mountain ranges, which is irrelevant for this study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3. Predictions made by the critical convergence claim (CCC)

Let us consider the geolects spoken in six localities situated on a line running from the Occitan Alps of France up to the Tuscan Apennines, passing through the Gallo-“Italic” area, as in Fig. 2. Moreover, let us concisely reiterate the CCC: ‘Gallo-“Italic” was linguistically more similar to the bordering Gallo-Romance sibling geolects but then became more similar to the bordering Italo-Romance second-cousin geolects’. Based on the link between objective linguistic similarity and intelligibility presented above, the CCC makes three testable predictions:

PREDICTION 1. Today, intelligibility is higher between geolects 5 and 6 (Romagnol and Tuscan) than between geolects 2 and 3 (Occitan of Italy and Piedmontese).

However, if the test results indicated that Prediction 1 is not borne out, this fact could be potentially explained by the following counter-conjecture, which makes those results logically reconcilable with the CCC:

- a. Occitan of Italy became linguistically closer to Piedmontese (assumed in turn to be linguistically closer to Italo-Romance than to Gallo-Romance) than to Occitan of France as an effect of its heteronomy towards Piedmontese and Italian. Possibly, Occitan of Italy could have become closer to Piedmontese than Romagnol to Tuscan.

However, if the counter-conjecture *a* were borne out, such supposed linguistic convergence of Occitan of Italy towards Piedmontese should necessarily correspond to a proportioned linguistic divergence of Occitan of Italy from Occitan of France, if one wants the counter-conjecture *a* to be reconcilable with both a pro-Italo-¹⁸ synchronological grouping for Gallo-“Italic” on the one hand, and a pro-Gallo- one for Occitan of France on the other.¹⁹ Consequently, this conjectural, modified geolinguistic picture would make a prediction that is a consistently modified version of Prediction 1. I will define it as ‘Prediction 1-a’ and will test it, too:

PREDICTION 1-a. Today, intelligibility is higher between geolects 5 and 6 (Romagnol and Tuscan) than between geolects 1 and 2 (Occitan of France and Occitan of Italy).

However, the CCC (e.g. von Wartburg, 1967) also understands

- i. that during the presumed shift, the entire Po valley plus Romagna area *remained linguistically united*, preserving the individuality of Gallo-“Italic” as a synchronological taxon; and
- ii. that *all* of the geolects of Po valley and Romagna got further from Occitan and closer to Tuscan, forming a new wider “Italian”/Italo-Romance *continuum* (i.e. a pro-Italo- synchronological taxon) from the Alps to Sicily (see e.g. von Wartburg, 1967: Map 10).

Therefore, paragraphs *i* and *ii* make a second prediction:

¹⁸ Pending a nomenclature exclusively dedicated to a (desirable) synchronological classification of the entire Romance domain, in the present study, the expression ‘pro-Italo- synchronological grouping for X’ will be used with the meaning of ‘synchronological classification that groups geolect X with the members of the genealogically-identified Italo-Romance group’. The seemingly alternative expression ‘Italo-Romance synchronological grouping’ would be inconsistent, since—as we have seen—a nomenclature that is normally used for a genealogical classification (Italo-Romance, Gallo-Romance, etc.) should not be used for a synchronological classification. The same applies to the expression ‘pro-Gallo- synchronological grouping’.

¹⁹ In this (conjectural) case, I deduce, Occitan of Italy would reasonably deserve a new synchronological classification.

PREDICTION 2. Today, intelligibility *decreases more abruptly* across geolects 1, 2, and 3 in the Alps (namely along the portion of the chain running from Occitan of France, through Occitan of Italy to Piedmontese) than across geolects 4, 5, and 6 in the Apennines (namely along the portion of the chain running from Emilian, through Romagnol to Tuscan).

Testing Predictions 1 (and/or 1-a) and 2 would be sufficient to test the CCC. However, in order to be as resolute as possible, I tested what is normally stated in the literature (e.g., in von Wartburg, 1967) about the unity of Gallo-“Italic” (assumed in point *i.* above). I considered that, if Gallo-“Italic” remained linguistically united as a synchronological taxon and became more similar to Tuscan than to Occitan, this makes the following prediction:

PREDICTION 3. Today, no line can be found in the Po valley and Romagna (i.e. between the Alps and the Apennines) across which intelligibility

- a. is lower
 - and
 - b. decreases more abruptly
- than across the western Alpine linguistic border.

One should note that, if the previous Predictions 1/1-a and 2 were not borne out, testing whether the Po valley and Romagna remained linguistically united would equate to testing a re-formulated version of Prediction 3, which I will call ‘Prediction 3*’:

PREDICTION 3*. Today, no line can be found in the Po valley and Romagna across which intelligibility

- a. is lower
 - and
 - b. decreases more abruptly
- than across the *Apennines*.

3.3.1. The experimental and the null hypotheses of the present study

I have argued that scholars have not provided quantitative evidence supporting the CCC. Therefore, in the current study, it is appropriate to consider the pro-Gallo- synchronological profile of Gallo-“Italic” as the traditional theory that is going to be challenged by a possible ‘new’ one. Therefore, it will correspond to the *null hypothesis*. Conversely, the CCC will correspond to what is usually considered to be the *experimental hypothesis*, namely the ‘new’ theory (‘new’ despite being formally elaborated by me on the basis of a claim dating back, at the latest, to Wartburg, 1950/1967)²⁰ that needs to be supported by evidence (i.e. that has ‘the burden of proof’).

3.4. Research questions

Testing the predictions above will answer the following research questions concerning intelligibility:

- 1) In a chain of six localities, running from the French Alps to the Tuscan Apennines, passing through the Po valley and Romagna, which is the pair of localities between whose geolects intelligibility is the lowest?
- 2) Which is the sequence of the three adjacent geolects along which intelligibility lowers most abruptly?

Based on the link between intelligibility and objective linguistic distance (Section 3.1), answering RQs 1 and 2 will correspond to answering the following research question concerning linguistic distance:

Which is the synchronological grouping of present-day Gallo-“Italic”: pro-Gallo- or pro-Italo-?

3.4.1. Topicality of testing the critical convergence claim

Testing the CCC, despite dating back at the latest to Wartburg (1950),²¹ is relevant for present-day and future research. Indeed, von Wartburg’s proposal of dividing the Romance domain in a western and an eastern part, once along the Apennine mountains ridge, and today along the southern slopes of the Alps, is normally accepted and re-proposed—with the same formal inconsistencies—in (much of) more recent scholarly literature focusing on language

²⁰ One should remember that the CC claimers did not propose the CC scenario as a testable (and tested) *hypothesis*, but rather *claimed* it without supporting it by quantitative evidence. Rather, it is I who wants to test the CCC, hence I express it in the form of a testable hypothesis.

²¹ In the present study, I cite the 1967 French translation of the original 1950 German version.

classification (e.g. Varvaro, 2001; Loporcaro, 2009). Consequently, most of the recent studies in linguistics and philology, including those not specifically focusing on language classification, refer to Gallo-“Italic” as an “Italo-Romance” sub-group, often as “Northern Italo-Romance” (e.g. Cardinaletti and Giusti, 2025; Bernini, 2025; GSPL Conference, 2020; Fanciullo, 2018; Garassino et al., 2017; Valentini, 2012; Filipponio et al., 2010).

4. METHODOLOGY

Over the past few years, literature dedicated to speech intelligibility increased and new methods were proposed. The scope and the varieties analysed in these studies are various: intelligibility of L2 speech (e.g. Munro and Derwing, 1995; Munro et al., 2006); intelligibility of Spanish and Portuguese (Jensen, 1989); Frisian, Afrikaans, and Dutch (van Bezooijen and Gooskens, 2005); Scandinavian languages (Gooskens, 2007; Moberg et al., 2007; Gooskens et al., 2008); Chinese geolects (Tang and van Heuven, 2009); German, Low German, and Dutch (Gooskens et al., 2011); Lombard and Italian (Tamburelli, 2014); and Bavarian and Standard German (Leonardi, 2016).

In these studies, the investigation of intelligibility is motivated by various interests: language planning or policy, language contact phenomena, L2 language teaching/learning, audiology, foreign language testing, speech technology, the relationship among accentedness, perceived comprehensibility and intelligibility (e.g. Munro and Derwing, 1995; Munro et al., 2006).

Several methods have been used to measure the intelligibility of either written or spoken languages: opinion testing (asking the participants “how well *they think* they understand” the target geolect, see Gooskens, 2013), functional testing (measuring how well the participants “*actually* understand” the target geolect, see Gooskens, 2013: 5), among which are transcription, picture-pointing tasks, summarization, multiple-choice questions, true/false questions, translation of single words, and Speech Perception in Noise (SPIN) test.

For the current aims, I considered the method used in Tang and van Heuven (2009) (i.e., SPIN test with high-predictability sentences, Kalikow et al., 1977) as being particularly suitable, since in that study intelligibility testing had similar intents to the current ones. Particularly, the following advantages were considered:

1. The SPIN test employs short sentences. These reduce the risk of the saturation of short-term memory (see Gooskens et al., 2008; Gooskens, 2013) thereby minimising nuisance variability due to difference in participants’ memory skills;
2. Functional intelligibility measures at sentence level (SPIN test) better reflect comparativists/dialectologists’ classifications
 - a. than functional intelligibility measures at word level, and
 - b. than opinion scores (Tang and van Heuven, 2009);
3. The results of a functional test at sentence level have greater ecological validity than those at word level. The latter methods employ isolated words out of context, which are the exception in natural language use (Tang and van Heuven, 2009);
4. The SPIN test “is a more realistic task than, for instance, semantically unpredictable sentences... in which all the words [of the sentence] are chosen to be semantically unpredictable.” (Valentini-Botinhao and Wester, 2014: 2063);
5. The large number of sentence meanings proposed in Kalikow et al. (1977) would allow for the stimuli of the current study to be as consistent as possible. Indeed, it allowed me to select a final set of 28 homogeneously long sentences for each target geolect, ending with a mono- or disyllabic noun (see Section 4.3.2).

Moreover, functional testing was preferred to opinion testing because this latter does not allow the researcher to assess whether the participants’ judgments are influenced by non-linguistic factors (e.g. attitudes, geographical knowledge of the variety in question; see Beijering et al., 2008).

4.1. Demographic and attitude variables

It has been claimed that, besides linguistic (structural and lexical) distance between the speaker’s and the listener’s variety (see Gooskens, 2007; Moberg et al., 2007), the following other variables could also affect intelligibility rate:

FOREIGN (AND LATIN) LANGUAGE COMPETENCE: “multilinguals have an advantage... for comprehension of a language” (Gooskens and Heeringa, 2014: 249; see Cenoz et al., 2001; Council of Europe, 2001); moreover, any foreign language competence (e.g. English, Spanish, Latin) might facilitate the understanding of words of the target geolect that are so-called²² “cognates” of words of the known foreign language (Gooskens, 2007: 446, 461, 464).

²² In historical linguistics, “cognates” has a different meaning than in the mutual intelligibility literature cited here. In the latter, it seems to mean ‘co-etymologic words, which share some phonetic material and semantic traits, regardless of whether they were both *inherited* from a common ancestor language, or *borrowed* by one or both of the two languages’.

CONTACT AND EXPOSURE: Previous literature argues that contact with the target area and exposure to the target language may improve intelligibility on the test (e.g., Gooskens, 2006; Moberg et al., 2007).

ATTITUDE. It has also been claimed that positive attitudes towards target languages and towards their speakers may encourage participants to try to understand the stimuli sentences, while negative attitudes should discourage them from making such an effort (Gooskens, 2006), causing false negatives in the intelligibility test (Wolff, 1959; see also Svanes, 1988). However, in experimental settings, the correlation between the attitude and intelligibility scores has often resulted as weak—hence difficult to demonstrate—(Gooskens, 2006; van Bezooijen and Gooskens, 2005; Schüppert and Gooskens, 2011; Tamburelli, 2014) or as non-significant (Leonardi, 2016).

In any case, in order to check the potential effect of attitudes and contact on the intelligibility rate, following previous research, I collected data from the participants via a questionnaire (Section 4.3.1), eliciting information about their attitudes towards the target varieties and geographic areas, about the amount of exposure to the target varieties and of familiarity with the target areas, and about their foreign and Latin language competence.

GENDER OF VOICE. There is some evidence that female voices tend to be more intelligible than male voices (Bradlow et al., 1996). In order to control for the gender of voice, I planned that for each geolect investigated, a male and a female voice would produce the entire list of stimuli each, and that each participant would be presented with an equal number of stimuli produced by the male and the female voice.

PARTICIPANT GENDER. There is also some evidence that, in conditions of language shift, females tend to abandon the endangered language before males (Labov, 1990; see an overview in Romaine, 2008). In order to control for participants' gender, I planned that an equal number of male and female participants would be selected, and that the different acoustic lists obtained through the Latin square design (Section 4.3.2) would be evenly distributed between male and female participants.

4.2. Design of the study

For intelligibility functional testing, I used a translated version of the SPIN test (Kalikow et al., 1977; Tang and van Heuven, 2009). I measured and compared intelligibility between pairs of geolects spoken in localities situated along two chains (A and B) of six localities each (see Fig. 3), both running from the Occitan-speaking Alps of France, to the Tuscan-speaking Apennines, passing through the Gallo-“Italic”-speaking area. Intelligibility was investigated independently along each chain.

STIMULI VS. FILLERS. Intelligibility data were collected by playing 28 recorded sentences (items) to each participant. 14 were spoken in the geolect that, for the sake of simplicity, I will define as ‘at west’ (i.e. spoken in the adjacent locality in

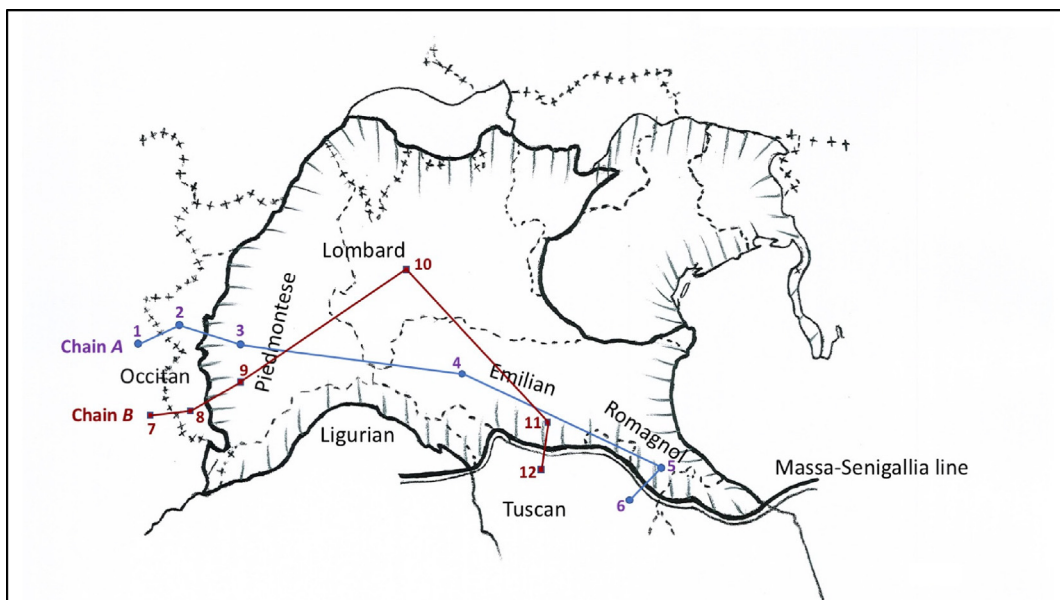


Fig. 3. Dots: Chain A data points; squares: Chain B data points.

the chain that is to the west of the participant's locality²³), while the other 14 items were spoken in the geolect 'at east'. However, looking at Fig. 3, it appears that:

1. Tuscan-6²⁴ data point was situated at the 'eastern' extremity of the chain,²⁵ therefore it had only one adjacent geolect (i.e. 'at west');
2. for Romagnol-5 participants, as speakers of Italian, too, only intelligibility scores on geolect 'at west' had experimental relevance, not scores on geolect 'at east' (Tuscan-6).²⁶

For these cases, to grant homogeneous experimental conditions for all participants (see Keating and Jegerski, 2015), filler sentences were used. The items spoken in the second nearest geolect in the chain were employed as fillers, as shown in Table 5. Moreover,

3. for Occitan-2 participants, intelligibility on geolect 'at east' (i.e. Piedmontese-3) had no experimental relevance. This is because Occitan-speaking dwellers of Piedmontese Alps (Occitan-2) are often trilingual (Occitan-Piedmontese-Italian) (see Sibille, 2012);
4. therefore, in data analysis, the unusable Occitan-2 (trilingual) participants' intelligibility scores on the Piedmontese-3 geolect were replaced with the Occitan-1 (France) participants' intelligibility scores on Piedmontese-3 geolect.

REASONS FOR STUDYING TWO CHAINS. I studied two independent chains of data points (*A* and *B*) in order to obtain a broader picture of intelligibility than what one single chain would have provided. I assumed that a possible parallelism in the sequence of the scores measured along two chains, in addition to the scholarly qualitative descriptions of the linguistic similarities among and within the linguistic areas crossed by the two chains, would have suggested that similar sequences of scores could be plausibly found in other similar chains of localities, having a similar direction and a similar distribution among the linguistic areas of interest. Moreover, investigating two chains made it possible to measure both Piedmontese-Emilian intelligibility (in Chain *A*) and Piedmontese-Lombard (and consequently also Lombard-Emilian) intelligibility (in Chain *B*) in both directions, improving the informativeness of the study. I included Lombard and Romagnol in only one chain each. This choice allowed studying all the possible combinations of bordering stimulus language – participant language with the minimum number of data points.

REASONS FOR INTERSECTING THE CHAINS. The two chains intersect in two points (Fig. 3). I chose this design in order to avoid what could be called the 'two-belts-continua interpretation'. Two non-intersecting chains of data points would have indeed left room for the logically possible interpretation of the results as if they concerned two distinct and bordering geolect continua. Such distinct continua could possibly be poorly intelligible to one another. Certainly, such an interpretation would be in contrast with the 'dialect' studies that I am aware of, which describe instead the geolects of Italy as clustering broadly around regional cores (Piedmontese, Lombard, Emilian, Romagnol, Tuscan, etc.) (see Ascoli, 1882–1885; Devoto and Giacomelli, 1972; Pellegrini, 1977; Maiden and Parry, 1997; Loporcaro, 2009). However, in order for the current results to be as exhaustive and resolute as possible, I preferred to preserve them even from the mere logical possibility of the two-belts-continua interpretation. In fact, intersecting the chains, a possible parallelism between the sequences of intelligibility scores of each chain, combined with possible high levels of intelligibility between at least the Gallo-Italic" pairs of data points in each chain, could only be interpreted as a general trend shown by two chains of geolects that are necessarily representative of the same (smooth) continuum. This fact allows a unitary interpretation of the two chains results.

DIVERGENCES BETWEEN 'ITEM LOCALITY' AND 'PARTICIPANT LOCALITY'. In four cases out of twelve (three in Chain *A*, one in Chain *B*), the locality where the sentences were translated and recorded (item locality) did not exactly coincide with the locality where the intelligibility data were collected (participant locality). This was due to practical impediments in data collection (for Occitan-1, Occitan-2, and Romagnol-5) or to my personal convenience (for Lombard-10). In each of these cases, the tight classificatory partnership of the item geolect and participant geolect was considered as a sufficient condition allowing the substitution, despite some obvious differences between the two geolects (detail in Brasca, 2023). Table 3 and Table 4 provide the item and participant localities. Based on previous literature, the geolect spoken in each locality

²³ Each locality is both a 'participant locality' (i.e. data point) and an 'item locality' (where the items were recorded). When this distinction is unnecessary in the text, the term 'data point' is generalized.

²⁴ These bullet points (1 to 4) only refer to Chain *A*, but identical considerations apply to Chain *B*.

²⁵ From a strict geographical standpoint, the Tuscan-6 data point is not to the east of the Romagnol-5 data point (see Fig. 3). However, for the sake of simplicity, I will refer to their reciprocal position following the general trend: low number= west, high number= east. The same applies to Chain *B* Tuscan-12 data point.

²⁶ Italian is based on Tuscan.

Table 3

Chain A: item locality and participant locality for each data point.

	<i>Item locality</i>	<i>Participant locality</i>
<i>Occitan-1</i>	La Ròcha de Rama/La Roche-de-Rame (PACA – F)	Lè Viàrà/Villar-Saint-Pancrace (PACA – F)
<i>Occitan-2</i>	Pradzalà/Pragelato (TO)	Chaumont/Chiomonte (TO)
<i>Piedmontese-3</i>	Vigon/Vigone (TO)	Vigon/Vigone (TO)
<i>Emilian-4</i>	Varan/Varano de' Melegari (PR)	Varan/Varano de' Melegari (PR)
<i>Romagnol-5</i>	Marchèt/Mercato Saraceno (FC)	Sêrsna/Sarsina (FC)
<i>Tuscan-6</i>	<i>No items needed</i> ^a	Poppi (AR)

^a Due to the fact that all the Gallo-“Italic”-speaking participants speak Italian, too, which is based on Tuscan.

Table 4

Chain B: item locality and participant locality/-ies for each data point.

	<i>Item locality</i>	<i>Participant locality/-ies</i>
<i>Occitan-7</i>	Barcelona/Barcelonnette (PACA – F)	Barcelona/Barcelonnette (PACA – F)
<i>Occitan-8</i>	Vinai/Vinadio (CN)	Vinai/Vinadio (CN)
<i>Piedmontese-9</i>	Vila Falèt/Villafalsetto (CN)	Vila Falèt/Villafalsetto (CN)
<i>Lombard-10</i>	Pioltell/Pioltello (MI)	Cernusc/Cernusco s.N., Gorgonzœla/Gorgonzola, Pessan/Pessano (MI)
<i>Emilian-11</i>	Pavoll/Pavullo nel Frignano (MO)	Pavoll/Pavullo nel Frignano (MO)
<i>Tuscan-12</i>	<i>No items needed</i>	Piteglio (PT)

Table 5

Item sets played to the participants for each data point.

Chain A			Chain B		
Items 'at west'	Data point (participants)	Items 'at east'	Items 'at west'	Data point (participants)	Items 'at east'
Piedmon.-3	Occitan-1	Occitan-2	Piedmon.-9	Occitan-7	Occitan-8
Occitan-1	Occitan-2	<i>Piedmon.-3 filler</i>	Occitan-7	Occitan-8	<i>Piedmon.-9 filler</i>
Occitan-2	Piedmon.-3	Emilian-4	Occitan-8	Piedmon.-9	Lombard-10
Piedmon.-3	Emilian-4	Romagnol-5	Piedmon.-9	Lombard-10	Emilian-11
Emilian-4	Romagnol-5	<i>Piedmon.-3 filler</i>	Lombard-10	Emilian-11	<i>Piedmon.-9 filler</i>
Romagnol-5	Tuscan-6	<i>Emilian-4 filler</i>	Emilian-11	Tuscan-12	<i>Lombard-10 filler</i>

presented in Tables 3 and 4 is representative of a sub-group of a smooth 'dialect continuum' clustering broadly around a regional core (each of these 'dialect continua' is defined as a *language of Italy* in the *UNESCO Atlas of the World's Languages in Danger* (Moseley, 2010) and in several scholarly studies²⁷). Particularly, all six Occitan geolects are ascribed to the Vivaro-Alpine sub-group of Occitan; both Piedmontese geolects are ascribed to the *Alto Piemontese* sub-group of Piedmontese (see Allasino et al., 2007: Fig. 2.1); both Emilian geolects are ascribed to the western sub-group of Emilian; all Lombard geolects are ascribed to Milanese, in turn part of the western sub-group of Lombard; the Marchèt and Sêrsna geolects are both ascribed to Romagnol (considered by some scholars as a sub-group of the Emilian-Romagnol group); the Poppi geolect is ascribed to the Casentinese sub-group of Tuscan; and the Piteglio geolect is ascribed to the Pistoiese sub-group of Tuscan (see Ascoli, 1882–1885; Devoto and Giacomelli, 1972; Pellegrini, 1977; Maiden and Parry, 1997; Loporcaro, 2009; Sibille, 2014).²⁸ Therefore, (1.) the localities investigated do not represent any sort of 'linguistic islands'; and (2.) in the four data points where item locality and participant locality do not coincide, the respective two (or more) geolects are very similar to one another, and their *Abstand* classification is equivalent.

²⁷ In Ethnologue (Lewis et al., 2013) and Glottolog (Hammarström et al., 2022), too. See other examples in Tamburelli and Tosco (2021).

²⁸ UNESCO lists Emilian and Romagnol as two distinct languages (ISO 639-3 codes egl and rgn respectively). In the current study, this distinction will be assumed.

For each item geolect ('at west' and 'at east'), 7 items were spoken by a male and 7 by a female²⁹ native speaker. Participants were required to write down the final word (target word) of each item in their state language, Italian or French. Each correctly translated target word generated a score of 1 point, while an incorrect translation and a blank response generated 0 points. The percentage of target words correctly translated represented the level of intelligibility of the geolect for the participant.

Before playing the 28 items, two practice items were played (see "task familiarity effect" in Keating and Jegerski, 2015: 17; see Kalikow et al., 1977). In order to prevent priming effects (see Morton, 1969; also Tang and van Heuven, 2009), each of the 28 items was played only once. The participant was not pre-advised about the number (i.e. two) nor about the identity of the geolects involved in the test.

4.3. Material

4.3.1. Background questionnaire

Through a questionnaire (Appendix A), information was collected about the participants' age, gender, and place of residence (Q[uestion] 10); the current language usage at home (Q3); and language usage during childhood (Q4). This information served to verify whether the participants met the prerequisite selection criteria (see Section 4.4.1).

In order to check their potential effect on intelligibility, information—as responses on a five-point scale—was also collected from participants about the following variables (see Section 4.1): Foreign language competence (Q1); Latin language competence (Q2); Contact with target areas (Q5); Exposure to target languages (Q6); Attitudes towards target areas (Q7); Attitudes towards target languages (Q8); Behavioural component of language attitudes (Q9).

In order to preserve the participants' naivety, in questions 5 to 9 I used distractors, asking participants to rate other areas and languages of Italy, France, or Switzerland (overall ratio: two target questions and four distractors). The participants were required to answer each question. The questionnaire was a variation of the one used by Tamburelli (2014).

4.3.2. Stimuli

Kalikow et al. (1977) provides a list of 200 English sentences with a highly predictable target word (henceforth 'high-predictability sentences', see example 1) and 200 English sentences with a minimally predictable target word (example 2):

1. Spread some butter on your bread.
2. He could discuss the bread.

High-predictability sentences provide the listener with semantic, syntactic, and prosodic cues, helping them in identifying the meaning of the target word (Kalikow et al., 1977; Wang, 2007). Therefore, a SPIN test run with high-predictability sentences focuses on the intelligibility of the entire sentence, rather than of the final word alone. This characteristic made high-predictability sentences appear preferable for the current study, whose aim is comparing the overall intelligibility between related geolects. In fact, in the case that all the first four words of the low-predictability sentence 2 were perfectly intelligible to a participant, but not the unpredictable last (target) word, the resulting 0 score would not properly represent the real overall level of intelligibility, but rather only the level of the participant's lexical competence, even limited to the sole last word. Instead, in the case of the high-predictability sentence 1, the correct identification of the last word could result for each of the following reasons: (a.) the intelligibility of the context (precedent words) helps the participant to 'guess' the unintelligible target word; or (b.) to 'grasp' the not perfectly intelligible target word (see Schotter et al., 2014, and references there); or (c.) the target word is intelligible regardless of the unintelligibility of the context. Therefore, a high-predictability sentences test measures intelligibility due to all the above reasons, all of which actually allow successful communication in real-life situations (see Miller et al., 1951 and Miller, 1962 cited in Kalikow et al., 1977; Valentini-Botinhao and Wester, 2014; see "interaction of bottom-up... and top-down... processes in... speech recognition" in Tang and van Heuven, 2009: 714).

4.3.2.1. *First selection and translation.* As a first step, 64 high-predictability English sentences (henceforth 'sentence meanings'³⁰) taken from Kalikow et al. (1977) were translated into the 10 non-Tuscan geolects. I chose these 64 sentence meanings for their applicability to the Occitan, Padanian,³¹ and Tuscan common cultural context. The sentences

²⁹ However, see in Section 4.3.2.2 the exception of Occitan-1 items.

³⁰ See McMahon and McMahon (2005: 29, 34).

³¹ I.e., relating to the Po valley and Romagna.

were translated by some native speakers (translators), self-reporting and/or reported by the community as endowed with good linguistic competence and practice of the item gelect.

4.3.2.2. *Recordings.* A male and a female native speaker recorded all of the 64 sentences each (see Section 4.1). In the Occitan-1 locality (France), no female native speakers were available, so two male native speakers were recorded. The sentences were uttered at a natural speed, loudness, and intonation, with no breaks between the words. The recordings were made in quiet spaces. I used a digital voice Dictaphone ZOOM H2 Handy Recorder, which produces high-quality WAV files.

4.3.2.3. *Final selection.* 28 sentence meanings were selected out of the 64 that were recorded and whose translations ended in a mono- or disyllabic noun (target word) in all 10 target gelects (henceforth ‘canonical sentence meanings’). Indeed, previous research shows that word length has an effect on word recognition (Wiener and Miller, 1946; see also Kürschner et al., 2008). For Romagnol-5 and Occitan-7 target gelects, inadequate translations (i.e. with last word longer than two syllables) of respectively one and two of the 28 canonical sentence meanings were replaced with adequate translations of other three sentence meanings (Table 6). These were taken from the group of 36 translated sentences. I did so on the assumption that each replacing non-canonical sentence was approximatively as intelligible as the replaced canonical one. This assumption was based on the opinion of some native speakers (e.g. the translators), and corroborated by their presence in local gelect dictionaries.

Table 6

Four levels Latin square design. *For the sentence meanings marked with an asterisk, see the penultimate paragraph of Section 4.3.2.4.

		<i>Acoustic lists</i>			
		1st	2nd	3rd	4th
1	I cut my finger with a knife.	WM	WF	EM	EF
2	The candle flame melted the wax.	WF	EM	EF	WM
3	The baby slept in his crib.	EM	EF	WM	WF
4	How long can you hold your breath?	EF	WM	WF	EM
5	Stir your coffee with a spoon.	WF	EM	EF	WF
6	Old metal cans were made with tin.	EM	EF	WM	EM
7	The beer drinkers raised their mugs.	EF	WM	WF	EF
8	Hold the baby on your lap.	WM	WF	EM	WM
9	The workers are digging a ditch.	EM	EF	WF	EM
10	Spread some butter on your bread.	EF	WM	EM	EF
11	The plow ^a was pulled by an ox.	WM	WF	EF	WM
12	The shepherd watched his flock of sheep.	WF	EM	WM	WF
13	The judge is sitting on the bench.	EF	WF	EM	WM
14	The cut on his knee formed a scab.	WM	EM	EF	WF
15	The farmer baled the hay.	WF	EF	WM	EM
16	She wore a feather in her cap.	EM	WM	WF	EF
17	At breakfast he drank some juice.	WF	EM	WM	WF
18	Raise the flag up the pole.	EM	EF	WF	EM
19	The landlord raised the rent.	EF	WM	EM	EF
20	Tom fell down and got a bad bruise.*	WM	WF	EF	WM
21	The chicks followed the mother hen.	EM	WM	WF	EM
22	A bear has a thick coat of fur.	EF	WF	EM	EF
23	His boss made him work like a slave.	WM	EM	EF	WM
24	The cigarette smoke filled his lungs.	WF	EF	WM	WF
25	The stale bread was covered with mold.	WM	WF	EM	EF
26	Our seats were in the second row.	WF	EM	EF	WM
27	That accident gave me a scare.*	EM	EF	WM	WF
28	I've got a cold and a sore throat.*	EF	WM	WF	EM

^a For “plow” and “mold” I use the original American spelling used in Kalikow et al. (1977).

Previous literature suggests that 28 stimulus sentences are sufficient to provide an accurate intelligibility picture (see [Leonardi, 2016](#) and examples in it, among which are [Dollaghan and Campbell, 1998](#); [Girbau, 2016](#)). Furthermore, in previous studies where a SPIN test was used to measure distance between related geolects, a number of stimulus sentences smaller than 28 was considered appropriate ([Tamburelli, 2014](#); [Leonardi, 2016](#)).³²

The raw file of each item was edited as follows: the volume was 'normalized' in order to homogenise the loudness across the entire study, and possible distracting environmental noises were deleted.

4.3.2.4. *Acoustic lists.* [Table 5](#) summarises what item sets were played to the participants of each data point as stimuli and possibly as fillers. As an example, read the fifth row of Chain A as follows: 'Romagnol-5 participants were played 14 stimuli spoken in Emilian-4 geolect, and 14 fillers spoken in Piedmontese-3 geolect'.

In [Table 5](#)—like in all the subsequent tables and figures—I generalized for the sake of convenience the use of the expressions 'items at west' and 'items at east',³³ also in cases (three out of twelve per chain) in which these expressions do not correspond to the reciprocal geographical position that they literally mean (compare with [Fig. 3](#)).

These were the variables studied and the respective levels:

- a. *intelligibility score* was the dependent variable;
- b. '*data point*' was the independent variable, with two levels (corresponding to the two data points whose intelligibility scores were compared in each individual test)³⁴;
- c. *gender of voice*, with two levels (male and female);
- d. *participant gender*, with two levels (male and female).

Adding the fillers allowed me to balance the item order in the acoustic lists according to a 4-level³⁵ Latin Square design for each data point. Latin square designs are normally chosen in order to avoid order effects (see "counterbalancing" in [Keating and Jegerski, 2015](#): 9). [Table 6](#) reports the 28 'canonical' sentence meanings whose translations finally constituted the items, in the order in which they were played to the participants of each data point. WM = male voice (speaking the geolect) 'at west'; WF = female voice 'at west'; EM = male voice 'at east'; EF = female voice 'at east'.

The inadequate Romagnol-5 translation of sentence 27 was replaced with the translation of 'Wash the floor with a mop'. The inadequate Occitan-7 translations of sentences 20 and 28 were replaced respectively with the translations of 'Instead of a fence, plant a hedge' and 'Wash the floor with a mop'.

Before running the tests, the entire set of 28 items of each item geolect (by two genders of voice) was played to two (non-participant) native speakers of the item geolect itself. This was done in order to test the validity of the stimuli. The two listeners of all localities scored 100% correct responses. Furthermore, a pilot study was run in the Lombard-10 data point among five Lombard native speakers ('pilot participants', mean age 70.4 years) to ensure that the questionnaire was clear, and that the translated sentences were suitable as stimulus material. No issues materialised in this respect, nor were any raised by the pilot participants.

4.3.3. Answer sheet

The final word (target word) of each of the 28 items had to be written down in the corresponding numbered blank space of an answer sheet.

4.4. Participants

4.4.1. Participant selection criteria

Participants had to be native speakers of the local geolect, have grown up with parents and grandparents who were native speakers of the local geolect, and who, moreover, spoke it at home. In order to control for participant gender, an equal number of male and female participants should be selected. The age of participants should be approximately that of between 65 and 80 years. In fact, all the regional languages involved, excepting Tuscan/Italian, are listed in the UNESCO Atlas of the World's Languages in Danger ([Moseley, 2010](#); see also [Soria, 2015](#)), meaning that their use is declining among younger generations ([Coluzzi, 2007, 2009](#); [Tosco, 2008, 2011](#); [Tamburelli, 2014](#); [Brasca, 2021](#); [Coluzzi et al., 2021](#)). Participants should have as low a level of foreign and Latin language competence as possible.

³² In the current study, each participant was played 14 items spoken in each target geolect. However, the four-level Latin square design of the item lists (see [Section 4.3.2.4](#)) allowed for the playing of 28 items spoken in each target geolect to the participants' group as a whole.

³³ In other passages, '*geolect/stimuli/area... at west*' and '*geolect/stimuli/area... at east*' will also be generalized.

³⁴ See in [Brasca \(2023: Section 4.2.5.2\)](#) a detailed presentation of the independent variable '*data point*'.

³⁵ Corresponding to the two levels in *b* plus the two levels in *c*.

They should be born and have possibly permanently lived in the data point locality, and had moved from their region as little as possible during their life. Participants should have had as low a contact as possible with the target areas, and as low exposure as possible to the target languages. Participants should present no hearing impairments.

4.4.2. Participants' recruitment and selection

For each data point, a local native speaker (contact person) helped the researcher by contacting and recruiting 24 to 30 participants (see "attrition" and "additional participants" in Keating and Jegerski, 2015: 27). The number of selected participants by data point is shown in Table 7.

131 recruited participants met the selection criteria for Chain A, and 143 for Chain B. Henceforth, these will be simply referred to as the participants of the study.

4.4.3. Participants: Descriptive statistics

4.4.3.1. Chain A. 131 participants were selected (mean age = 71.3 years, min-max = 37-92, SD = 6.856). See age by gender in Table 8 and Fig. 4.

Table 7
Number of selected participants (whose responses constituted the data).

Chain A		Chain B	
Occitan-1	11 (3F)	Occitan-7	23 (11F)
Occitan-2	24	Occitan-8	24
Piedmontese-3	24	Piedmontese-9	24
Emilian-4	24	Lombard-10	24
Romagnol-5	24	Emilian-11	24
Tuscan-6	24	Tuscan-12	24
Tot	131	Tot	143

Table 8
Chain A participants' gender, and age by gender.

N	68 male	63 female
Percentage	51.91%	48.09%
Mean age in years (SD)	71.47 (7.789)	71.11 (5.739)

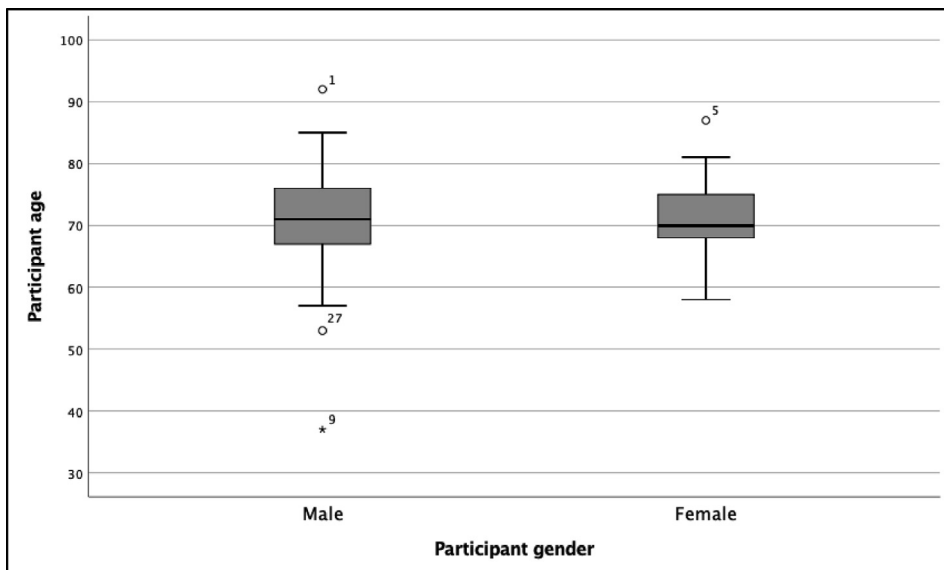


Fig. 4. Chain A participants' age by gender. The small circles and the asterisk indicate, respectively, the outlier and extreme outlier age values of participants n. 1, 5, 9 and 27, out of the 131 participants.

In each of the five data points situated in the Italian administrative territory there were 12 male and 12 female participants. In these data points as a whole, the participants' mean age was 70.88 years (min–max = 53–81; $SD = 6.042$). Except for one participant aged 53, the distribution of the age substantially matched the selection criterion.

In the Occitan-1 data point (France), there were 11 participants (mean age = 75.82 years, min–max = 37–92, $SD = 14.4$; 3 female). Given the limited number of possible participants in this data point, due to the advanced language shift situation, I extended recruitment to younger and older people. Therefore, four people aged 92, 87, 85, and 37 years were selected.

4.4.3.2. *Chain B*. 143 participants were selected (mean age = 71.91 years, min–max = 38–98, $SD = 6.664$). See age by gender in [Table 9](#) and [Fig. 5](#).

In each of the five data points situated in the Italian administrative territory there were 12 male and 12 female participants. In these data points as a whole, the participants' mean age was 71.42 years (min–max = 65–80, $SD = 4.859$). In the Occitan-7 (France) data point, there were 23 participants (11 female; mean age = 74.43 years, min–max = 38–98, $SD = 12.288$). As in Chain A, I extended recruitment to younger and older people. Therefore, in this data point, 6 people aged 98, 90, 89, 87, 86, and 38 years were selected.

4.4.4. Ethics statement

Participants gave their written informed consent. This study has received ethical approval from Bangor University's Research Ethics Committee.

4.5. Procedure

The tests were conducted in individual sessions and were held in a quiet room, furnished with a table and two chairs. Each participant was asked to complete the consent form and, after that, the background questionnaire ([Appendix A](#)). These were both provided in the state language (Italian or French). It took 10–15 min to complete the questionnaire. I

Table 9
Chain B participants' gender and age by gender.

N	72 male	71 female
Percentage	50.35%	49.65%
Mean age in years (SD)	72.06 (4.878)	71.76 (8.118)

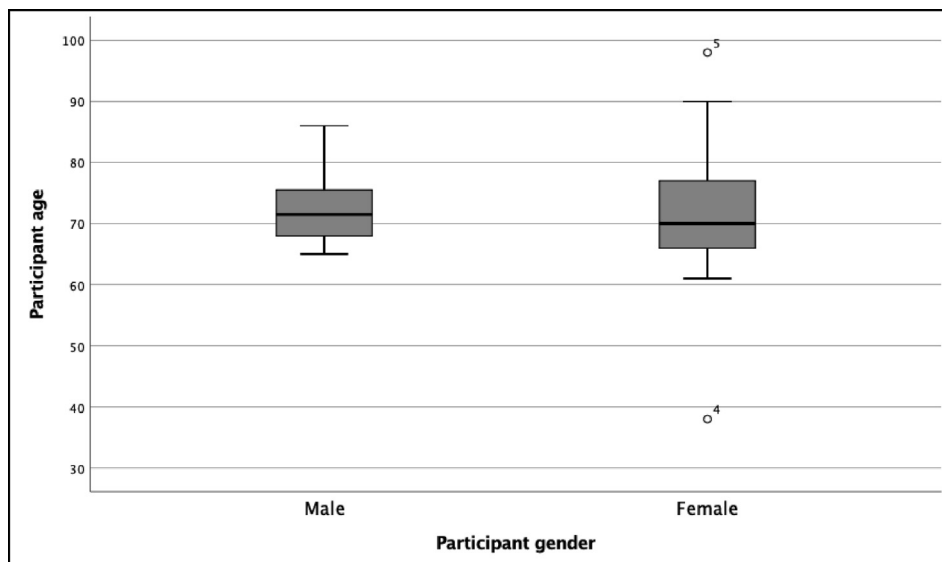


Fig. 5. Chain B participants' age by gender. The small circles indicate the outlier age values of participants n. 4 and 5, out of the 143 participants of Chain B.

collected 10 responses per participant, for a total of 2740 responses (10 questions³⁶ × 274 participants) in the entire study (both chains); 1310 responses in Chain A and 1430 in Chain B.

Next, each participant was instructed in detail about the intelligibility test (task): they would listen to 30 sentences, and after each sentence they would have to record the Italian (or French) translation of the final word (target word) in the corresponding blank space of the answer sheet. Participants were allowed to leave the blank space empty if they were not able to provide a translation, and they could decide when they were ready to listen to the next item. Once the start button had been pressed, there was a 2.0-second pause, and then the participant would listen to the new item (see e.g. [Thierry and Wu, 2007](#); [Gooskens et al., 2008](#); [Gooskens and Heeringa, 2014](#)).

The items (high-quality mp3 stereo 44,100 Hz, 320 kbps files) were played by a Sony Walkman digital audio player, and listened to through high-fidelity SONY MDR-ZX110NC headphones. The two initial practice items also served to adjust the headphones and the volume. I obtained 5656 intelligibility responses³⁷ in the whole two-chain study.

5. RESULTS

For all statistical tests an alpha level of 0.05 was used.

5.1. Chain A results

5.1.1. Intelligibility data³⁸ descriptive statistics

Intelligibility scores are in [Table 10](#), [Table 11](#), and [Fig. 6](#).

Table 10

Mean percentage intelligibility scores, percent confidence interval (95%) and SD. Small letters (*a – i*) refer to the boxplots in [Fig. 6](#).

	<i>Stimuli</i>				
	Occitan-1	Occitan-2	Piedmontese-3	Emilian-4	Romagnol-5
<i>Participants</i>	–	51.93	55.21	–	–
Occitan-1 ^a		45.86–58.07 1.272 <i>b</i>	46.86–63.50 1.737 <i>a</i>		
Occitan-2	66.36 62.86–69.86 1.160 <i>c</i>	–	–	–	–
Piedmontese-3	–	52.07 49.36–54.86 0.908 <i>d</i>	–	85.43 79.71–91.14 1.899 <i>e</i>	–
Emilian-4	–	–	78.00 74.64–81.29 1.100 <i>f</i>	–	83.93 79.43–88.43 1.482 <i>g</i>
Romagnol-5	–	–	–	83.36 78.93–87.79 1.465 <i>h</i>	–
Tuscan-6	–	–	–	–	36.64 31.79–41.43 1.597 <i>i</i>

^a For some (currently irrelevant) considerations about the not significantly different scores of Occitan-1 (Fr) participants on Occitan-2 stimuli and on the geographically more distant Piedmontese-3 stimuli, see [Brasca \(2023: Sections 4.3.A.5.2 and 4.3.B.5.2\)](#).

³⁶ The questions were formally ten (Q1 to Q10), but some of them required a double response, namely about geolect/area 'at west'

³⁷ 7672 responses, if one includes those given to the fillers.

³⁸ Data are available on request to the author.

Table 11

Mean raw intelligibility scores (out of 14 stimuli), confidence interval for mean (95%) and min-max. Small letters (a – i) refer to the boxplots in Fig. 6.

	<i>Stimuli</i>				
	Occitan-1	Occitan-2	Piedmontese-3	Emilian-4	Romagnol-5
<i>Participants</i>	–	7.27	7.73	–	–
Occitan-1	–	6.42–8.13	6.56–8.89	–	–
		6–10	6–11		
		<i>b</i>	<i>a</i>		
Occitan-2	9.29	–	–	–	–
	8.80–9.78				
	7–12				
	<i>c</i>				
Piedmontese-3	–	7.29	–	11.96	–
		6.91–7.68		11.16–12.76	
		6–9		8–14	
		<i>d</i>			
Emilian-4	–	–	10.92	–	11.75
			10.45–11.38		11.12–12.38
			9–13		8–14
			<i>f</i>		
Romagnol-5	–	–	–	11.67	–
				11.05–12.29	
				8–14	
				<i>h</i>	
Tuscan-6	–	–	–	–	5.13
					4.45–5.80
					2–8
					<i>i</i>

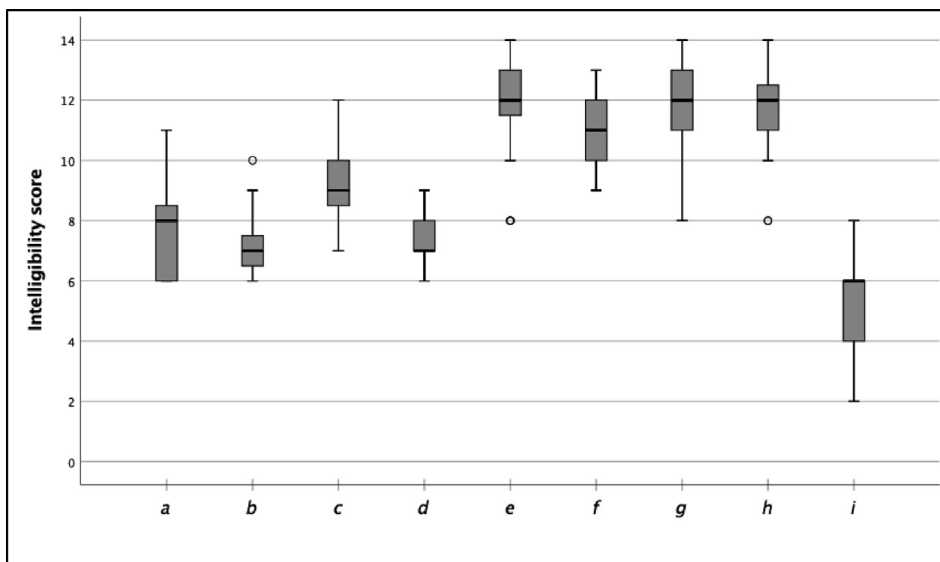


Fig. 6. Boxplots for the raw intelligibility scores compared in Chain A study, out of 14 stimuli. The small circles correspond to outlier scores of participants n. 11 (b), 37 and 41 (e), 87 and 99 (h), out of the 131 participants of Chain A. Small letters (a – i) refer each boxplot to the corresponding values in Table 10 and Table 11.

5.1.2. Gender

5.1.2.1. *Gender of voice.* Differently from all the other localities of both chains, in the Occitan-1 locality (France) the stimuli were produced by two male speakers (Section 4.3.2.2). When I checked the effect of gender of voice on intelligibility

scores, I consequently excluded the scores of the Occitan-2 participants. This is the only group of participants that listened to Occitan-1 stimuli.

A Wilcoxon Signed-Rank two-related-samples test was conducted. No significant difference was found between the two sets of scores obtained respectively on the male and on the female voices ($Z = -1.806$, $p = 0.071$, *ns*). No evidence was found that gender of voice had an effect on intelligibility score.

5.1.2.2. Participant gender. A Mann-Whitney U two-independent-samples test was conducted. No significant difference was found between the two sets of scores obtained respectively from male and female participants ($U = 1909.5$; $p = 0.282$, *ns*). No evidence was found that participant gender had an effect on intelligibility score.

5.1.3. Language competence, exposure, attitudes, and contact (LCEAC) variables³⁹

Two multiple regressions analyses were run⁴⁰ in order to check the potential effect of language competence, exposure, and attitudes, and contact with target area (henceforth LCEAC) variables (corresponding to questions 1, 2, 5, 6, 7, 8, and 9, see Appendix A) on the intelligibility score. One regression was run with scores on stimuli 'at west' and one with scores on stimuli 'at east' as a dependent variable. This is both because questions 5, 6, 7, 8, and 9 asked for separate responses concerning area/language 'at west' and 'at east', and because in three data points out of six, items 'at east' were fillers. The multiple regressions were run while controlling for 'data point'.⁴¹

Concerning the multiple regression with 'intelligibility score on stimuli *at west*' as dependent variable, using the 'enter' method, it was found that the model explains a significant amount of the variance in the intelligibility score ($F(14, 116) = 31.497$, $p < 0.001$, $R^2 = 0.792$, $R^2_{\text{Adjusted}} = 0.767$). The analysis also showed that 'data point' alone had an effect on intelligibility score, explaining a significant amount of the variance in the intelligibility score ($F(5, 125) = 81.265$, $p < 0.001$, $R^2 = 0.765$, $R^2_{\text{Adjusted}} = 0.755$). There is no evidence that any of the LCEAC variables individually had an effect on intelligibility score (Table 12).

Then, an identical multiple regression, but with 'intelligibility score on stimuli *at east*' as dependent variable, showed that the model explains a significant amount of the variance in the intelligibility score ($F(11, 47) = 6.588$, $p < 0.001$, $R^2 = 0.607$, $R^2_{\text{Adjusted}} = 0.515$). The analysis also showed that 'data point' alone had an effect on intelligibility score, explaining a significant amount of the variance in the intelligibility score ($F(2, 56) = 35.245$, $p < 0.001$, $R^2 = 0.557$, $R^2_{\text{Adjusted}} = 0.541$). In this case, too, there is no evidence that any of the LCEAC variables had individually an effect on intelligibility score (see Table 12):

In summary, the results of both multiple regressions did not provide evidence that any of the LCEAC variables individually affected intelligibility.

Table 12

Chain A: results of multiple regressions with *intelligibility score on stimuli 'at west' and 'at east'* as dependent variable.

LCEAC variable	'at west'			'at east'		
	Beta	t	p	Beta	t	p
Foreign language 1 competence	-0.080	-1.334	0.185	0.140	0.872	0.388
Foreign language 2 competence	0.125	1.744	0.084	0.005	0.023	0.982
Foreign language 3 competence	0.084	1.394	0.166	0.128	0.602	0.550
Latin language competence	0.040	0.804	0.423	-0.101	-0.850	0.399
Contact with target area	-0.082	-1.210	0.229	0.082	0.444	0.659
Exposure to target language	-0.031	-0.585	0.560	-0.330	-1.900	0.064
Attitude towards target area	-0.047	-0.896	0.372	0.049	0.417	0.679
Attitude towards target language (Q8)	-0.062	-1.192	0.236	0.014	0.134	0.894
Attitude towards target language (Q9)	0.060	1.071	0.286	0.133	1.204	0.235

³⁹ See descriptive statistics in Brasca (2023), where they are named "control variables".

⁴⁰ All assumptions of multiple regression were met (see Brasca, 2023).

⁴¹ Please remember that each data point corresponds to a different participant's native gelect, and to two different target gelects, respectively 'at west' and 'at east', hence the need to control for 'data point'. See Brasca (2023: Section 4.2.5.2) for a presentation of the independent variable 'data point'.

5.1.4. Testing the predictions made by the CCC

PREDICTION 1

To test Prediction 1,⁴² the following was compared:

- the intelligibility score obtained by Tuscan-6 participants on Romagnol-5 stimuli with the intelligibility score obtained
 - a. by Occitan-1 participants on Piedmontese-3 stimuli, and
 - b. by Piedmontese-3 participants on Occitan-2 stimuli.

In comparison *a*, Occitan-1 participants were used instead of Occitan-2 (usually tri-lingual) participants.

A Mann-Whitney U two-independent-samples test was conducted between each of the two pairs of scores⁴³ concerned. The results are displayed in Table 13.

Tuscan-6 participants' intelligibility scores on Romagnol-5 stimuli were statistically significantly different from both the other intelligibility scores, such that Tuscan-6 participants scored lower than Occitan-1 participants and lower than Piedmontese-3 participants (see Table 10, also for the next Chain A results).

PREDICTION 1-a

To test Prediction 1-a,⁴⁴ the following was compared:

- the score obtained by Tuscan-6 participants on Romagnol-5 stimuli with the score obtained
 - a. by Occitan-1 participants on Occitan-2 stimuli, and
 - b. by Occitan-2 participants on Occitan-1 stimuli.

A Mann-Whitney U two-independent-samples test was therefore conducted between each of the pairs of scores concerned. The results are in Table 14.

Table 13

Results of Mann-Whitney U two-independent-samples test.

<i>Test for the difference between Tuscan-6 participants' score on Romagnol-5 stimuli, and</i>	<i>U</i>	<i>p</i>	<i>% score difference^a</i>	<i>Reciprocal^b % score difference</i>
<i>a. Occitan-1 participants' score on Piedmontese-3 stimuli</i>	35.0	<0.001	+18.57	+50.68
<i>b. Piedmontese-3 participants' score on Occitan-2 stimuli</i>	63.5	<0.001	+15.43	+42.11

^a This value corresponds to mean score in *a*. (or in *b*.) minus Tuscan participants' mean score. This value, and the 'reciprocal percentage score difference' reported alongside in this table, are overabundant information. They just intend to facilitate comparison of scores in Table 10. I will report these values—in tables or in footnotes—only when comparisons involve Tuscan.

^b This value answers the following question: 'How much is the score in *a*. (or *b*.) bigger than Tuscan-6 participants' score if this latter is taken as 100%?' E.g., the value for *a*. is: $\{[(100 \times 55.21) : 36.64] - 100\} = 50.68$. Here, 55.21 and 36.64 are the mean percentage scores in *a*. and of Tuscan participants respectively (find them in Table 10).

Table 14

Results of Mann-Whitney U two-independent-samples test.

<i>Test for the difference between Tuscan-6 participants' score on Romagnol-5 stimuli, and</i>	<i>U</i>	<i>P</i>	<i>% score difference</i>	<i>Reciprocal % score difference</i>
<i>a. Occitan-1 participants' score on Occitan-2 stimuli</i>	33.5	<0.001	+15.29	+41.73
<i>b. Occitan-2 participants' score on Occitan-1 stimuli</i>	7.0	<0.001	+29.72	+81.11

⁴² Prediction 1: today intelligibility is higher between Romagnol-5 and Tuscan-6 than between Occitan-2 and Piedmontese-3 (see Section 3.3).

⁴³ Henceforth, in Section 5.1.4, 'score' means 'intelligibility score'.

⁴⁴ Prediction 1-a: today intelligibility is higher between Romagnol-5 and Tuscan-6 than between Occitan-1 and Occitan-2 (see Section 3.3).

Tuscan-6 participants' intelligibility score on Romagnol-5 stimuli was statistically significantly different from both the other intelligibility scores, such that Tuscan-6 participants scored lower than Occitan-1 and Occitan-2 participants.

PREDICTION 2

To test Prediction 2,⁴⁵ the comparisons presented in I and II were made.

- I. First, the following was compared:
 - a. the magnitude of the *possible* difference between
 - o Tuscan-6 participants' score on Romagnol-5 stimuli,
 - and
 - o Emilian-4 participants' score on Romagnol-5 stimuli,
 - b. with the magnitude of the *possible* difference between
 - o Occitan-1 participants' score on Occitan-2 stimuli,
 - and
 - o Piedmontese-3 participants' score on Occitan-2 stimuli.

Concerning *a*, a Mann-Whitney U two-independent-samples test revealed that the two sets of scores differed significantly ($U = 1.0$; $p < 0.001$).⁴⁶ The magnitude of the difference resulted in a value of Cohen's $d = 1.72$, approximately corresponding to what is described as a "huge"⁴⁷ effect (Sawilowsky, 2009).

Concerning *b*, a Mann-Whitney U two-independent-samples test was conducted. No significant difference was found ($U = 120.0$; $p = 0.654$, *ns*).

The results demonstrated that the magnitude of the difference in *a* (i.e. across the Apennines) is larger than in *b* (i.e. across the Alps).

- II. Second, the following was compared:
 - a. the magnitude of the *possible* difference between
 - o Tuscan-6 participants' score on Romagnol-5 stimuli,
 - and
 - o Romagnol-5 participants' score on Emilian-4 stimuli,
 - b. with the magnitude of the *possible* difference between
 - o Occitan-2 participants' score on Occitan-1 stimuli,
 - and
 - o Piedmontese-3 participants' score on Occitan-2 stimuli.

Concerning *a*, a Mann-Whitney U two-independent-samples test revealed a statistically (highly) significant difference between the two sets of scores ($U = 2.0$, $p < 0.001$).⁴⁸ The magnitude of the difference resulted in a value of Cohen's $d = 1.73$ (i.e. approximately "huge" effect; Sawilowsky, 2009).

Concerning *b*, a Mann-Whitney U two-independent-samples test was conducted, revealing a statistically significant difference between the two sets of scores ($U = 52.5$, $p < 0.001$). The magnitude of the difference resulted in a value of Cohen's $d = 1.43$ (i.e. approximately "very large" effect; Sawilowsky, 2009).

Results of comparison II parallel those of comparison I. They demonstrated that the magnitude of the difference in *a*, namely across the Apennines, was larger than the magnitude of the difference in *b*, namely across the Alps.

⁴⁵ Prediction 2: today intelligibility decreases more abruptly across the Alps (namely along the portion of chain running from Occitan-1, through Occitan-2 to Piedmontese-3) than across the Apennines (namely along the portion of chain running from Emilian-4, through Romagnol-5 to Tuscan-6) (see Section 3.3).

⁴⁶ Score difference: +47.29%; reciprocal score difference: +129.10%.

⁴⁷ Cohen's d magnitudes: 0.2= small; 0.5= medium; 0.8= large (Cohen, 1988); 1.2= very large; 2.0= huge (Sawilowsky, 2009).

⁴⁸ Score difference: +46.72%; reciprocal score difference: +127.51%.

PREDICTION 3*

To test Prediction 3*,⁴⁹ firstly, the following was compared:

- the score obtained by Tuscan-6 participants on Romagnol-5 stimuli

with the score obtained

- by Piedmontese-3 participants on Emilian-4 stimuli,
- by Emilian-4 participants on Piedmontese-3 stimuli,
- by Emilian-4 participants on Romagnol-5 stimuli (already compared for Prediction 2), and
- by Romagnol-5 participants on Emilian-4 stimuli.

A Mann-Whitney U two-independent-samples test was therefore conducted between each of the pairs of scores concerned. The results are displayed in Table 15.

Tuscan-6 participants' intelligibility score on Romagnol-5 stimuli was statistically significantly different from all the other intelligibility scores, such that Tuscan-6 participants scored lower than Emilian-4 participants on both Romagnol-5 and Piedmontese-3 stimuli, lower than Piedmontese-3 participants, and lower than Romagnol-5 participants.

To test Prediction 3* there was also a comparison of:

- the magnitude (already calculated for Prediction 2) of the difference between the scores obtained:
 - by Tuscan-6 participants on Romagnol-5 stimuli, and
 - by Emilian-4 participants on Romagnol-5 stimuli,
- with the magnitude of the *possible* difference between the scores obtained
 - by Piedmontese-3 participants on Emilian-4 stimuli, and
 - by Romagnol-5 participants on Emilian-4 stimuli,
- and with the magnitude of the *possible* difference between the scores obtained
 - by Occitan-1 participants on Piedmontese-3 stimuli, and
 - by Emilian-4 participants on Piedmontese-3 stimuli.

As was done for Prediction 1, the trilingual *Occitan-2* participants' data were replaced by the *Occitan-1* (Fr) participants' data.

Concerning the comparison between *a* and *b*, a Mann-Whitney U two-independent-samples test was performed between Piedmontese-3 participants' score on Emilian-4 stimuli, and Romagnol-5 participants' score on Emilian-4 stimuli. No significant difference was found between the two groups ($U = 233.50$, $p = 0.244$, *ns*).

Concerning the comparison between *a* and *c*, a *t*-test⁵⁰ was conducted between Occitan-1 participants' score on Piedmontese-3 stimuli, and Emilian-4 participants' score on Piedmontese-3 stimuli. The two groups differed significantly ($t(33) = -6.606$, $p < 0.001$). The intelligibility mean score of Occitan-1 participants was statistically significantly lower than the intelligibility mean score of Emilian-4 participants.

Table 15

Results of Mann-Whitney U two-independent-samples tests.

Test the difference between	U	p	% score difference	Reciprocal % score difference
Tuscan-6 participants' score on Romagnol-5 stimuli, and				
a. Piedmontese-3 participants' score on Emilian-4 stimuli	3.0	<0.001	+48.79	+133.16
b. Emilian-4 participants' score on Piedmontese-3 stimuli	0.0	<0.001	+41.36	+112.88
c. Emilian-4 participants' score on Romagnol-5 stimuli	1.0	<0.001	+47.29	+129.07
d. Romagnol-5 participants' score on Emilian-4 stimuli	2.0	<0.001	+46.72	+127.51

⁴⁹ Prediction 3*: today no line can be found in the Po valley and Romagna, between the Alps and the Apennines, across which intelligibility is lower and decreases more abruptly than across the Apennines (namely along the portion of chain running from Emilian-4, through Romagnol-5 to Tuscan-6) (see Section 3.3).

⁵⁰ When the parametric assumptions were met, i.e. in three cases in the whole study, a *t*-test was run in order to enjoy its advantages (e.g. more statistical power). However, in order to grant incontestable comparability of the results, a Mann-Whitney U two-independent-

Therefore, the magnitude of the difference in *c* just tested was calculated, resulting in a Cohen's *d* value of 1.17 (i.e. "very large" effect size, Sawilowsky, 2009).

These results demonstrate that the magnitude of difference in *a* (i.e. across the Apennines, "huge", see Prediction 2, *l*) is larger than the magnitudes of the difference in *b* and in *c*.

5.1.5. Interim discussion of Chain A results

This section presents the linguistic relevance of the results of each prediction testing and considers whether they support or are incompatible with the critical convergence claim (CCC). Note that in the present and next (Discussion and Conclusion) sections, when referring to the 'linguistic relevance of the results' intended as evidence for 'linguistic similarity', the term 'linguistic' refers to "lexical similarity" and "phonological correspondence", since these are the domains of the language for which the correlation between the SPIN test results and the objective linguistic distance was studied in previous literature (Tang and van Heuven, 2009: 724). The same applies for the expressions 'linguistic distance' and 'Abstand'. On the other hand, traditional classification is mostly based on these domains of the language, and, consequently, the grouping "built on" the SPIN test results "...adequately... reflect[s] the taxonomy of... dialects postulated by traditional dialectologists" (Tang and van Heuven, 2009: 726).⁵¹

The results show that Prediction 1 is not borne out, providing evidence, on the contrary, that today intelligibility is *lower* between Tuscan and Romagnol than between Piedmontese and Occitan of Italy. Based on the link between intelligibility and *Abstand* (Section 3.1), these results provide evidence that Piedmontese is more similar to Occitan of Italy than Romagnol is to Tuscan. This is incompatible with the CCC and is direct evidence against it. The unusable intelligibility data of Occitan-2 trilingual participants (Italy) on Piedmontese-3 stimuli were replaced by data of the geographically more distant Occitan-1 participants (France) on Piedmontese-3 stimuli (Section 4.2). Therefore, the results provide evidence that Piedmontese is more similar not only to Occitan of Italy but even to (the geographically further) Occitan of France, than Romagnol is to Tuscan.

The results show that Prediction 1-a is not borne out either, providing evidence, on the contrary, that today intelligibility is *lower* between Tuscan and Romagnol than between Occitan of Italy and Occitan of France. This in turn indicates that the results provide evidence that Occitan of France is more similar to Occitan of Italy than Romagnol is to Tuscan. Given that prediction 1 was not borne out either, this is incompatible with the CCC.

Moreover, Prediction 2 is not borne out either. On the contrary, the results provide evidence that today intelligibility decreases *less* abruptly across the Alps than across the Apennines. In *Abstand* terms, this indicates that today linguistic similarity decreases more abruptly across the Apennines than across the Alps.

Finally, let's recall Prediction 3*: Today no line can be found in the Po valley and Romagna, between the Alps and the Apennines, across which intelligibility (*i*) is lower than and (*ii*) decreases more abruptly than across the Apennines. The results show that both Prediction 3*(*i*) and (*ii*) are borne out. This in turn indicates that there is evidence that the whole of Po valley and Romagna remained linguistically united, forming, with Alpine Occitania of France, a continuum which is smoother than the continuum it forms with Tuscany.

The fact that Romagnol-5 participants were recruited in Sêrsna instead of Marchèt (item locality) strengthens the informativeness of the results. Indeed, Sêrsna is even further from the Emilian-4 stimuli locality (Varan) than Marchèt is, and in particular it is located in an upper part of the Apennines (in the same valley), where the literature unanimously indicates that more conservative Gallo-Italic geolects are spoken than those spoken in the lower Apennines (where both Varan and Marchèt are situated) (e.g. Loporcaro, 2009). An exploratory investigation in Brasca (2023) confirmed that this general trend is observable in the specific localities at issue, too. Therefore, it appears likely that had Marchèt participants been used, they would have scored even higher on Emilian-4 (Varan) stimuli than Sêrsna participants did.⁵²

In summary, Chain A results provide direct evidence against the critical convergence claim.

5.2. Chain B results

Identical analyses as for Chain A were run for Chain B. They are more concisely presented in this section.

5.2.1. Intelligibility data descriptive statistics

Intelligibility scores are in Table 16 and Fig. 7.

⁵¹ In this respect, recall that the current study wants to contribute to solve a disagreement present in the traditional comparative and dialectological literature.

⁵² This argument quantitatively strengthens the pro-Gallo- interpretation of the results. It is in any case overabundant in order to falsify the CCC.

Table 16

Mean percentage intelligibility scores, percent confidence interval (95%), and SD. Small letters (*a – i*) refer to the boxplots in Fig. 7.

	<i>Stimuli</i>				
	Occitan-7	Occitan-8	Piedmontese-9	Lombard-10	Emilian-11
<i>Participants</i>	–	85.71	49.71	–	–
Occitan-7	–	80.86–90.57 1.567 <i>b</i>	44.36–55.00 1.718 <i>a</i>	–	–
Occitan-8	73.21 68.57–77.86 1.539 <i>c</i>	–	–	–	–
Piedmontese-9	–	88.99 85.55–92.43 1.141 <i>d</i>	–	85.12 80.86–89.38 1.412 <i>e</i>	–
Lombard-10	–	–	70.24 66.72–73.76 1.167 <i>f</i>	–	82.44 78.67–86.21 1.250 <i>g</i>
Emilian-11	–	–	–	83.93 80.57–87.29 1.113 <i>h</i>	–
Tuscan-12	–	–	–	–	37.50 33.60–41.40 1.294 <i>i</i>

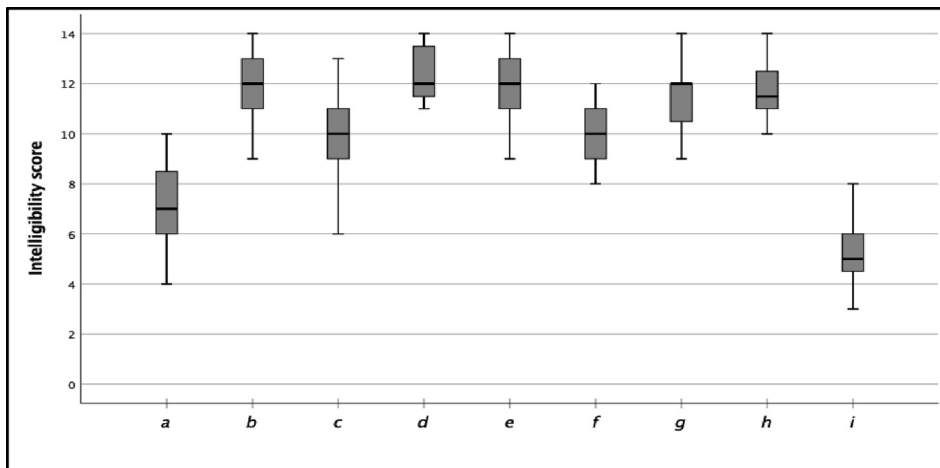


Fig. 7. Boxplots for the intelligibility scores compared in Chain B study, out of 14 stimuli. Small letters (*a – i*) refer each boxplot to the corresponding values in Table 16.

5.2.2. Gender

Gender of voice. A Wilcoxon Signed-Ranks two-related-samples test was conducted. No significant difference was found ($Z = -0.847$, $p = 0.397$, *ns*).

Participant gender. A Mann-Whitney U two-independent-samples test did not reveal a significant difference ($U = 2409.5$, $p = 0.551$, *ns*).

Table 17

Chain B: results of multiple regressions with *intelligibility score on stimuli 'at west' and 'at east'* as dependent variable.

LCEAC variables	'at west'			'at east'		
	Beta	t	p	Beta	t	p
Foreign language 1 competence	0.040	0.709	0.480	-0.022	-0.156	0.877
Foreign language 2 competence	-0.028	-0.589	0.557	0.196	1.450	0.152
Latin language competence	0.032	0.679	0.498	0.042	0.286	0.776
Contact with target area	0.020	0.376	0.707	-0.004	-0.025	0.980
Exposure to target language	-0.003	-0.064	0.949	0.079	0.576	0.567
Attitude towards target area	0.039	-0.824	0.412	0.163	1.082	0.284
Attitude towards target language (Q8)	0.034	0.691	0.491	0.181	1.209	0.231
Attitude towards target language (Q9)	-0.044	-0.777	0.438	-0.136	-0.814	0.419

5.2.3. Language competence, exposure, attitudes, and contact (LCEAC) variables⁵³

A multiple regression with 'intelligibility score on stimuli *at west*' as dependent variable was conducted on the LCEAC variables while controlling for 'data point'. Using the 'enter' method, it was found that the model explains a significant amount of the variance in the intelligibility score ($F(13, 129) = 38.531, p < 0.001, R^2 = 0.795, R^2_{Adjusted} = 0.775$). The analysis also showed that 'data point' alone had an effect on intelligibility score, explaining a significant amount of the variance in the intelligibility score ($F(2, 68) = 95.762, p < 0.001, R^2 = 0.738, R^2_{Adjusted} = 0.730$). On the contrary, there is no evidence that any of the LCEAC variables individually had an effect on intelligibility score (Table 17).

An identical multiple regression but with 'intelligibility score on stimuli *at east*' as dependent variable was also conducted. No evidence was found that the model significantly affected the intelligibility score ($F(10, 60) = 0.690, p < 0.730, R^2 = 0.103, R^2_{Adjusted} = -0.046$). Moreover, no evidence was found that 'data point' alone had an effect on intelligibility score ($F(2, 68) = 0.708, p < .496 ns, R^2 = 0.020, R^2_{Adjusted} = -0.008$). Furthermore, there is no evidence that any of the LCEAC variables individually had an effect on intelligibility score (Table 17):

In summary, neither multiple regression provided evidence that any of the LCEAC variables individually affected intelligibility.

5.2.4. Testing the predictions

PREDICTION 1

In order to test Prediction 1,⁵⁴ this section compares:

- the score⁵⁵ obtained by Tuscan-12 participants on Emilian-11 stimuli

with the score obtained

- by Occitan-7 participants on Piedmontese-9 stimuli, and
- by Piedmontese-9 participants on Occitan-8 stimuli.

The results of the respective tests are in Table 18. Tuscan-12 participants' intelligibility score on Emilian-11 stimuli was statistically significantly different from both the other scores, such that Tuscan-12 participants scored lower than Occitan-7 and Piedmontese-9 participants (see Table 16, also for the next Chain B results).

PREDICTION 1-a

To test Prediction 1-a,⁵⁶ the following was compared:

- the score obtained by Tuscan-12 participants on Emilian-11 stimuli with the score obtained
 - by Occitan-7 participants on Occitan-8 stimuli, and
 - by Occitan-8 participants on Occitan-7 stimuli.

⁵³ See descriptive statistics in Brasca (2023), where they are named "control variables".

⁵⁴ Prediction 1: today intelligibility is higher between Emilian-11 and Tuscan-12 than between Occitan-8 and Piedmontese-9 (see Section 3.3).

⁵⁵ In Section 5.2.4, 'score' means 'intelligibility score'.

⁵⁶ Prediction 1-a: today intelligibility is higher between Emilian-11 and Tuscan-12 than between Occitan-7 and Occitan-8 (see Fig. 3).

Table 18
Results of tests of significance.

	Two-independent-sample t-test ^a	Mann-Whitney U two-independent-samples test	Sig.	% score difference	Reciprocal % score difference
<i>Test for the difference between Tuscan-12 participants' score on Emilian-11 stimuli, and</i>	<i>t(45)</i>	<i>U</i>	<i>p</i>		
a. Occitan-7 participants' score on Piedmontese-9 stimuli	3.857		<0.001	+12.21	+32.56
b. Piedmontese-9 participants' score on Occitan-8 stimuli		0.0	<0.001	+51.49	+137.31

^a A Mann-Whitney U two-independent-samples test was also conducted, yielding equivalent (for the current aims) results: $U = 118.500$, $p < 0.001$.

Table 19
Results of Mann-Whitney U two-independent-samples tests.

<i>Test for the difference between Tuscan-12 participants' score on Emilian-11 stimuli, and</i>	<i>U</i>	<i>p</i>	<i>% score difference</i>	<i>Reciprocal % score difference</i>
a. Occitan-7 participants' score on Occitan-8 stimuli	0.0	<0.001	+48.21	+128.56
b. Occitan-8 participants' score on Occitan-7 stimuli	7.0	<0.001	+35.71	+95.23

The results of the tests are in Table 19. Tuscan-12 participants' intelligibility score on Emilian-11 stimuli was statistically significantly different from both the other intelligibility scores, such that Tuscan-12 participants scored lower than Occitan-7 and Occitan-8 participants.

PREDICTION 2

In order to test Prediction 2,⁵⁷ this section makes two comparisons (I and II):

- I. First, the following was compared:
 - a. the magnitude of the *possible* difference between
 - i. Tuscan-12 participants' score on Emilian-11 stimuli, and
 - ii. Lombard-10 participants' score on Emilian-11 stimuli,
 - b. with the magnitude of the *possible* difference between
 - iii. Occitan-7 participants' score on Occitan-8 stimuli, and
 - iv. Piedmontese-9 participants' score on Occitan-8 stimuli.

The results of the tests are in Table 20. Results demonstrated that the magnitude of the difference in a (i.e. across the Apennines) is larger than the magnitude of the difference in b (i.e. across the Alps).

- II. Second, the following was compared:
 - a. the magnitude of the *possible* difference between
 - i. Tuscan-12 participants' score on Emilian-11 stimuli, and
 - ii. Emilian-11 participants' score on Lombard-10 stimuli,
 - b. with the magnitude of the *possible* difference between
 - iii. Occitan-8 participants' score on Occitan-7 stimuli, and
 - iv. Piedmontese-9 participants' score on Occitan-8 stimuli.

⁵⁷ Prediction 2: today intelligibility decreases more abruptly across the Alps (namely along the portion of chain running from Occitan-7, through Occitan-8 to Piedmontese-9) than across the Apennines (namely along the portion of chain running from Lombard-10, through Emilian-11 to Tuscan-12) (see Section 3.3).

Table 20
Results of tests of significance and corresponding magnitudes of difference.

	Two-independent-samples t-test ^a	Mann-Whitney U two-independent-samples test	Sig.	Magnitude of difference
<i>Test for the difference between</i>	<i>t(46)</i>	<i>U</i>	<i>p</i>	<i>Cohen's d</i>
a. Tuscan-12 participants' score on Emilian-11 stimuli, and Lomb.-10 participants' score on Emil.-11 stimuli ^b	17.131		<0.001	1.84 "huge"
b. Occitan-7 participants' score on Occitan-8 stimuli, and Piedm.-9 participants' score on Occitan-8 stimuli		232.5	0.343	–

^a A Mann-Whitney U two-independent-samples test was also conducted, which yielded equivalent (for the current purposes) results: $U = 0.0$, $p < 0.001$; *Cohen's d* = 1.73 (i.e. "huge" effect size, Sawilowsky, 2009).

^b Score difference: +44.94%; reciprocal score difference: +119.84%.

Table 21
Results of Mann-Whitney U two-independent-samples tests and corresponding magnitudes of difference (*Cohen's d*).

<i>Test for the difference between</i>	<i>U</i>	<i>p</i>	<i>Cohen's d</i>
a. Tuscan-12 participants' score on Emilian-11 stimuli, and Emilian-11 participants' score on Lombard-10 stimuli ^a	0.0	<0.001	1.73 "huge"
b. Occitan-8 participants' score on Occitan-7 stimuli, and Piedmontese-9 participants' score on Occitan-8 stimuli	71.5	<0.001	1.31 "very large"

^a Score difference: +46.43%; reciprocal score difference: +123.81%.

The results of the tests are in Table 21. The results of comparison II confirmed the results of comparison I. Indeed, they demonstrated that the magnitude of the difference in *a* (i.e. across the Apennines) is larger than the magnitude of the difference in *b* (i.e. across the Alps).

PREDICTION 3*

In order to test Prediction 3*,⁵⁸ this section will compare:

- the score obtained by Tuscan-12 participants on Emilian-11 stimuli

with the score obtained

- by Piedmontese-9 participants on Lombard-10 stimuli,
- by Lombard-10 participants on Piedmontese-9 stimuli,
- by Lombard-10 participants on Emilian-11 stimuli (already compared, Table 20), and
- by Emilian-11 participants on Lombard-10 stimuli.

The results of the respective tests of significance are in Table 22. The Tuscan-12 participants' intelligibility score on Emilian-11 stimuli differed statistically significantly from all the other intelligibility scores, such that Tuscan-12 participants scored lower than Piedmontese-9 participants, lower than Lombard-10 participants (on both Emilian-11 and Piedmontese-9 stimuli), and lower than Emilian-11 participants.

To test Prediction 3*, the following was compared, too:

- the magnitude (already calculated, Table 20) of the difference between scores obtained
 - by Tuscan-12 participants on Emilian-11 stimuli, and
 - by Lombard-10 participants on Emilian-11 stimuli,

⁵⁸ Prediction 3*: today no line can be found in the Po valley and Romagna, between the Alps and the Apennines, across which intelligibility is lower and decreases more abruptly than across the Apennines (namely along the portion of chain running from Lombard-10, through Emilian-11 to Tuscan-12) (see Fig. 3).

Table 22
Results of Mann-Whitney U two-independent-samples tests.

<i>Test for the difference between Tuscan-12 participants' score on Emilian-11 stimuli, and</i>	<i>U</i>	<i>p</i>	<i>% score difference</i>	<i>Reciprocal % score difference</i>
a. Piedmontese-9 participants' score on Lombard-10 stimuli	0.0	<0.001	+47.62	+126.99
b. Lombard-10 participants' score on Piedmontese-9 stimuli	2.0	<0.001	+32.74	+87.31
d. Emilian-11 participants' score on Lombard-10 stimuli	0.0	<0.001	+46.43	+123.81

- b. with the magnitude of the *possible* difference between the scores obtained
 iii. by Piedmontese-9 participants on Lombard-10 stimuli, and
 iv. by Emilian-11 participants on Lombard-10 stimuli,
- c. and with the magnitude of the *possible* difference between the scores obtained
 v. by Occitan-7 participants on Piedmontese-9 stimuli, and
 vi. by Lombard-10 participants on Piedmontese-9 stimuli.

Note that in point *c*, as was done in Chain A, the trilingual Occitan-8 participants' data are replaced by the Occitan-7 (France) participants' data.

The results of the tests are in Table 23. The results demonstrated that the magnitude of the difference in *a* (i.e. across the Apennines) is larger than the magnitude of the difference in *b* and *c*.

5.2.5. Interim discussion of Chain B results

The Chain B results answer the research questions in a substantially identical way as Chain A ones. Against the CCC, they provide evidence that both Piedmontese and Occitan of France are more similar to Occitan of Italy than Emilian is to Tuscan. Moreover, against the CCC, the results provide evidence that today linguistic similarity decreases more abruptly across the Apennines than across the Alps. They also show that the whole of Po valley remained united, forming with Occitan of France a continuum which is smoother than the continuum that it forms with Tuscan.

5.3. Corollary: the individuality and strong internal Abstand unity of Gallo-“Italic”

As a corollary⁵⁹ of the current study, results also provide empirical evidence that the internal *Abstand*-based unity of Gallo-“Italic” is *strong* with respect to both the bordering Tuscan and—albeit in a smaller measure—Occitan areas (see Table 24 and compare with Tamburelli, 2014).

The strong internal *Abstand*-based unity of Gallo-“Italic” and its individuality with respect to the bordering linguistic areas emerges even more clearly if we weight intelligibility mean scores per geographical distance between stimulus and participant localities, as shown in Table 25.

Table 23
Results of tests of significance and corresponding magnitudes of difference.

<i>Test for the difference between</i>	<i>Two-independent-samples t-test^a</i>	<i>Mann-Whitney U two-independent-samples test</i>	<i>Sig.</i>	<i>Magnitude of difference</i>
	<i>t(46)</i>	<i>U</i>	<i>p</i>	<i>Cohen's d</i>
a. Tuscan-12 score on Emilian-11 stimuli, and Lomb.-10 score on Emil.-11 stimuli ^b	17.131		<0.001	1.84 “huge”
b. Piedm.-9 score on Lomb.-10 stimuli, and Emilian-11 score on Lomb.-10 stimuli		249.00	0.407	–
c. Occitan-7 score on Piedm.-9 stimuli, and Lombard-10 score on Piedm.-9 stimuli		49.00	<0.001	1.43 “very large”

^a Already calculated and here repeated from Table 20. See also alternative Mann-Whitney U two-independent-samples test results in the corresponding footnote above.

^b Score difference: +44.94%; reciprocal score difference: +119.84%.

⁵⁹ That the internal linguistic unity of Gallo-“Italic” is currently strong or not is not relevant in order to test the CCC.

Table 24

Mean percentage intelligibility scores and confidence intervals. The scores of Gallo-“Italic”-speaking participants on Gallo-“Italic” stimuli are in boldface.

Chain A				Chain B			
Stimuli	Participants	M%	C.I.	Stimuli	Participants	M%	C.I.
Occitan It	Occitan Fr	51.93	45.86–58.07	Occitan It	Occitan Fr	85.71	80.86–90.57
Piedmon.	Occitan Fr	55.21	46.86–63.50	Piedmon.	Occitan Fr	49.71	44.36–55.00
Occitan Fr	Occitan It	66.36	62.86–69.86	Occitan Fr	Occitan It	73.21	68.57–77.86
Occitan It	Piedmon.	52.07	49.36–54.86	Occitan It	Piedmon.	88.99	85.55–92.43
Emilian	Piedmon.	85.43	79.71–91.14	Lombard	Piedmon.	85.12	80.86–89.38
Piedmon.	Emilian	78.00	74.64–81.29	Piedmon.	Lombard	70.24	66.72–73.76
Romagnol	Emilian	83.93	79.43–88.43	Emilian	Lombard	82.44	78.67–86.21
Emilian	Romagnol	83.36	78.93–87.79	Lombard	Emilian	83.93	80.57–87.29
Romagnol	Tuscan	36.64	31.79–41.43	Emilian	Tuscan	37.50	33.60–41.40

Table 25

Geographical distance between each stimulus locality and the corresponding participant locality, mean percentage intelligibility score, and mean percentage intelligibility score weighted per geographical distance. In each chain, the mean scores corresponding to the greatest geographical distance in the chain were kept constant (i.e. multiplier= 1) and the other scores were weighted. Mean score x multiplier = weighted score.

Chain A						Chain B					
Stimuli	Participants	Km	Mean % score	Multiplier	Weighted % score	Stimuli	Participants	Km	Mean % score	Multiplier	Weighted % score
Occitan It	Occitan Fr	43	51.93	0.1616	8.39	Occitan It	Occitan Fr	67	85.71	0.2863	24.54
Piedmon.	Occitan Fr	92	55.21	0.3459	19.10	Piedmon.	Occitan Fr	102	49.71	0.4359	21.67
Occitan Fr	Occitan It	69	66.36	0.2594	17.21	Occitan Fr	Occitan It	67	73.21	0.2863	20.96
Occitan It	Piedmon.	63	52.07	0.2368	12.33	Occitan It	Piedmon.	53	88.99	0.2265	20.16
Emilian	Piedmon.	266	85.43	1	85.43	Lombard	Piedmon.	234	85.12	1	85.12
Piedmon.	Emilian	266	78.00	1	78.00	Piedmon.	Lombard	234	70.24	1	70.24
Romagnol	Emilian	257	83.93	0.9662	81.09	Emilian	Lombard	231	82.44	0.9872	81.38
Emilian	Romagnol	260	83.36	0.9774	81.48	Lombard	Emilian	231	83.93	0.9872	82.85
Romagnol	Tuscan	57	36.64	0.2143	7.85	Emilian	Tuscan	52	37.50	0.2222	8.33

Table 26

Mutual intelligibility percentage value for each pair of Gallo-“Italic” adjacent geolects in Chain A and in Chain B.

Chain A Gallo-“Italic” geolects	Mutual intelligibility % mean	Chain B Gallo-“Italic” geolects	Mutual intelligibility % mean
Piedmontese – Emilian	81.715	Piedmontese – Lombard	77.68
Emilian – Romagnol	83.645	Lombard – Emilian	83.185

The unitary interpretation of the results from the two chains, provided by the intersection of the chains (Section 4.2), allows in turn the strong *Abstand*-based unity of Gallo-“Italic” to surface even more clearly with respect to the bordering linguistic areas. Table 26 displays the mutual intelligibility value for each pair of Gallo-“Italic” adjacent geolects in Chains A and B. *Mutual intelligibility* is defined here as “the average (mean) of the intelligibility of speaker [x] for speaker [y] and vice versa (Cheng, 1997)” (Tang and van Heuven, 2009: 710).

6. DISCUSSION

The aim of the present study was to test the critical convergence claim (CCC). The critical convergence scenario and the consequent pro-Italo- synchronological classification of present-day Gallo-“Italic” are quantitative in nature but were stated in the literature without the support of quantitative evidence, and seem to be taken as a given in much recent works. The CCC can be summarised as follows:

A linguistic convergence occurred which made the (genealogically)⁶⁰ Gallo-Romance geolects historically spoken in the Po valley and Romagna (Gallo-“Italic”) more similar to the bordering Italo-Romance geolects (south of the Massa-Senigallia line) than to the bordering Gallo-Romance geolects (Occitan). As a consequence, the synchronological continuum that Gallo-“Italic” today forms with the bordering Italo-Romance geolects is smoother than the one that it forms with the bordering Gallo-Romance geolects. This in turn equates to saying that the line dividing Gallo-“Italic” from Occitan on the Western Alps became a more substantial bundle of isoglosses than the Massa-Senigallia line.

On the basis of the link between intelligibility and *Abstand* (Section 3.1), I argued that the CCC makes some testable predictions in terms of intelligibility rates (Section 3.3). I therefore compared intelligibility between geolects spoken in each of two chains of localities running from the Occitan-speaking Alps of France to the Tuscan-speaking Apennines. Based on such an experimental design, testing the CCC predictions equates to answering the following research questions (Section 3.4): 1) In a chain of six localities, running from the French Alps to the Tuscan Apennines, passing through the Gallo-“Italic” area, which is the pair of localities between whose geolects intelligibility is the lowest? 2) Which is the sequence of the three adjacent geolects along which intelligibility lowers most abruptly?

Among the methods that have been used to measure and compare levels of intelligibility, I chose the SPIN test with high predictability sentences (Section 4). Among various advantages, the fact that its results were demonstrated to better reflect comparativists’ and dialectologists’ classifications than results obtained by other methods was crucial for this choice. In order for my study to be as representative as possible of the intelligibility situation of the Gallo-“Italic” area as a whole, it involved testing two independent chains of six localities each.

In Interim Discussion sections, I concluded that results from both chains provide evidence against the CCC. In *Abstand* terms, based on the correlation between SPIN test results on the one hand, and “lexical similarity” and “phonological correspondence” on the other (Tang and van Heuven 2009: 724), this indicates that Gallo-“Italic” did not become—at least in these two domains of the language—more similar to the bordering Tuscan geolects than to the bordering Occitan geolects. Instead, it still forms with the latter a continuum that is smoother than the one that it forms with Tuscan. Moreover, the results indicate that, from the specific standpoint of “traditional dialectology”, the stance of the pro-Gallo- “traditional dialectologists” (e.g. Hull, 1982/2017, see Section 1) is the tenable one, despite whatever linguistic convergence Gallo-“Italic” could possibly undergo towards Italian/Tuscan during the past centuries. This conclusion is based on previous literature, which found that the grouping “built on sentence-intelligibility scores [obtained by the SPIN test] . . . adequately . . . reflect[s] the taxonomy of . . . dialects postulated by traditional dialectologists” (Tang and van Heuven, 2009: 726).

A HIGHER LEVEL OF GEOGRAPHICAL GENERALIZATION. In the Interim Discussion sections, recapped just above, I presented what could be called a ‘first level of generalization’, namely the one allowed by the statistical tests themselves. In this respect, the results of the statistical analyses are sufficient to falsify the CCC, since they show that *at least* along the stripe of territory defined by each chain, the intelligibility data are incompatible with the CCC.

However, a higher level of geographical generalization of the results can be suggested, as based on some geolinguistic considerations which prevent the results of the statistical analyses from the following potential conjectural interpretation, reconcilable in part with the pro-Italo- stance:

- a. each chain (*A* and/or *B*) could be a sort of *special and thin conservative corridor* which is not representative of the current synchronological profile of some/most of the Po valley and Romagna.

⁶⁰ Specifying ‘(genealogically)’ in brackets before the label ‘Gallo-Romance’ should not be necessary, since, as presented above, labels of this form are normally intended—hence, consistently, should exclusively be intended—as members of a genealogical nomenclature (see Ibero-Romance, too). However, given the frequently ambiguous use of the label ‘Italo-Romance’, which is among the objects of this study, I shall specify it in this way where useful.

Conjecture (a) does not stand up to scrutiny. Indeed, we have seen (Section 4.2) that, based on the dialectological literature, none of the geolects investigated can be considered a ‘linguistic island’; rather, each of them is fully representative of one of the major regional continua (or regional languages) of Italy. This suggests that the trends currently identified are highly likely to be found in other similar chains of localities, having a similar direction and a similar distribution among the linguistic areas investigated (see examples in Brasca, 2023). These considerations broaden the geographical representativeness of the results from each chain.

INTERSECTING THE CHAINS. A higher level of generalization is also allowed by my choice to intersect the chains in two points (Fig. 3). Indeed, in the two chains, similarly high intelligibility scores were found between pairs of Gallo-“Italic” data points, and similarly low scores were found across the Apennines (Table 24). Thanks to the intersection of the chains, these results cannot be interpreted as a possibly fortuitous resemblance between two distinct and reciprocally poorly intelligible parallel continua (the ‘two-belts-continua interpretation’, Section 4.2), but rather as a general trend shown by geolect chains that are necessarily representative of the same smooth continuum (i.e. Gallo-“Italic”). This means that intersecting the chains made the study of each chain more informative than the study of each of two non-intersecting chains, allowing a unitary interpretation of the two chains results. Based on these considerations and on the geolinguistic ones presented in the precedent point, I consider it reasonable to generalize the linguistic conclusions drawn about the geolects investigated to the entire linguistic regions (i.e. regional languages) of which those geolects are local expressions.

A FURTHER LEVEL OF GENERALIZATION: FROM ‘OCCITAN’ TO ‘GALLO-ROMANCE’. It should be pointed out that what this intelligibility study results strictly indicate is that Gallo-“Italic” groups with *Occitan* and not with *Tuscan*, but this does not allow us to draw any conclusion about the position of Gallo-“Italic” with respect to the wider sub-groups of the Western Romance domain. However, the pro-Gallo- interpretation of the results, and more precisely the idea that *the studies that use this nomenclature, as those that I currently cited*,⁶¹ would be more consistently expected to classify Gallo-“Italic” as *still synchronologically grouped with the rest of the (genealogically) Gallo-Romance geolects*, derives necessarily from the following three facts: in such literature, (1.) Occitan is ascribed to Western Romance and in particular to Gallo-Romance; (2.) no supplementary and different synchronological grouping is proposed for it; (3.) not only between Eastern and Western Romance, but also more specifically between Italo-Romance and Gallo-Romance, *tertium non datur*⁶² from the genealogical and synchronological points of view; and (4.) Gallo-“Italic” is described as sharing all the inherited traits that are considered relevant in defining Western Romance and specifically *Gallo-Romance* (detail in Brasca, 2023). Indeed, according to von Wartburg (as an example of the synchronological pro-Italo- stance), after that Gallo-“Italic” geolects of the Po plain reached the supposedly ‘critical’ level of their convergence towards Tuscan, “. . . almost all the upper valleys of the rivers that go down the Western Alps towards the Po valley have maintained their Gallo-Romance nature. . . [and, w]hat happened on the western edge of the Po plain also happened on its northern edge” (Section 2, quote 1 [emphasis added]). As far as the results of the current study show that Gallo-“Italic” remained more similar to Occitan, and as far as points 1 to 4 above are true, these results necessarily indicate, following von Wartburg’s premises, that present-day Gallo-“Italic” “maintained its Gallo-Romance nature” in a sufficient *quantity* to make it unjustifiable, even from a mere synchronological standpoint, to separate it from the rest of Gallo-Romance and group it with Italo-Romance or ‘Italian’.

CONVERGENCE OF TUSCAN TOWARDS CISALPINE. Further considerations *quantitatively* strengthen the pro-Gallo- interpretation of the current results. As we have seen, what would purportedly determine the grouping of Gallo-“Italic” with Italo-Romance or ‘Italian’ are—citing again von Wartburg as an example of that stance—“waves of linguistic changes *progress[ing] from South to North*. . . from the 13th century, when the linguistic supremacy of Tuscan became unopposed” (von Wartburg, 1967: 130 [emphasis added]). This means that Gallo-“Italic” *convergence towards* Italo-Romance and its consequent *divergence from* Gallo-Romance would have finally made it synchronologically closer to the former than to the latter. In this respect, intelligibility results are sufficient to falsify the CCC (which is the specific object of the current research questions). However, if one wanted to have a finer-grained idea of ‘how much’ Gallo-“Italic” is synchronologically close to *Italo-Romance*, one should consider the following, on which scholars universally agree. At least from the Early Middle Ages, Tuscan accepted many Gallo-Romance loans from Cisalpine, especially as far as the Apennine Tuscan geolects are concerned (like those involved in the current study). This made them *converge towards* Gallo-Romance and *diverge from* a ‘more Italo-Romance’ original form of Tuscan (e.g. Rohlf, 1949–1966; Hull, 1982/2017). This in turn entails that, in order to determine ‘how much’ Gallo-“Italic” is synchronologically close to

⁶¹ E.g.: Lausberg (1965); von Wartburg (1967); Hall (1976); Pellegrini (1975); Posner (1996); Lopocaro (2009). This nomenclature is still standard today, in any case vastly dominant.

⁶² In Latin it means ‘a third option is not possible’.

Italo-Romance through its intelligibility to *Tuscan* speakers, a subtraction be made. Indeed, from the Tuscan participants' intelligibility score collected, we should necessarily subtract the share (not measured in this study) due to the convergence of Tuscan towards Cisalpine. These considerations *quantitatively* strengthen the pro-Gallo- interpretation of the results, which in any case are, as recalled above, sufficient in their own right to falsify the CCC.

SHARING ITALIAN AS AN OFFICIAL LANGUAGE. All the participants of the politically Italian side of the study share Italian as the language of literacy and media. It could be argued that this could facilitate the intelligibility task, suggesting to the participant the meaning of a word or of a grammatical trait which is shared by the target geolect and Italian but not by the participant's geolect (see 'language competence' as a variable in Section 4.1). This would consequently make it impossible to 'measure' objectively the effect of the mere *Abstand* on intelligibility. However, this should not be regarded as an issue for the current aims. Indeed, the current research questions, rather than seeking to 'measure' intelligibility levels in absolute terms, sought to 'compare' them. These being the current aims, some considerations do in fact make the case for the informativeness of the current results, and in particular strengthen their pro-Gallo- interpretation. On the one hand, concerning the politically Italian side of the study, the fact that all participants share Italian as a standard language should allow us to consider 'standard Italian competence' as an essentially controlled variable, and its effects as broadly minimised. Indeed, standard Italian competence is likely to have facilitated participants of all groups, including Tuscan participants. In fact, the Italian (or 'Italians') of literacy and media shares with Gallo-*"Italic"* a good deal of lexicon and structures that are not shared with various local Tuscan geolects (see e.g. Berruto et al., 1993/2002; see the "[skopa]/[gra'nata]" example in Brasca, 2023). A finer-grained control of the 'standard Italian competence' variable, however, would require considering that such a variable is likely to have had a bigger effect on the intelligibility score when the stimulus geolect was more similar to Italian. In this respect, it turns out that similarity indexes from dialectometric analyses show that standard Italian is more similar to Emilian and Romagnol than to Piedmontese and the Occitan of Italy (Goebel et al., 2019). Therefore, we can assume that the possible effect of 'standard Italian competence' is likely to have been greater on Tuscan participants' scores (obtained on Emilian and Romagnol stimuli) than on Piedmontese and Occitan participants' scores (obtained on each other's stimuli). These considerations do in fact make the case for the informativeness of the results and in particular strengthen their pro-Gallo- interpretation.

A further consideration strengthens the pro-Gallo- interpretation of the results. Indeed, Occitan-speaking participants from France do not share Italian as their standard language with the other groups. This suggests that, if it was possible to fully control 'standard Italian competence' as a variable, the differences in scores and magnitude of difference between the Alpine and the Apennine range would have been greater than the ones that actually resulted.

MEASURING VS. COMPARING INTELLIGIBILITY RATES. Gooskens (2013) points out that the percentage of correct answers varies considerably across intelligibility studies according to the method used (true/false questions, multiple choice questions, word translation, etc.). However, two other considerations in Gooskens (2013) suggest that this fact does not represent a problem for the current aims. Firstly, if, on the one hand, it is impossible to state a unique summarizing value/percentage as far as one wants to measure intelligibility "in an absolute sense. . . it may [on the other hand] be possible to *compare* the relative intelligibility of various" geolects (Gooskens, 2013: 13 [emphasis added]). In this respect, we have seen that the current research questions sought to *compare* scores and magnitudes of difference. Secondly, as to the question "which method is best?", Gooskens (2013: 10–11) answers that "the choice of method. . . depends on the precise aim of the investigation. . . [as] different tests measure different aspects of intelligibility". In this regard, we have seen that the SPIN test with high predictability sentences was found to produce results that correlate with results from traditional comparativists/dialectologists better than other kinds of intelligibility testing. Given that the current problem statement concerns a classificatory disagreement appearing in the traditional comparative and dialectological literature, this fact made the SPIN test results more informative than those obtainable by other methods.

LACK OF GALLO-*"ITALIC"* PARTICIPANTS' SCORES ON TUSCAN STIMULI. We have seen that Romagnol (in Chain A) and Emilian (in Chain B) participants' intelligibility scores on Tuscan stimuli were not collected. However, some considerations suggest that the data actually collected are sufficient to falsify the CCC. Indeed, in "most (all?) cases" presented in the literature, the asymmetry in intelligibility scores observed is best explained as a consequence of asymmetric learning (Hammarström, 2008: 35). This suggests that if one wants to determine the level of mutual intelligibility as exclusively due to objective linguistic similarity, it is appropriate to assume the lowest of the two possibly asymmetric values (collected in the Gallo-*"Italic"* and Occitan areas) as the most representative one. In this respect, it turns out that in both chains the scores obtained from Tuscan participants were significantly lower than *all* the other scores (in both directions). Analogously, all the respective magnitudes of difference were higher across the Apennines than on the western Alpine front. This very fact prevents the lack of Gallo-*"Italic"* scores on Tuscan stimuli from being a problem for the current purposes, even if one conjectured that (no longer existing) non-Italian-speaking Emilian and Romagnol participants could score higher on Tuscan stimuli than what Tuscan participants actually did on Emilian and Romagnol stimuli.

It is in any case worth recalling that, formally, the aim of the current empirical study was to test the CCC, considered as the experimental hypothesis, namely the claim ‘with the burden of proof’ (i.e., of empirical evidence). Therefore, if for whatever other possible reasons the lack of Emilian and Romagnol participants’ scores on Tuscan stimuli was regarded as a weak point of the current experimental design, this should be properly regarded as jeopardising the possibility of providing evidence *in favour of* the CCC, leaving the pro-Gallo- stance still in the strongest position, as supported by the (also pro-Italo-) scholars’ genealogical and dialectological descriptions.

EPISTEMOLOGICAL VALUE OF THE CURRENT SPIN TEST RESULTS. Scholars (McMahon and McMahon, 2005; Tang and van Heuven, 2009) consider the methods used in this study (SPIN test) and in dialectometric studies (Levenshtein algorithm used in Tamburelli and Brasca, 2018, see Section 1; Salzburg dialectometry in Goebel, 2008, see Section 2) as endowed with heuristic effectiveness in settling classificatory disagreements. This is because their results mirror those of comparativists and dialectologists in parts of the classification where these generally agree. This, in turn, would indicate that the authoritativeness of the comparative method and the traditional dialectology grants for the effectiveness itself of these—in a sense ‘ancillary’—methods, and the interpretation of their results. In this respect, what follows should be pointed out.

First, Brasca (2020, 2021, 2023) demonstrated that, notwithstanding the fact that Gallo-“Italic” has been predominantly referred to as part of Italo-Romance so far in the literature, the number of comparative and dialect studies where the pro-Italo- stance is formally grounded on *Abstand* ‘arguments’ is minute. Basically, these are von Wartburg (1967), where *Abstand* and *Ausbau* criteria appear to not be mixed, and various single statements made by some ‘mixed criterion’ scholars (Section 1) who generically refer to a convergence on Italian due to heteronomy as an *additional* pro-Italo- reason (e.g. Loporcaro, 2009—see Sections 2 and 6.1 below; also Lausberg, 1965; more examples in Brasca, 2023). This means that in these studies, the Gallo-Romance genealogical profile of Gallo-“Italic” is not contested, rather only its current pro-Gallo- synchronological profile. Moreover, we have seen that the pro-Italo- synchronological stance is quantitative in nature but the above scholars did not provide quantitative evidence (i.e. by actual measurements) supporting it. This means that, also from the synchronological standpoint, the method (SPIN) used in the current study cannot be said to oppose a largely majoritarian number of pro-Italo- *Abstand*-based (dialectological) studies, simply because no pro-Italo- *Abstand*-based quantitative evidence at all was ever provided.

Second, the abovementioned (purported) ‘ancillary’ status of these methods with respect to the traditional dialectological and comparative methods could be possibly held as far as the genealogical interpretation of their results is concerned, for which the fine-grained linguist’s knowledge of the languages compared can be considered irreplaceable in assessing, for instance, the relative chronology of changes. Instead, these methods can no longer be reductively said to be ‘ancillary’ with respect to the traditional methods as far as the synchronological interpretation of their results is concerned, where all possible traits should be considered, regardless of whether they are inherited or borrowed. Indeed, the effectiveness of the SPIN test in assessing the *Abstand* is based on the fact that its results mirror those of dialectometric methods (Section 3.1). The latter, in turn, compared to the traditional dialectological methods (see McMahon and McMahon, 2005), have the considerable advantage of avoiding the ‘arbitrary trait selection’ problem (see also Tamburelli and Brasca, 2018).

6.1. Conclusion

The results of all the statistical tests and comparisons conducted in both Chain A and Chain B studies falsified the critical convergence claim (CCC). Indeed, they did not provide any evidence that intelligibility is higher and decreases less abruptly across the Apennines than across the Occitan–Gallo-“Italic” border, nor across the western Alpine ridge. On the contrary, they provided evidence that intelligibility is lower and decreases more abruptly across the Apennines than across any line cutting the rest of the territory investigated, from the Occitan Alps in France to the Romagnol Apennines. In *Abstand* terms, the results indicated that there is no evidence that contact with the Tuscan/Italian literary language rendered Gallo-“Italic” more similar to the bordering Italo-Romance second-cousin language (i.e. Tuscan) than to the bordering uncontested Gallo-Romance sister language (i.e. Occitan). On the contrary, the results indicated that Gallo-“Italic” is still more similar to the latter than to the former. Moreover, the implicit stand in the CCC (e.g. in Wartburg, 1967), namely that Gallo-“Italic” remained linguistically united during the centuries, was borne out. However, the results indicated a grouping which is opposite to the one proposed by the pro-Italo- scholars. Indeed, rather than indicating that Gallo-“Italic” became linguistically closer to Tuscan—moving as a whole far from Occitan—the results indicated that Gallo-“Italic” remained as a whole linguistically closer to Occitan than to Tuscan.

Besides these conclusions, which are based on the intelligibility data, and are sufficient to falsify the CCC, other considerations broaden the geographical representativeness of the data themselves. Indeed, scholarly studies (describing each of the geolects investigated as a member of a smooth dialect continuum), and an aspect of the experimental design (i.e. the intersection of the two data point chains), allow one to generalize the results across a wider area than the stripes of territory linking the localities investigated.

Assuming then, based on the scholarly literature, that Occitan is (genealogically) Gallo-Romance and still synchronologically grouped with the other Gallo-Romance varieties, and that between Italo-Romance and Gallo-Romance *tertium non datur*, the results support the pro-Gallo- classification of Gallo-“Italic”.

The historical convergence of Tuscan upon Cisalpine strengthens the pro-Gallo- interpretation of the results, suggesting that the intelligibility scores across the Apennines would likely be even lower than the ones actually collected if Tuscan maintained a more original Italo-Romance *Abstand* profile.

Therefore, even assuming that the label “Italian”, often attributed to the set of Peninsular plus Cisalpine geolects, is meant as a member of a synchronological classification and not a genealogical one, like, e.g., in von Wartburg (1967) (see Brasca, 2023), the results indicate that classifying/labelling Gallo-“Italic” as “Italian”—as much as “Italo-Romance”—has to be rejected. Instead, it would be expected that scholars refer to Gallo-“Italic” simply as Gallo-Romance, in harmony with how they currently and normally define any other (genealogically) Gallo-Romance variety in non-genealogical studies, too, if it does not need a complementary synchronological grouping different from the genealogical one.

The synchronological classification of Gallo-“Italic” resulting from the current quantitative study is in line with that proposed in Hull’s (1982/2017), as resulting from his qualitative dialectological (as well as historical-linguistic) analyses. It is also in line with the Salzburger similarity and cluster (dialectometric) analyses of the AIS data, which were collected one century before the current (intelligibility) ones. With respect to the Salzburger analyses, the current study adds comparisons with Occitan of France (whose data are not included in the AIS).

The results of this empirical study ultimately respond to the following Loporcaro (2009) statement. In this concluding section, this Loporcaro statement seems apt for summarizing the pro-Italo- synchronological claim and its inconsistency, and for further clarifying the reasons for my current methodological choices:

“Among the reasons why we do not define Turinese as a Gallo-Romance dialect, and instead define it as an Italo-Romance one—like Sicilian—is also a strictly linguistic reason. Between Turin and Sicily there extends a dialectal continuum, that is a territory on which the dialects that evolved locally, starting from Latin, are linked together by a ‘chain of intercomprehensibility’: in whatever locality, the speakers of the local dialect understand those of the immediately adjacent localities, and vice versa. If such a continuum did not exist, in the presence of structural differences such as those exemplified above [common to all of Gallo-“Italic”; Loporcaro, 2009: 10–12, also 16, 82–91; detail in Brasca, 2023], we would have *much more serious difficulty* in classifying both Turinese and Sicilian as Italo-Romance.” (Loporcaro, 2009: 13 [emphasis added]).

It is important to note that in this statement no reference is made to what happens between Turinese and Occitan (of Italy and of France) in terms of intelligibility (“intercomprehensibility”). However, such a missing piece of information, in particular whether intelligibility is lower between Turinese and Occitan than between any other adjacent varieties of Italy, would be actually necessary in order for Loporcaro to draw his conclusion above. The current empirical study fills the gap made exactly by this lack of information. Indeed, it demonstrates that in the Romance *continuum* (where, by definition, geolects are supposed to be mutually intelligible to some however small extent, even across the most substantial internal bundle of isoglosses), between the Western Alpine and the Apennine bundles of isoglosses, the one across which intelligibility is lower and lowers more abruptly is the Apennine bundle. This means that, based on Loporcaro’s very line of reasoning, we have to necessarily conclude that “classifying Turinese [and more in general Gallo-“Italic”]. . . as Italo-Romance” is untenable. Instead we have to conclude that Gallo-“Italic” groups more closely with Gallo-Romance than with Italo-Romance, also in a synchronological classification.

6.1.1. Future directions

These empirical findings urge researchers to critically re-examine various statements found in the literature (see some developments already in Brasca, 2023, 2021, and 2026). For instance, they are irreconcilable with the widespread assumption that Gallo-“Italic” is a group of “Italian dialects”. On the contrary, and in line with UNESCO, Ethnologue (Lewis et al., 2013), Glottolog (Hammarström et al., 2022), and an increasing number of scholars, the results indicate that Gallo-“Italic” still shows its *Abstand* identity of a (group of) Gallo-Romance language(s), and as such separate from Italian. Moreover, the intelligibility of Gallo-“Italic” to Tuscan speakers was particularly low. This provides empirical basis for specifically excluding that standard Italian could be considered the “[protecting] roof language” for Gallo-“Italic” (as Loporcaro, 2009 assumes, already criticised in Brasca, 2021, 2023 with other arguments).

The results also showed that the internal *Abstand* unity of Gallo-“Italic” is strong (see Corollary section), and from this arises the question whether Piedmontese, Lombard, Emilian, and Romagnol are different *Abstand* languages or rather geographical varieties of the same *Abstand* language (see some initial considerations in Brasca, 2023).

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Lissander Brasca: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

Data will be made available on request.

Declaration of competing interest

I have nothing to declare.

ACKNOWLEDGEMENTS

I wish to thank Simona Scuri. I also would like to thank the people who kindly participated in the tests, those who helped me to translate and record the stimuli, and those who helped me to recruit the participants. It was a pleasure to meet so many people willing to help in scientific research. Thank you very much.

APPENDIX A

Questionnaire for the Occitan-1 participants



Enquête “Langues et dialectes”

Département de linguistique
Université de Bangor (Royaume-Uni)

Nous vous sommes reconnaissants d’avoir accepté de répondre aux questions suivantes concernant vos pratiques linguistiques, dans le cadre d’une enquête menée par le Département de Linguistique de l’Université de Bangor, au Pays de Galles (Royaume-Uni). Pour la pertinence et le succès de l’enquête, nous vous remercions d’apporter des réponses aussi précises que possible. Nous vous remercions pour votre participation.

Q1. Quelles langues étrangères connaissez-vous ? (entourez le niveau de compétence pour chaque langue mentionnée)

1. _____ niveau : école – élémentaire – intermédiaire – avancé
2. _____ niveau : école – élémentaire – intermédiaire – avancé
3. _____ niveau : école – élémentaire – intermédiaire – avancé

Q2. Quelle connaissance avez-vous du latin ? (entourez le niveau correspondant)
aucune – élémentaire – bonne – excellente

Q3. Quels langues/patois/dialectes sont utilisés dans votre famille ? (cochez d’une croix une ou plusieurs options)

- français
 patois de _____ ou d’un de ses hameaux (précisez lequel : _____)
 un autre patois/dialecte (précisez lequel : _____)
 autre (précisez lequel : _____)

Q4. Quels langues/patois/dialectes utilisaient vos parents et grands-parents avec vous pendant votre enfance ? (cochez d’une croix une ou plusieurs options)

- français
 patois de _____ ou d’un de ses hameaux (précisez lequel : _____)
 un autre patois/dialecte (précisez lequel : _____)
 autre (précisez lequel : _____)

Q5. Combien de fois vous êtes-vous rendu(e) dans les zones suivantes ? (cochez d'une croix la fréquence)

	Jamais / Rarement	Quelquefois	Souvent	Très souvent
a. Haute-Savoie				
b. Alpes piémontaises				
c. Côte ligure				
d. Côte corse				
e. Plaine piémontaise				
f. Puy-de-Dôme				

Q6. Avec quelle régularité fréquentez-vous ou avez-vous fréquenté des personnes parlant ou ayant parlé en votre présence les patois/dialectes suivants ? (cochez d'une croix la fréquence)

	Jamais / Rarement	Quelquefois	Souvent	Très souvent
a. Franco-provençaux				
b. Occitans (Piém.)				
c. Ligures				
d. Corses				
e. Piémontais				
f. Auvergnats				

Q7. Aimeriez-vous habiter dans l'une des zones suivantes ? (notez selon l'échelle de 1 à 5)

	1. Pas du tout	2.	3.	4.	5. Beaucoup
a. Haute-Savoie					
b. Alpes piémontaises					
c. Côte ligure					
d. Côte corse					
e. Plaine piémontaise					
f. Puy-de-Dôme					

Q8. Comment qualifiez-vous les dialectes/patois suivants ? (notez selon l'échelle de 1 à 5)

	1. Horribles	2.	3.	4.	5. Très beaux
a. Franco-provençaux					
b. Occitans (Piém.)					
c. Ligures					
d. Corses					
e. Piémontais					
f. Auvergnats					

Q9. Indiquez votre degré d'accord avec les énoncés suivants selon l'échelle proposée (cochez d'une croix le chiffre choisi).

Complètement en désaccord	Plutôt en désaccord	Ni d'accord ni en désaccord	Plutôt d'accord	Complètement d'accord
1	2	3	4	5

a. J'aimerais savoir parler une variété franco-provençale.	1	2	3	4	5
b. J'aimerais savoir parler une variété occitane (Piém.).	1	2	3	4	5
c. J'aimerais savoir parler une variété ligure.	1	2	3	4	5
d. J'aimerais savoir parler une variété corse.	1	2	3	4	5
e. J'aimerais savoir parler une variété piémontaise.	1	2	3	4	5
f. J'aimerais savoir parler une variété auvergnate.	1	2	3	4	5

Q10. DONNÉES PERSONNELLES

Âge : _____

Sexe : H / F

Région et province d'origine : _____

NOUS VOUS REMERCIONS DE VOTRE COLLABORATION !

APPENDIX B

Sentence meanings and their Lombard translations.

<i>Sentence meanings (Kalikow et al. 1977)</i>	<i>Lombard translations (in IPA)</i>
I cut my finger with a knife.	ma sun ⁶³ taja(da) ⁶⁴ l 'did kunt un kur'tɛl
The candle flame melted the wax.	la 'fjama da la kan'dila l a dezlen'gwa la 'sila
The baby slept in his crib.	al fju'lin al dur'miva indala sua 'kyna
How long can you hold your breath?	per kwantu 'temp ta 'rjɛset a tegni l 'fja:
Stir your coffee with a spoon.	ryga l tɔ ka'fɛ kunt uŋ ky'dʒa:
Old metal cans were made with tin.	i 'vɛdʒ tulɛt da mɛ'tal eren fa da 'stɔŋ
The beer drinkers raised their mugs.	i bevidur da 'bira tiren 'sy i sɔ bi'tʃɛ:r
Hold the baby on your lap.	'tɛŋ al fju'lin in 'skɔs
The workers are digging a ditch.	i lau'rɔnt in a'dre a skava un 'by:z
Spread some butter on your bread.	'mɛt un pɔ da by'ter in syl tɔ 'pa:ŋ
The plow was pulled by an ox.	la si'lɔrja l era tirada di 'bɔ
The shepherd watched his flock of sheep.	al pegu'rat al ky'rava l sɔ rɔf da 'bɛ
The judge is sitting on the bench.	al 'dʒydez l ɛ se'ta in sy la 'banʃka
The cut on his knee formed a scab.	al 'taj in syl sɔ dʒɛ'nɔdʒ l a fa 'sy la 'krusta
The farmer baled the hay.	al pai'zan l a muntuna l 'fɛ:ŋ
She wore a feather in her cap.	la pur'tava una 'pjyma in syl ka'pɛl
At breakfast he drank some juice.	a kula'sjun l a be'y uŋ tʃi'tʃin da 'latʃ
Raise the flag up the pole.	tira 'sy la ban'dera in syl 'pa:l
The landlord raised the rent.	al padrun da 'ka l a krɛsy l 'fitʃ
Tom fell down and got a bad bruise.	al tu'ma:z l ɛ burla 'dʒu e g ɛ vɛny 'sy una 'nɔla
The chicks followed the mother hen.	i pure'ziŋ g andaven a'dre a la 'pita
A bear has a thick coat of fur.	'l urs al g a un man'tɛl 'ɛrtig da 'pɛ:l
His boss made him work like a slave.	al sɔ 'kap l a fa lau'ra kumpɔŋ d un 'azin
The cigarette smoke filled his lungs.	al 'fym da la siga'rɛta l a impien'di i sɔ pul'mu:ŋ
The stale bread was covered with mold.	al pan 'pɔs l era kwata da 'myfa
Our seats were in the second row.	i nɔs 'pɔst eren in se'kunda 'fila
That accident gave me a scare.	kal iŋt'fident 'la al m a fa tʃa'pa un 'ku:lp
I've got a cold and a sore throat.	ɔ tʃa'pa l fre'dʒu e l mal da 'gula

References

- Allasino, E., Ferrier, C., Scamuzzi, S., Telmon, T., 2007. *Le lingue del Piemonte*. IRES.
- Ascoli, G.I., 1882–1885. *L'Italia dialettale*. Arch. Glottol. Ital. 8, 98–128.
- Bailey, K.D., 1994. *Typologies and Taxonomies: An Introduction to classification Techniques*. Sage, London.
- Bartoli, M.G., 1936. *Caratteri Fondamentali delle Lingue Neolatine*. Arch. Glottol. Itali. 28, 97–133.
- Bec, P., 1970–1971. *Manuel pratique de philologie romane*. Picard, Paris.
- Beijering, K., Gooskens, C., Heeringa, W., 2008. Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. *Linguist. Neth.* 25, 13–24.
- Bernini, G., 2025. *Avverbi posizionali in italiano: deissi orientata al parlante e al territorio*. *Verbum Analecta Neolat.* 26 (1), 35–49.
- Berruto, G., Bettoni, C., Francescato, G., Sobrero, A.A., 1993/2002. *Introduzione all'italiano contemporaneo: la variazione e gli usi*. Laterza, Roma-Bari.

⁶³ The nasalization of vowels is not marked in this list.

⁶⁴ /tjada/ is the grammatically feminine form that was used by the female speaker.

- Bradlow, A.R., Torretta, G.M., Pisoni, D.B., 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Commun.* 20 (3–4), 255–272.
- Brasca, L., 2020. Un problema nella classificazione delle lingue del gruppo gallo-“italico” all’interno della famiglia romanza. In: Trizzino, M. (Ed.), ΑΙΝΙΓΜΑΤΟΣ ΑΝΟΙΓΜΑ. Il varco della Sfinge. Nuove etimologie nell’odierno orizzonte linguistico-etnografico. *Miscellanea di studi etimologici ed etnografici in memoria di Remo Bracchi*. LAS, Roma, pp. 153–170.
- Brasca, L., 2021. Mixing methods in linguistic classification: a hidden agenda against multilingualism? The contestedness of Gallo-“Italic” languages within the Romance family. In: Tamburelli, M., Tosco, M. (Eds.), *Contested Languages: The Hidden Multilingualism of Europe*. Studies in World Language Problems. John Benjamins Publishing Company, Amsterdam and Philadelphia, pp. 59–86.
- Brasca, L., 2023. Classification of Gallo-“Italic”: Current Issues in the Literature and Proposals for a Solution. Doctoral dissertation. Bangor University.
- Brasca, L., 2026. Italian nationalist ideology and the inconsistent classification of Gallo-“Italic” contested languages. *Cogn. Stud./Études Cogn.* 2026 (26), 3487. <https://doi.org/10.116649/cs.3487>.
- Campbell, L., 1998. *Historical Linguistics: An Introduction*. MIT Press, Cambridge, Massachusetts (reprinted 2006).
- Cardinaletti, A., Giusti, G., 2025. La posizione dei pronomi clitici nella ristrutturazione: variazione diatopica e diacronica in italo-romanzo. *Verbum-Analecta Neolat.* 26 (1), 51–71.
- Cenoz, J., Hufeisen, B., Jessner, U. (Eds.), 2001. *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Multilingual Matters, Clevedon, UK.
- Cheng, C.C., 1997, February. Measuring relationship among dialects: DOC and related resources. In: *International Journal of Computational Linguistics & Chinese Language Processing*, Volume 2, Number 1, February 1997: Special Issue on Computational Resources for Research in Chinese Linguistics, pp. 41–72.
- Cohen, J., 1988. *Statistical power analysis for the behavioral sciences*. Erlbaum, Hillsdale, NJ.
- Coluzzi, P., 2007. *Minority language planning and micronationalism in Italy*. Peter Lang, Bern.
- Coluzzi, P., 2009. Endangered minority and regional languages (‘dialects’) in Italy. *Mod. Italy* 14 (1), 39–54.
- Coluzzi, P., Brasca, L., Scuri, S., 2021. Revitalizing contested languages: the case of Lombard. In: Tamburelli, M., Tosco, M. (Eds.), *Contested Languages: The Hidden Multilingualism of Europe*. Studies in World Language Problems. John Benjamins Publishing Company, Amsterdam and Philadelphia, pp. 163–182.
- Council of Europe, 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- Derwing, T., Munro, M.J., 1997. Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Stud. Second Lang. Acq.* 19 (1), 1–16.
- Devitt, M., 2008. Resurrecting biological essentialism. *Philos. Sci.* 75 (3), 344–382.
- Devoto, G., Giacomelli, G., 1972. *I dialetti delle regioni d’Italia*. Sansoni, Firenze.
- Dollaghan, C., Campbell, T.F., 1998. Nonword repetition and child language impairment. *J. Speech Lang. Hear. Res.* 41 (5), 1136–1146.
- Fanciullo, F., 2018. Problemi di sostrato. Sicuro che i suoi effetti debbano manifestarsi in copia conforme? *Lingua e Stile.* 53 (2), 199–218.
- Filipponio, L., 2010. La quantità vocalica nei proparossitoni etimologici al confine tra toscano e gallo-italico. In *XXVe CILPR Congrès International de Linguistique et de Philologie Romanes*. De Gruyter Mouton, pp. 2–67.
- Garassino, D., Loporcaro, M., Schmid, S., 2017. La quantità vocalica in due dialetti della Liguria. Conference: *Fattori sociali e biologici nella variazione biologica/Social and biological factors in phonetic variation*. Volume 3. University of Zurich, Open Repository and Archive.
- Girbau, D., 2016. The non-word repetition task as a clinical marker of specific language impairment in Spanish-speaking children. *First Lang.* 36 (1), 30–49.
- Goebel, H., Winterleitner, N., 1971. Projekt einer sprachstatistischen Auswertung von in Sprachatlanten gespeicherter linguistischer Information mit Hilfe elektronischer Rechenanlagen. *Linguist. Ber.* 14, 60–61.
- Goebel, H., 2008. La dialettometrizzazione integrale dell’AIS: presentazione dei primi risultati. *Revue de Linguist. Rom.* 285, 25–113.
- Goebel, H., Haimel, E., Smečka, P., Castellazzi, B., 2019. *Dialektkarten.ch*. <http://dialektkarten.ch/dmviewer/index.it.html>.
- Gooskens, C., 2006. Linguistic and extra-linguistic predictors of Inter-Scandinavian intelligibility. In: van de Weijer, J., Los, B. (Eds.), *Linguistics in the Netherlands*. John Benjamins, Amsterdam, pp. 101–113.
- Gooskens, C., 2007. The contribution of linguistic factors to the intelligibility of closely related languages. *J. Multiling. Multicult. Dev.* 28 (6), 445–467.
- Gooskens, C., Heeringa, W., 2004. The position of Frisian in the Germanic language area. In: Gilbers, D., Schreuder, M., Knevel, N. (Eds.), *On the Boundaries of Phonology and Phonetics*. University of Groningen, pp. 61–87.
- Gooskens, C., Heeringa, W., 2014. The role of dialect exposure in receptive multilingualism. *Appl. Linguist. Rev.* 5 (1), 247–271.
- Gooskens, C., Heeringa, W., Beijering, K., 2008. Phonetic and lexical predictors of intelligibility. *Int. J. Hum. Art Comput.* 2 (1–2), 63–81.
- Gooskens, C., 2013. Methods for measuring intelligibility of closely related language varieties. In: Bayley, R., Cameron, R., Lucas, C. (Eds.), *Handbook of Sociolinguistics*. Oxford University Press, Oxford, pp. 195–213.
- Gooskens, C., Kürschner, S., van Bezooijen, R., 2011. Intelligibility of low and high German to speakers of Dutch. *Dialectol. Special Issue II*, 35–63.

- GSPL, 2020. Conference: La presenza dei dialetti italo-romanzi nel paesaggio linguistico: Ricerche e riflessioni. Bergamo, giovedì 10 dicembre 2020, Università degli Studi di Bergamo, Dipartimento di Lettere, Filosofia, Comunicazione.
- Hall, R.A., 1976. *Proto-Romance Phonology*. Elsevier, New York, London, Amsterdam.
- Hammarström, H., 2008. Counting languages in dialect continua using the criterion of mutual intelligibility. *J. Quant. Linguist.* 15 (1), 34–45.
- Hammarström, H., Forkel, R., Haspelmath, M., Bank, S., 2022. Glottolog 4.6. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.6578297> (Available online at <http://glottolog.org>).
- Hull, G.S., 1982. *The Linguistic Unity of Northern Italy and Rhaetia*. Doctoral dissertation. University of Sydney. Revised version published in 2017. Beta Crucis Editions, Sydney.
- Jaberg, K., Jud, J., 1928–1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen.
- Jensen, J., 1989. On the mutual intelligibility of Spanish and Portuguese. *Hispania* 72 (4), 848–852.
- Kalikow, D.N., Stevens, K.N., Elliott, L.L., 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am.* 61 (5), 1337–1351.
- Keating, G.D., Jegerski, J., 2015. Experimental designs in sentence processing research: a methodological review and user's guide. *Stud. Second. Lang. Acquis.* 37 (1), 1–32.
- Kloss, H., 1967. 'Abstand languages' and 'ausbau languages'. *Anthropol. Linguist.*, 29–41.
- Kürschner, S., Gooskens, C., van Bezooijen, R., 2008. Linguistic determinants of the intelligibility of Swedish words among Danes. *Int. J. Hum. Arts Comput.* 2 (1–2), 83–100.
- Labov, W., 1990. The intersection of sex and social class in the course of linguistic change. *Lang. Var. Chang.* 2, 205–254.
- Lausberg, H., 1965. *Lingüística Románica* Tr. J. Pérez Riesco y E. Pascual Rodríguez (Madrid: Gredos 19762). Original German title: *Romanische Sprachwissenschaft*. Zweite, durchgesehene Auflage. De Gruyter, Berlin, 1963. First published by De Gruyter, Berlin, 1956¹.
- Leonardi, M., 2016. *Bilingualism or Trilingualism? Social versus linguistic views: evidence from the Germanic-speaking language group in South Tyrol (Italy)*. Doctoral dissertation.
- Lewis, P., Simons G., Fennig C. (Eds.), 2013. *Ethnologue: Languages of the World, eighteenth ed.*, SIL International, Dallas, Texas. Available at: <https://www.ethnologue.com/>.
- Loporcaro, M., 2009. *Profilo linguistico dei dialetti italiani*. Laterza, Bari.
- Maiden, M., Parry, M. (Eds.), 1997. *The Dialects of Italy*. Routledge, London, New York.
- McMahon, A., McMahon, R., 2005. *Language Classification by Numbers*. Oxford University Press on Demand (Reprinted 2011).
- Miller, G.A., Heise, G.A., Lichten, W., 1951. The intelligibility of speech as a function of the context of the test materials. *J. Exp. Psychol.* 41 (5), 329–335.
- Moberg, J., Gooskens, C., Nerbonne, J., Vaillette, N., 2007. Conditional entropy measures intelligibility among related languages. In: Dirix, P., Schuurman, I., Vandeghinste, V., Van Eynde, F. (Eds.), *Computational Linguistics in the Netherlands 2006: Selected Papers from the 17th CLIN Meeting*. LOT, Utrecht, pp. 51–66.
- Morton, J., 1969. Interaction of information in word recognition. *Psychol. Rev.* 76, 165–178.
- Moseley, C. (Ed.), 2010. *Atlas of the World's Languages in Danger*. Third ed. UNESCO Publishing, Paris.
- Munro, M., Derwing, T., 1995. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Lang. Learn.* 45, 73–97.
- Munro, M., Derwing, T.M., Morton, S., 2006. The mutual intelligibility of L2 speech. *Stud. Second. Lang. Acquis.* 28 (1), 111–131.
- Pellegrini, G.B., 1970. La classificazione delle lingue romanze e i dialetti italiani. In: *Forum italicum*, 4. SAGE Publications, London, pp. 211–237.
- Pellegrini, G.B., 1975. *Saggi di linguistica italiana: storia, struttura, società*. Cap.2: I cinque sistemi linguistici dell'italo-romanzo. Boringhieri, pp. 55–87. First published in 1973, I cinque sistemi linguistici dell'italo-romanzo. *Revue roumaine de linguistique*. Pellegrini, G.B., 1977. *Carta dei dialetti d'Italia*. Pacini.
- Pellegrini, G.B., 1992. Il 'Cisalpino' e l'italo-romanzo. *Archivio Glottologico Italiano*. 77, 272–296.
- Posner, R., 1996. *The Romance Languages*. Cambridge University Press.
- Rohlf, G., 1949–1966. *Grammatica storica della lingua italiana e dei suoi dialetti*. Einaudi.
- Romaine, S., 2008. Variation in language and gender. In: Holmes, J., Meyerhoff, M. (Eds.), *The Handbook of Language and Gender*. John Wiley & Sons.
- Sanga, G., 1997. Lombardy. In: Maiden, M., Parry, M. (Eds.), *The Dialects of Italy*. Routledge, London and New York.
- Sawilowsky, S.S., 2009. New effect size rules of thumb. *J. Mod. Appl. Stat. Methods* 8 (2), 26.
- Schmid, H., 1956. Über Randgebiete und Sprachgrenzen. *Vox Romanica* 15 (2), 19–80.
- Schotter, E.R., Bicknell, K., Howard, I., Levy, R., Rayner, K., 2014. Task effects reveal cognitive flexibility responding to frequency and predictability: evidence from eye movements in reading and proofreading. *Cognition* 131 (1), 1–27.
- Schüppert, A., Gooskens, C., 2011. Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia* 7, 119–140.
- Shukla, S., Connor-Linton, J., 2008. Il mutamento linguistico. Il mulino.
- Sibille, J., 2012. Les clitiques sujets dans le parler occitan de Chiomonte et des Ramats (Italie). *Revue de linguistique romane*. (303–304), 401–435. <hal-00952429>.
- Sibille, J., 2014, September. Contatti linguistici nell'Alta Val Susa. L'esempio di Chiomonte. In: XLVIII Congresso internazionale di studi della Società di Linguistica Italiana (SLI). Bulzoni, pp. 395–414. hal-01399448.

- Simpson, G.G., 1945. The Principles of Classification and a Classification of Mammals. *Bulletin of American Museum of Natural History*, New York, vol. 85, reprinted 1962.
- Soria, C., 2015. Assessing the effect of official recognition on the vitality of endangered languages: a case study from Italy. In: Jones, M.C. (Ed.), *Policy and Planning for Endangered Languages*. Cambridge University Press, pp. 123–137.
- Svanes, B., 1988. Attitudes and 'cultural distance' in second language acquisition. *Appl. Linguis.* 9 (4), 357–371. <https://doi.org/10.1093/applin/9.4.357>.
- Tamburelli, M., 2014. Uncovering the 'hidden' multilingualism of Europe: an Italian case study. *J. Multiling. Multicult. Dev.* 35 (3), 252–270.
- Tamburelli, M., Brasca, L., 2018. Revisiting the classification of Gallo-Italic: a dialectometric approach. *Digit. Scholarsh. Hum.* 33 (2), 442–455.
- Tamburelli, M., Tosco, M., 2021. What are contested languages and why should linguists care. In: Tamburelli, M., Tosco, M. (Eds.), *Contested Languages: The Hidden Multilingualism of Europe*. Studies in World Language Problems. John Benjamins Publishing Company, Amsterdam and Philadelphia, pp. 1–16.
- Tang, C., van Heuven, V.J., 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119 (5), 709–732.
- Thierry, G., Wu, Y., 2007. Brain potentials reveal unconscious translation during foreign language comprehension. *PNAS* 104 (30), 12530–12535.
- Tosco, M., 2008. Introduction: Ausbau is everywhere! *Int. J. Sociol. Lang.* 2008 (191), 1–16.
- Tosco, M., 2011. Between endangerment and Ausbau. In: Miola, E., Ramat, P. (Eds.), *Language Contact and Language Decay. Socio-Political and Linguistic Perspectives*. IUSS Press, pp. 227–246.
- Trager, G.L., 1934. On the classification of the Romanic languages. *Rom. Rev.* 25 (2), 129–136.
- Trudgill, P., 1992. Ausbau sociolinguistics and the perception of language status in contemporary Europe. *Int. J. Appl. Linguist.* 2 (2), 167–177.
- Valentini, A., 2012. Per una tipologia della struttura informativa: il caso delle frasi scisse in un dialetto italo-romanzo. Università di Bergamo.
- Valentini-Botinhao, C., Wester, M., 2014. Using linguistic predictability and the Lombard effect to increase the intelligibility of synthetic speech in noise. In: *Proc. Interspeech*, Singapore, pp. 2063–2067.
- Van Bezooijen, R., Gooskens, C., 2005. How easy is it for speakers of Dutch to understand Frisian and Afrikaans, and why? *Linguist. Neth.* 22 (1), 13–24.
- Varvaro, A., 2001. *Linguistica romanza: corso introduttivo*. Liguori Ed. Ristampa 2006.
- Voegelin, C.F., Harris, Z.S., 1951. Methods for determining intelligibility among dialects of natural languages. *Proc. Am. Philos. Soc.* 95 (3), 322–329.
- Von Wartburg, W., 1936. *La posizione della lingua italiana nel mondo neolatino: tre conferenze*. Keller, Leipzig.
- Von Wartburg, W., 1967. *La fragmentation linguistique de la Romania*. Klincksieck, Paris. Translated from German: (Die Ausgliederung der Romanischen Sprachräume, Verlag Francke, Bern, 1950) by Jacques Allières and Georges Straka.
- Wang, H., 2007. English as a lingua franca: Mutual Intelligibility of Chinese, Dutch and American speakers of English. LOT Dissertation, LOT, Utrecht.
- Wiener, F.M., Miller, G.A., 1946. Some characteristics of human speech. Transmission and reception of sounds under combat conditions. Summary Technical Report of Division 17, National Defense Research Committee, 58–68.
- Wolff, H., 1959. Intelligibility and inter-ethnic attitudes. *Anthropol. Linguist.*, 34–41.