



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Investigating 'Other' in Cross-Linguistic Perspective: Theoretical and Methodological Challenges

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Gianollo, C., Mauri, C. (2026). Investigating 'Other' in Cross-Linguistic Perspective: Theoretical and Methodological Challenges. Leiden : Brill [10.1163/9789004744196_009].

Availability:

This version is available at: <https://hdl.handle.net/11585/1030110> since: 2025-11-25

Published:

DOI: http://doi.org/10.1163/9789004744196_009

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Investigating ‘other’ in cross-linguistic perspective: theoretical and methodological challenges

Chiara Gianollo & Caterina Mauri (Università di Bologna)

1. Introduction

The cross-linguistic investigation of the expression of ‘other’ is made challenging by a series of factors, among which the variety of lexicalization patterns, the widespread polysemy, and the strong context-dependence of the disambiguation strategies play a particularly relevant role. In recent literature, various dimensions of variation have been discussed (Beck 2000, Breban 2003, Eguren & Sánchez 2004, Oxford 2010, Charnavel 2015, Cinque 2015, Brugè 2017, 2018 a.o.). However, cross-linguistic observations have remained sporadic and limited to cases where the existence of different lexicalizations in one language is interpreted as evidence for ambiguity of the single lexeme of another language: see e.g. Beck (2000) on German *verschieden* and *ander-* corresponding to English *different*, Tovena & Van Peteghem (2003) on the division of work between words for ‘other’ in French, English and German, or Gianollo & Mauri (2020) on German *ander-* and *weiter* corresponding to Italian *altro*.

This contribution lays out some theoretical and methodological preliminaries to the systematic cross-linguistic investigation of the dimensions of variation connected to the semantic space covered by ‘other’ and similar terms. It does so by presenting the construction and the first results of a survey aimed at investigating the language-internal and cross-linguistic extent of the ambiguity between the non-identity and the additive reading of ‘other’.

This dimension of variation has been discussed in the literature in particular for Romance languages (in Spanish by Eguren & Sánchez 2003, 2004; Gutiérrez Rodríguez 2011, Brugè 2017, 2018; in Italian by Cinque 2015, Gianollo & Mauri 2020; in Portuguese by Brugè & Giusti 2021), but is also known from other languages (see the cross-linguistic overview in Cinque 2015, and Mantenido 2020 for San Sebastián del Monte Mixtec). It concerns the possibility for the word meaning ‘other’ in a language to receive additive readings, in which the dimension of non-identity is discursively backgrounded and the use of ‘other’ is motivated as a purely anaphoric device. The Italian example in (1.a) shows the non-identity reading, (1.b) the additive reading; contrastive and additive adverbs are used to test and disambiguate the interpretations.

(1) Italian (from Gianollo & Mauri 2020: 132):

- a. *Si ritiene che il virus sia stato trasmesso agli esseri umani dal pipistrello.*
‘The virus is believed to have been transmitted to humans by the bat.’
*Un’**altra** ipotesi riconduce il salto di specie, **invece** / #**inoltre**, al pangolino.*
‘**Another** hypothesis traces the spillover back to the pangolin **instead** / #**furthermore**’
- b. *Il tracciamento dei contatti è essenziale per sconfiggere la pandemia.*
‘Contact tracing is essential in order to control the pandemics.’
*Un’**altra** misura da mettere in atto è, **inoltre** / # **invece**, il distanziamento sociale.*
Another measure that should be implemented is, **furthermore** / #**instead**, social distancing’

The main goal of our survey is to reach an improved understanding of this ambiguity in a comparative perspective: how widespread is the possibility for ‘other’ to receive additive readings cross-linguistically? Which factors contribute to contextual disambiguation? Which

strategies are used to express the additive reading by languages that cannot express it with ‘other’?

To this end, we designed a questionnaire based on the findings of Gianollo & Mauri (2020) over a corpus of Italian spontaneous conversation and we tested it on a sample of 12 Indo-European and non-Indo-European languages.

The discussion will proceed as follows. In Section 2 we present the distinction between the non-identity and the additive reading and we summarize the results of the corpus work on Italian conducted by Gianollo & Mauri (2020). The corpus study provides insights on the distribution of the two readings, in particular on the interaction with (in)definiteness and quantification, that guide the design of the questionnaire. In Section 3, we introduce the questionnaire, discussing the methodological choices we made and presenting in detail the items used for elicitation. In constructing our questionnaire, we followed a descriptive grid that does not exhaust the possible dimensions of variation, but encompasses the relevant variation points that we argue to be sufficient in order to provide an adequate empirical coverage of individual languages for the purposes of cross-linguistic comparison. The main challenge consists in finding a way to elicit the various possible ways of expressing ‘other’ in a language, taking into consideration subtle, highly context-dependent interpretive factors. Since the inferential path deciding between the non-identity and the additive reading is dynamically determined by elements in the (preceding and following) discourse context, the questionnaire format has to overcome the limitations created by the absence of a shared discourse background. In Section 4 we present the main preliminary conclusions emerging from the cross-linguistic comparison: we show that the survey, although limited at this stage to a small number of languages and speakers, validates the cross-linguistic relevance of the non-identity vs. additive ambiguity and uncovers interesting interactions with (in)definiteness and quantification, also pointing to some recurrent diachronic paths relating to ordinality and reciprocity. In Section 5 we conclude by summarizing the main findings and formulating an agenda for the next research steps.

2. Non-identity and additivity in the interpretation of ‘other’

2.1 Defining the ambiguity

The word for ‘other’ can mean ‘further, additional’ in some languages but not in others. Thus, while Italian *altro* (2) can either have a non-identity (2.a) or an additive (2.b) reading (similarly to English *other*), German *ander-* (3) only has the non-identity reading in (3.a); the additive reading must be expressed through different formal means, e.g. the adjective *weiterer* in (4) or the additive focus particle *noch* in (5) (cf. Amaral et al., this volume):

(2)	<i>Posso</i>	<i>avere</i>	<i>un’</i>	<i>altra</i>	<i>borsa?</i>
	can. 1SG	have	ART.INDEF.F	ALTRO	bag
	a. ‘May I have a different bag?’ (potential context: the one you gave me is too fragile)				
	b. ‘May I have a further bag?’ (potential context: the one I have is not enough)				

(3)	<i>Darf</i>	<i>ich</i>	<i>eine</i>	<i>andere</i>	<i>Tüte</i>	<i>haben?</i>
	can. 1SG	I	ART.INDEF.F	ANDER	bag	have
	a. ‘May I have a different bag?’					
	b. * ‘May I have a further bag?’					

(4)	<i>Darf</i>	<i>ich</i>	<i>eine</i>	<i>weitere</i>	<i>Tüte</i>	<i>haben?</i>

	can.1SG	I	ART.INDEF.F	WEITER	bag	have
	‘May I have an additional bag?’					

(5)	<i>Darf</i>	<i>ich</i>	<i>noch</i>	<i>eine</i>	<i>Tüte</i>	<i>haben?</i>
	Can.1SG	I	again	ART.INDEF.F	bag	have
	‘May I have an additional bag?’					

This observation raises two questions: the first question concerns the semantic-pragmatic relation between the non-identity and the additive meaning of ‘other’; the second question has to do with the way the two meanings are disambiguated in context in languages like Italian, which allow both readings.

Gianollo & Mauri (2020) address these questions on the basis of converging evidence coming, on the one hand, from diachronic considerations on the etymology and the diachrony of the Latin ancestor for Italian *altro*, and, on the other hand, from a corpus study of spoken Italian (spontaneous conversation in the KIParla corpus, Mauri et al. 2019).

With respect to the first question, the two readings are defined as follows: in the non-identity reading, a contrastive opposition is established between (two) compared entities that are in a relation of potential exclusion or substitution in discourse (cf. 1.a); in the additive reading, the compared entities are presented as belonging to a set that is homogeneous with respect to some salient property and they are not mutually exclusive in discourse, but incrementally added to the context (cf. 1.b).

The ambiguity emerging in certain languages is due to the fact that the word for ‘other’ realizes in fact two distinct meanings, differing with respect to the status of the comparative component involved in the contrastive opposition, which is either discursively foregrounded (in the non-identity reading) or backgrounded (in the additive reading). The absence of ambiguity in languages that do not allow for the additive reading is due to the impossibility of backgrounding the comparative component in discourse. This raises the issue of what enables, in certain languages but not in others, the backgrounding of the comparative component, as we call it here. One possibility is to derive this through a pragmatic mechanism, starting from a single lexical entry for ‘other’. Another possibility is to assume lexical ambiguity, that is, two distinct lexical entries, in one of which (additive ‘other’) the comparative component is absent.

Gianollo & Mauri (2020) assume lexical ambiguity as a working hypothesis. They propose that both the non-identity and the additive reading share an anaphoric component, which is identified as the presupposition of existence of an antecedent alternative x to the alternative y introduced by ‘other’ in the context C (cf. 6) (Gianollo & Mauri 2020: 165-168). The antecedent can be explicitly introduced in the previous context or be derived from encyclopedic knowledge and accommodated.

$$(6) \exists x [x \in C \wedge x \neq y]$$

Note that distinctness of alternatives is intrinsic to this presupposition (similarly to what is assumed for additive particles and other anaphoric elements).

Crucially, in order to create an alternative set (minimally containing x and y), the speaker will rely on inferential mechanisms that retrieve a contextually salient property P_0 common to all the members of the set (e.g. the property of being a hypothesis in (1.a), the property of being a bag in (2); cf. Gianollo & Mauri 2020: 169-171; see also Mauri & Sansò 2018). The way in which the necessary coherence relations are established is similar to what has been observed with additive particles (see the overview in Striegnitz 2005: chapter 5).

Following Beck (2000) and Charnavel (2015), Gianollo & Mauri (2020) furthermore assume that the non-identity meaning emerges from a comparative component, lexically encoded in addition to the presuppositional component in (6): in the non-identity reading, ‘other’ involves a comparison between two entities x and y on the basis of a pragmatically determined property P_1 in the context C : If P_1 holds for the term of comparison (the antecedent alternative), it cannot hold for y (cf. 7).

(7) $\lambda x \lambda y. \exists P_1 \in C, P_1(x) = 1 \ \& \ P_1(y) = 0$

The relevant property P_1 is a sub-property of P_0 (the property on which the alternative set is based). It can distinguish individuals or classes of individuals (cf. the type vs. token distinction in Cinque 2015). The way in which the relevant property P_1 is chosen is guided by discourse-prominence considerations that involve various contextual factors, as will be discussed in Section 2.2. For instance, in (2) under the reading in (2.a) the relevant property P_1 may be the fragility of the bag, and ‘other’ is used to refer to a bag that contrasts with its anaphoric antecedent in the value for this property.

In the case of the additive reading, instead, the homogeneity of the alternatives is not disputed; this homogeneity (a form of identity with respect to the value of the property P_0) becomes salient since, as seen above, a common property P_0 is the cohesive factor justifying the anaphoric link. Thus, in (2) under the reading in (2.b) the relevant property is being a bag. The only dimension of diversity (multiple distinct instantiations of a token or type) is the distinctness of alternatives encoded within the existential presupposition (cf. 6), and is hence backgrounded. In sum, in Gianollo & Mauri’s (2020) analysis, distinctness of alternatives is intrinsic to the anaphoric core of ‘other’ (in both the additive and the non-identity reading), and is distinguished from a proper comparative component that only emerges in the non-identity reading.

A fundamental discourse property of the additive reading is its ability to incrementally build a reference set, by adding into the discourse an alternative that shares the relevant property. As Cinque (2015: 22) notes, this is an essential communicative function: “If you ordered a beer and later you ask the same waiter for a beer, you are virtually forced to say “Can I have another beer?””. In other words, there is a strong communicative preference for signaling the anaphoric link.¹

Gianollo & Mauri (2020), thus, assume two different lexical entries for the non-identity and the additive meanings, which in some languages, like Italian, Spanish or English, would be realized by the same item: the entry for the non-identity reading contains the comparative component in (7), the entry for the additive component doesn’t. Ideally, a unified treatment in which one of the two interpretations is derived from one and the same lexical entry under specific pragmatic conditions is of course preferable (cf. Del Prete, this volume). However, the fact that cross-linguistically not all words for ‘other’ can receive an anaphoric reading points to an irreducible difference in some languages: future research will have to establish whether this difference can be attributed to pragmatic factors systematically inhibiting the ambiguity, e.g. to lexical blocking, thus supporting a unified treatment. It is also plausible that the additive reading may emerge diachronically as a form of bleaching connected to grammaticalization (cf. Section 4 and Gianollo and Mauri *forthcoming*; see also Oxford 2010 on English *other*).²

¹ In signaling the anaphoric link, homogeneity is the prominent dimension: see Milner’s (1984: 62) observation on French, commented by Bortolussi (2015: 173) in his study of ‘other’ in Latin, that in ordering a second coffee, it is equivalent to say *Donnez-moi la même chose!* ‘Give me the same thing’ or *Donnez m’en un autre!* ‘Give me another one’.

² In the case of Romance, historically, both the deictic-anaphoric and the comparative component are present since the beginning in Latin *alter* ‘other in a binary set’, which is traced back to a demonstrative root combined with

In this contribution we do not discuss this theoretical issue further. Our aim is to reach an improved understanding the language-internal and cross-linguistic extent of the ambiguity between the non-identity and the additive reading and to construct a tool for comparative research. In this respect, nothing hinges on the decision between a lexical ambiguity analysis and a unified one. The discussion surrounding this issue, however, helped us define the content of the non-identity and the additive reading.

2.2 Factors of disambiguation

The choice between the non-identity and the additive reading is fundamentally a choice concerning the communicative relevance of non-identity vs. identity of the involved alternatives with respect to a selected property, a choice that is strongly dependent on contextual cues of various nature. This point directly connects with the second question emerging from the Italian corpus data in Gianollo & Mauri (2020), namely the way in which the non-identity and the additive meaning are disambiguated in context.

The results from the corpus study on Italian show that the disambiguation of the non-identity and additive readings is systematically related to five main clusters of factors: (in)definiteness; quantification; cardinality of the set; ordering within a set; discourse prominence of the comparative component.

In what follows, we synthetically present these factors, since they have been instrumental in guiding the construction of the questionnaire for the cross-linguistic survey (Section 3). See further Amaral et al., this volume, for discussion and validation of these factors on a different, broader corpus by means of computational methods.

Concerning (in)definiteness, the corpus data from Italian indicate that the ambiguity is higher if *altro* is preceded by an indefinite article. In that case, the occurrences are distributed quite evenly between the non-identity and the additive reading (respectively, 58% and 42%). In the cases, instead, where *altro* is preceded by the definite article or by no article, the non-identity reading is predominant, reaching 94% in case of co-occurrence with a definite article. Overall, in the corpus, the additive reading occurs with a definite article in only 11% of cases, while in 89% of occurrences the additive reading is found in combination with an indefinite article or in a bare nominal phrase (the latter mostly represented by plural indefinite nominal phrases).

The corpus examples in (8) show how the combination of indefinite article + *altro* is compatible with either an additive (8.a) or a non-identity (8.b) interpretation (disambiguated in 8.a-b by the surrounding discourse context). In the combination of definite article + *altro* (8.c), instead, the incremental component is absent and what is foregrounded is the token diversity.

(8) Italian (KIParla corpus)

a. TO084:	<i>ho</i>	<i>sempre</i>	<i>vissuto</i>	<i>in</i>	<i>quella</i>	<i>casa</i>	<i>lì</i>			
	have.1S G	always	lived	in	that	house	there			
	<i>poi</i>	<i>vabbè</i>	<i>ho</i>	<i>un'</i>	<i>altro</i>	<i>casa,</i>	<i>la</i>	<i>casa</i>	<i>al</i>	<i>mare</i>
	then	well	have.1S G	ART.INDE F.F	ALTRO :F	house, e,	ART.DEF .F	house e	by	sea

the suffix *-tero-* expressing binary contrast and found in various Indo-European languages as comparative morpheme. English *other* has the same etymology. See Gianollo & Mauri (2020: 136-138) for discussion and references.

	I have always lived in that house there, then well I have another house, the house by the sea (another house = an additional house)									
(TOD2004)										

b. TO058:	[...]	<i>perché</i>	<i>ehm</i>	<i>quando</i>	<i>quando</i>	<i>stavamo</i>	<i>(li)</i>			
		because	er	when	when	stay:IND.I MPF:1PL	(there)			
	<i>avevamo</i>	<i>un'</i>	<i>altra</i>	<i>cas</i> <i>a</i>	<i>e</i>	<i>poi</i>	<i>abbiamo</i>	<i>preferito</i>	<i>prendere</i>	<i>questa</i> .
	have:IND.I MPF:1PL	ART.IN DEF.F	ALT RO:F	house	and	then	AUX:IND.P RES:1PL	preferred	take	this
	because er when when we were (there), we had another house and then we preferred to take this (another house = a different house)									
(TOD2013)										

c. BO073:	[.. .]	<i>lui</i>	<i>viveva</i>	<i>nell'</i>	<i>altra</i>	<i>casa</i>	<i>pe</i> <i>r</i>	<i>e</i> <i>h</i>	<i>cerc</i> <i>are</i>	<i>d</i> <i>i</i>	<i>ave</i> <i>re</i>	[.. .]
		he	live:IND:IMP F:3SG	in:ART.D EF:F	ALTR O:F	house	to	e h	try	t o	have	
	<i>pe</i> <i>r</i>	<i>dicia</i> <i>mo</i>	<i>un</i>	<i>distacco</i>	<i>meno</i>	<i>traumatico</i>						
	to	let's. say	ART.INDEF. M	detachment	less t	traumatic						
	he lived in the other house to eh try to have to let's say a less traumatic detachment											
(BOD2005)												

In Gianollo & Mauri (2020) we classified instances such as (8.c) as non-identity readings. Context is thus less relevant to disambiguate the meaning of definite ‘other’, and, as will become clear in Section 3.2.1 and Section 4, this played a role when we had to identify the type of situation that could be employed in the questionnaire to elicit occurrences as (8.c). When ‘other’ occurs in a definite nominal phrase, the presuppositionality of the set of alternatives indeed makes incrementality not-at-issue, ultimately causing a neutralization of the ambiguity. In these cases, the uniqueness associated with the interpretation of the definite article in fact leads to a bipartition of the set: there are only two alternatives, the antecedent and the one introduced by ‘other’, and they are already given, hence the set is already established and ‘other’ cannot incrementally contribute to its construction. Also in case of multiple individuals belonging to the set (plural nominal phrases), the nominal phrase containing ‘other’ behaves as a form of complement anaphora (Moxey & Sanford 1993, Nouwen 2003). We will come back to this point below, when discussing dual sets, and in Section 4.

When *altro* occurs in quantified nominal phrases, in the corpus there is an almost systematic association between the presence of quantifiers or numerals imposing a cardinality >2 and an additive interpretation, cf. (9).³

³ Cinque (2015: 22-23) observes that with measures in general the additive reading is preferred (see also Amaral et al., this volume), and that this in Italian correlates with a natural position of *altro* before the cardinal numeral. Cinque (2015) discusses aspects of the additive vs. non-identity ambiguity in terms of a distinction between tokens

(9) Italian (KIParla corpus)

TO028:	<i>eh</i>	<i>poi</i>	<i>ci sono</i>	<i>altri</i>							
	eh	the n	there.a re	ALTRO:P L.M							
	<i>altre</i>	<i>tre</i>	<i>doman de,</i>	<i>una</i>	<i>è</i>	<i>se</i>	<i>quan do</i>	<i>siamo</i>	<i>i n</i>	<i>erasm us</i>	[. .]
	ALTRO: PL.F	thr ee	questio ns,	one	i s	whet her	when	be.IND.PRES .1PL	i n	erasm us	
	Eh then there are three more questions, one is whether when we are in Erasmus [...]										
(TOA10 02)											

Contexts like (9) are the prototypical cases in which ‘other’ is used for incremental set construction, that is, to add to the discourse context alternatives that are anaphorically linked to previously introduced discourse referents. In such discourse contexts, what is foregrounded is the incremental addition, as in (9), where the questions share a relevant homogeneous property (their being questions), which is nonetheless compatible with further sub-differentiations (the questions will each have to do with different matters) that are non-at-issue. It must be said that most of the quantified occurrences in the corpus have an indefinite interpretation. As will be argued in Section 3.2.2, in the light of the ambiguity neutralization observed in definite contexts (see above), in the questionnaire we decided to also elicit occurrences of *definite* quantified nominal phrases, in order to monitor the interaction between quantification and definiteness.

Concerning the cardinality of the set, the relevance of this parameter was originally suggested by the fact that Latin lexicalizes the difference between ‘other in a binary set’ (*alter*) and ‘other in a set of cardinality greater than 2’ (*alius*) (cf. Bertocchi & Orlandini, this volume), and that in the Classical Latin data analyzed in Gianollo & Mauri (2020) the ambiguity between the non-identity and the additive reading is often neutralized in dual contexts. These observations on Latin correlate with the general effect of definiteness discussed above: ultimately, even in a language without articles like Latin, otherness in a presupposed binary set most naturally maps to definiteness, by virtue of the resulting uniqueness of the introduced alternative.⁴

The keenness of ‘other’ to enter binary relations is also observable in its cross-linguistically frequent use in reciprocal constructions and reciprocal pronouns (cf. Italian *l’un l’altro*, English *each other*, German *einander*, Greek *alilon* historically deriving from the reduplication of the root for ‘other’; see Kulikov 2014). This is in turn connected to the ordinal value that the word for ‘other’ assumes in various languages, where it is also used to express the meaning ‘second’ in a numerical, hierarchical, or temporal series (see Section 4 and Kaiser, this volume; see also Bar-Asher Siegal 2015 for the relation between ‘other’ and ‘second’ in the Modern Hebrew reciprocal ‘numeral’ construction). Also in this case, the additive value is predominant, since the function of ‘other’ consists in incrementally building a reference set containing elements that have a relevant property in common. The use of ‘other’ in incremental lists, enumerations and as indication of hierarchical succession within a series is a particularly clear example of its correlative nature, highlighted by Van Peteghem (2000).

vs. types. We believe that the token vs. type distinction is a further dimension of variation in the interpretation of *other*, which cross-cuts the distinction we are interested in. See further Section 3.2.

⁴ Gianollo & Mauri (2020: 147) model the duality condition of *alter* as an exhaustivity presupposition: the considered alternatives *x* and *y* exhaust the relevant alternative set. This way, uniqueness is ensured for the anaphorically linked alternative.

Finally, the discourse prominence of the comparative component is a factor that is less easy to systematically capture in corpus annotation but favors a non-identity interpretation. In general, the prominent inferential path is dynamically determined by elements of various nature in the surrounding discourse context. A non-identity reading is especially favored by elements (predicates, adverbs, particles) that signal contrast, opposition, choice. One of the clearest contexts for the non-identity reading involves substitution of one alternative with another (cf. 2.a and 8.b). In many cases, however, the factors leading to disambiguation have to do with world knowledge and are difficult to classify in a systematic way: cf. (10), where the disambiguation towards a non-identity reading in (10.a) is due to the impossibility of living multiple lives, and the additive reading in (10.b) is due to the fact that people typically do not swap babies.

(10) Italian

a.	<i>Gianni</i>	<i>vuole</i>	<i>un'</i>	<i>altra</i>	<i>vita</i>	
	Gianni	want:IND.PRES.3SG	INDEF:F	ALTRO:F	life	
‘Gianni wants another life’						

b.	<i>Gianni</i>	<i>vuole</i>	<i>un</i>	<i>altro</i>	<i>bambino</i>	
	Gianni	want:IND.PRES.3SG	INDEF:M	ALTRO:M	child	
‘Gianni wants another child’						

The Italian corpus shows interesting cases of online disambiguation, such as (11) *qualcosa d’altro e qualcosa di più* ‘something else and something more’, where the speaker, by adding *qualcosa di più* ‘something more’ in the second conjunct provides a cue to the non-identity interpretation of the preceding conjunct containing the word for ‘other’, *qualcosa d’altro* ‘something else’.

(11) Italian (KIParla corpus)

BO1 00:	[.. .]	<i>la</i>	<i>capacità</i>	<i>di</i>	<i>mett ere</i>	<i>in</i>	<i>relazi one</i>	<i>le</i>	<i>informazi oni</i>	<i>pe r</i>	<i>otten ere</i>
		DEF:F	abilit y	of	put	in	relatio n	DEF:P L.F	informatio n:PL	to	obtai n
		<i>qualco sa</i>	<i>d’</i>	<i>altro</i>	<i>e</i>	<i>qualco sa</i>	<i>di</i>	<i>più</i>			
		someth ing	of	ALTR O:M	and	someth ing	of	more			
“the ability to connect information to get something else and something more (BOD1007)”											

3. The questionnaire: aims, methods and design

3.1 Methodological rationale for a questionnaire

As seen in Section 2, Gianollo & Mauri (2020) identify a number of features that correlate to a lower degree of ambiguity for Italian *altro*. Our hypothesis is that cross-linguistic variation in the encoding of the semantic domain of ‘other’, in particular with respect to the ambiguity between the non-identity and the additive reading, could be sensitive to these features. Therefore, we designed a questionnaire aimed to elicit which expressions for ‘other’ are typically used in contexts having different properties with respect to definiteness,

quantification, ordinality and duality, in order to identify the cutoff points at which languages may vary (i.e., points of variation).

In so doing, we concentrated the elicitations on the lexeme that, in a language, contributes to the establishment of reference by imposing a non-identity requirement with respect to the anaphoric antecedent, and we explored its ability to also receive additive interpretations, as well as alternative realizations of the additive reading. The contribution of ‘other’ to the establishment of reference is the most important factor that leads to treatments of ‘other’ in the literature as an element of the functional lexicon (part of a complex determiner for Van Peteghem 1997, Breban et al. 2011). We left aside, instead, lexemes that have the predicative function of comparing individuals and stating either non-identity of individuals or non-identity of the kinds to which the individuals belong (following Beck’s 2000: 102-103 definition of *different*). For this reason, our questionnaire does not contain elicitations for the use of the word for ‘other’ in predicative structures.

Before we describe the questionnaire in detail with respect to the linguistic features it explores (Section 3.2), let us spell out some methodological considerations that guided our work.

First, at the moment the questionnaire is designed to be used with expert informants. Although no specific background in linguistics is necessary, the informants are expected to have a high degree of literacy in their native language and in English and to be familiar with metalinguistic reasoning.⁵ This choice is guided by the questionnaire’s main aim at this stage of research, which is to explore the cross-linguistic relevance of some selected points of variation. We understand this stage as a necessary preliminary step to a broader survey with typological import, also involving non-trained informants. In its current form, the questionnaire has the advantage that it can be filled out by informants remotely with no need for the researcher’s presence. We believe that the researcher’s presence would allow for it to be used also with non-trained informants, but we wish to underline that the questionnaire, in its present form, is not designed to be a tool for semantic fieldwork proper.

A second point has to do with the exhaustivity of the collected information for the individual languages, which is connected to the way the questionnaire has been distributed. As it was filled out by the informants remotely, we cannot be sure that informants always provide all the possible ways of saying a certain thing (namely, of utilizing all the potentially relevant ‘other’ expressions). In order to limit the possible drawbacks related to this point, we added a note in the instructions telling the informants that we expect that there may be multiple ways of reacting to a certain context, and asking them to provide us with all the ways they can think of; moreover, at the end of the questionnaire we added three additional questions that invite the informants to review their output and integrate it if necessary (cf. Section 3.2).

A further crucial point that guided the design of our questionnaire is the strong context-dependence of the interpretation of ‘other’. The elicitation tasks have to be embedded in a setting that allows for the precise determination of the context of utterance relative to the aspects of variation that we want to investigate (that is, the non-identity vs. the additive reading). After considering and testing various solutions, we opted for the addition of introductory paragraphs that describe the utterance contexts and provide the cue to the informant’s elicitation, using English as a metalanguage. This solution has two disadvantages: it restricts the use of the questionnaire to individuals with high literacy in English and it can potentially bias the informant’s output due to interference with the metalanguage (but see Matthewson 2004: 271 for the low impact of this kind of interference). This is particularly true

⁵ This also means that, for the pilot study presented here, we were able to largely disregard inter-speaker variation, since we expect expert informants to provide us with all the possible ways to express ‘other’ in a certain context in their language. See further on in this section for the explicit prompts contained in the questionnaire in this respect.

for the cases in which the cue to the informant's elicitation is similar to a translation task. We tried to reduce the impact of this aspect by never presenting the request as a translation task. What we ask, after presenting the utterance context with a short paragraph and providing the elicitation cue, is: "tell us how you would say the same thing in your language". In some cases, we left a blank in the English-language elicitation cue, asking the informants to fill this blank prior to providing their native language equivalent of the elicitation cue. This way, we tried to avoid a mechanic translation of English *other*, allowing the informant to think of various possible ways to express the concept in their language (something that, as we will see, is particularly important in the case of additive interpretations). The items used in the questionnaire and a more detailed description of their structure are provided in Section 3.2. A further measure to reduce the negative impact of translation was to build a certain degree of redundancy in the questionnaire, in order to enable the researcher to double-check certain interpretations. While we are aware of the drawbacks of using elicitation tasks in form of translations, we believe that they are an effective way of carrying out the task we are interested in at this research stage: the validation of the cross-linguistic significance of the points of variation that we selected on the basis of our research on Italian. We start from the best possible data (large amount of spoken conversation) and extract from this data the diagnostics for cross-linguistic investigation, which for feasibility reasons are employed in a questionnaire containing contextualized elicitation tasks.

A final important methodological point concerns the choice of the points of variation to be investigated by means of the questionnaire. We chose to focus on the contrast between the non-identity and the additive reading, which was the point of departure for the investigation of the Italian corpus in Gianollo & Mauri (2020). The main questions the questionnaire wants to answer is: does the language allow for both non-identity and additive readings of the word(s) for 'other'? if not, what are the strategies to express the additive reading? Answering this research questions is only part of a larger cross-linguistic investigation on the semantics of 'other'. Our impression is that such larger investigation can only be attempted once some basic aspects concerning the lexical and structural means used in the semantic domain of 'other' in a certain language are covered, and this is what we aim to do with our questionnaire. This choice allows us to keep the questionnaire to a manageable length. The questionnaire can be used as preparatory to further in-depth investigations of individual languages that will necessarily take into consideration broader aspects of their syntax and semantics, such as general constraints on word order and on the expression of quantification and reference. Further aspects of variation that need to be investigated by means of in-depth individual language description and were deliberately left out of our questionnaire are, in particular, the respective position of the word for 'other' and the cardinal numerals (cf. Cinque 2015); the formal possibilities for explicitly expressing a term of comparison (cf. Beck 2000); a systematic investigation of relational adjectives related to "otherness" with a more predicative function (cf. Van Peteghem 1997, Beck 2000, Tovena & Van Peteghem 2003); the possibilities of co-variation in bound readings and the relation with reciprocity and distributivity in this respect (Carlson 1987, Moltmann 1992, Beghelli & Stowell 1997, Beck 2000).

3.2 Points of variation

The questionnaire is organized in 15 contextualized elicitation tasks, followed by three questions addressing the informants' metalinguistic competence. The main challenge in constructing the questionnaire consisted in appropriately representing and controlling the contextual factors that emerged as relevant for the disambiguation in the corpus study on Italian. During the first test trials for the elicitation tasks, we realized that the dichotomy between the non-identity and the additive readings was not fine-grained enough to present

unambiguous contexts to the informants. We therefore decided to also include the token vs. type dimension, since it allowed us to define the discourse contexts more precisely. The token vs. type dimension, which is central in Cinque’s (2015) approach, rests on observations by Muliačić (1970) and Tekavčić (1972), who remarked that ‘other’ can indicate diversity of individual (token) or qualitative diversity (type). As seen in Section 2.2, token diversity does not necessarily amount to a non-identity reading: distinctness at the level of the individual alternatives can be factored into the discourse as a form of identity at the level of a superordinate property (different individuals belonging to the same type), resulting in an additive reading.

We therefore decided to construct our contexts applying the dimension of qualitative comparison as ‘± identical type’ (abbreviated as ± id). This way, we included in the construction of the items the factor of disambiguation that in Section 2.2 we identified as the discourse prominence of the comparative component.

A further decision we took was to elicit additive readings in their clearest form as resulting in incremental set construction: hence, the relevant contextual dimension was coded as ‘± incrementality’ (abbreviated as ± incr). Where possible, we tried to create contexts biased towards a [- incremental] reading in their clearest form as contexts involving substitution. ‘Other’ expressions may indeed refer to elements belonging to the same or a different type, thus involving qualitative comparison, but at the same time they may refer to elements in addition or in substitution, thus involving the construction of the set and its internal composition (i.e. in the case of substitution, the total number of items in the set ultimately remains the same).

The resulting scheme is shown in Table 1:

<i>OTHER</i>		Qualitative comparison	
		IDENTICAL TYPE	DIFFERENT TYPE
Set construction	INCREMENTAL	[+id +incr]	[-id +incr]
	NON-INCREMENTAL	[+id -incr]	[-id -incr]

Table 1. ‘Other’: the two dimensions of qualitative comparison (±id) and set construction (±incr).

The prototypical non-identity reading corresponds to the combination [-incremental, -identical type], whereas the prototypical additive reading corresponds to the combination [+incremental, +identical type]. However, additivity can also be found in a discourse setting in which diversity of type correlates with incremental set construction [+incremental, -identical type]: in this case, we want to test if languages that do not allow ‘other’ in prototypical additive readings use it more readily in contexts characterized by diversity of type. Conversely, the context [-incremental, +identical] allows us to test the tolerance of the word for ‘other’ to contexts involving contrast between tokens of the same type, and to keep apart the token dimension and incrementality (differing, in this respect, from Cinque 2015).

In what follows, we describe in detail the various contexts and utterances employed in the questionnaire to elicit the verbalization of ‘other’, by avoiding as much as possible direct translations, in those cases where it could lead to a relevant bias.

For each sentence, the informant is provided with:

- a) a CONTEXT, which can be:
 - a1) biased towards an incremental or non-incremental reading
 - a2) biased towards an identity or non-identity reading
 - a3) neutral, i.e., not biased towards any particular reading;

- b) an ENGLISH UTTERANCE that is likely to be produced in that context. The utterance may be:
 - b1) complete, i.e., with no empty slots – for those cases in which we foresee a low risk of translational bias
 - b2) incomplete, i.e. with one or more empty slots to fill – for those cases in which it is advisable to avoid a direct translation of English ‘other’ expressions. The identification of the appropriate English expression is left to the informants;
- c) a BLANK SPACE where the informants are asked to use *their own language* and write the utterance that is likely to be produced in that context, ideally providing also interlinear glosses or at least word-to-word translations. As stated in the instructions, the utterance should not consist of a literal translation from English but should rather correspond to a natural way of expressing the meanings conveyed by the English sentence.

The reason why we decided to include an English utterance, to be taken as a starting point for the identification of the appropriate expression in the informant’s language, is to favor the production of ‘other’ expressions, and to limit the risk of random production of the most diverse utterances that the specific context would allow.

After the 15 items constructed this way, the informants are asked to answer three additional questions, that are conceived to make them aware of the aims of the questionnaire and thus activate a control phase, bringing them back to the answers just given. Such a control phase is aimed to improve the reliability, consistency, and comprehensiveness of the answers provided. The three metalinguistic questions are:

- How would you translate in your language English *other*?
- Were you able to use this (or these) expression(s) in the questions above? Please go back and add it/them, if necessary.
- Can you think of other contexts, not covered in the above cases, where you could use the expressions you listed?

Before we describe the items in detail, it is important to remark that the various elicitation items are best proposed in a random order, to avoid – or at least to limit – the priming effect. Therefore, the order that will be followed below in motivating and discussing each question does not reflect the order eventually found in the questionnaire; rather, it clusters the items according to the points of variation that we expect to find across languages. The items are assigned an identification number according to this order.

3.2.1 Definiteness

Let us start by discussing the items aimed at capturing the relation between ‘other’-words and (in)definiteness. We first of all aim to know if the language has a definite and an indefinite article at all and whether articles would be employed in a neutral context, as shown in Item (1) and (2):

ITEM (1)

AIM - indefinite article	Neutral CONTEXT:	You are in a supermarket and you realize that you need a bag, you ask the shop assistant:
	ENGLISH:	a. <i>Please, can I have bag?</i> b. <i>Please, can I have bags?</i>
	tell us how you would say the same thing in	a.

	YOUR LANGUAGE:	b.
--	----------------	---------

ITEM (2)

AIM - definite article	Neutral CONTEXT:	Mary and Frank are in a shop
	ENGLISH:	a. <i>Frank saw a bag. The bag was black and white.</i> b. <i>Frank saw two bags. The bags were black and white.</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

Indefinite reference has been argued not to play a significant role in disambiguating between the additive and the non-identity meaning. According to the findings of Gianollo & Mauri (2020) summarized in Section 2, the occurrences where ‘other’ is combined with indefinite determiners are indeed those with the highest degree of ambiguity. We therefore aim to know which ‘other’ expressions would be used with indefinite (both singular and plural) phrases.

Given the expected ambiguity, we provide the informant with five different context types according to the dimensions shown in Table 1 (both a singular and a plural version is provided, in order to control for the effect of number):

- a *neutral* context, i.e. a context that is not biased towards the incremental (vs. non-incremental) or the identity (vs. non-identity) meaning but is compatible with either value for each parameter, as in Item (3). In Item (3), the issues talked about can either be issues that are already on the table [- incremental] or issues that are going to be added [+ incremental]; they can be issues of the same type or issues of a different type:

ITEM (3)

AIM - Indefinite referent	Neutral CONTEXT:	You are discussing with your colleague about who is going to write to John, you say
	ENGLISH:	a. <i>I'll do it, because I was going to write him anyway about another issue</i> b. <i>I'll do it, because I was going to write him anyway about other issues</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

- A context *biased* towards identity and incrementality, which clearly favors an additive meaning and introduces a referent belonging to the same type (i.e. in addition and identical), as in Item (4):

ITEM (4)

AIM - Indefinite referent	[+id +incr] CONTEXT:	You are in a supermarket and you realize that the bag that the shop assistant gave you is not enough, you ask the shop assistant:
---------------------------	----------------------	---

	ENGLISH:	a. <i>Please, can I have bag?</i> b. <i>Please, can I have bags?</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

- A context *biased* towards identity and non-incrementality, which favors a substitution interpretation and introduces a referent belonging to the same type (i.e. in substitution and of identical type), as in Item (5):

ITEM (5)

AIM - Indefinite referent	[+id -incr] CONTEXT:	You are in a supermarket and you realize that the bag that the shop assistant gave you is broken, you ask the shop assistant:
	ENGLISH:	a. <i>Please, can I have bag?</i> b. <i>Please, can I have bags?</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

- a context *biased* towards non-identity and non-incrementality, which favors a substitution interpretation and introduces a referent belonging to a different type (i.e. in substitution and different), as in Item (6):

ITEM (6)

AIM - Indefinite referent	[-id -incr] CONTEXT:	You are in a supermarket and you realize that the type of bag that the shop assistant gave you is too small and too fragile, you ask the shop assistant:
	ENGLISH:	a. <i>Please, can I have bag?</i> b. <i>Please, can I have bags?</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

- a context *biased* towards non-identity and incrementality, which favors an additive meaning and introduces a referent belonging to a different type (i.e. in addition and different), as in Item (7):

ITEM (7)

AIM - Indefinite referent	[-id +incr] CONTEXT:	When a friend asks you what you like about travelling, you say:
	ENGLISH:	a. <i>I like to open up to another way of living</i> b. <i>I like to open up to other cultures</i>

	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.
--	--	--------------------

The four *biased* contexts allow us to identify the expressions employed for the four possible combinations of qualitative comparison (identical type vs. different type) and set construction (incrementality vs. non-incrementality), as summarized in Table 2.

OTHER		Qualitative comparison	
		IDENTICAL TYPE	DIFFERENT TYPE
Set construction	INCREMENTAL	Item (4)	Item (7)
	NON-INCREMENTAL	Item (5)	Item (6)

Table 2. ‘Other’: the two dimensions of qualitative comparison (\pm id) and set construction (\pm incr) as represented in the questionnaire’s elicitation items.

While great potential ambiguity is observed for the interpretation of ‘other’ in indefinite nominal phrases, based on the corpus evidence in Gianollo & Mauri (2020), we expect definiteness to correlate with a lower ambiguity, favoring a non-identity reading of ‘other’. As already argued in Section 2.2, definiteness indeed implies that the alternative introduced by ‘other’, or the set to which it belongs, is given, hence making incrementality less relevant. This means that, in a language having different lexemes for the additive and the non-identity meanings, we expect the ‘non-identity-*other*’ lexeme to be more likely with definite referents. As a consequence, given the lower ambiguity, it is not necessary to provide several context types as we did for indefinite ‘other’ phrases, but it is sufficient to elicit ‘other’ expressions referring to definite referents in neutral contexts, as in Item (8):

ITEM (8)

AIM - definite referent	Neutral CONTEXT:	You are discussing with your colleague about who is going to call John, you say
	ENGLISH:	a. <i>I’ll do it, because I’m going to call him anyway about the other issue</i> (you know what I am talking about) b. <i>I’ll do it, because I’m going to call him anyway about the other issues</i> (the ones we discussed before)
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

More examples with definite referents will be discussed in the next sections, together with quantification, ordinality, and duality.

3.2.2 Quantification

Definiteness interacts with quantification, which is another feature that has been argued to play a role in predicting ‘other’ interpretations. When ‘other’ expressions refer to quantified referents, they are more likely to receive an additive interpretation (Gianollo & Mauri 2020; cf. Section 2.2). However, this could be a specificity of Italian, or it could be due to the high

frequency of indefinite quantified referents, compared to definite ones, in the corpus data. Therefore, this is a further potential point of variation that needs to be addressed in the questionnaire.

We prepared three questions aimed at eliciting the use of ‘other’ expressions with quantified referents, both definite and indefinite, distinguishing between exact quantification (sets of cardinality $|2|$, as in Item (9), and $|>2|$, as in Item (10)) and indeterminate quantification (e.g. *many*, *a few*, as in Item (11)). In these cases, the context is not biased towards the non-identity or the additive interpretation, but two English utterances are provided, in Item (9) and (10), the first with indefinite referents and the second with definite ones, in order to monitor the interaction between definiteness and quantification.

ITEM (9)

AIM - exact quantity (=2)	Neutral CONTEXT:	It is very late in the evening, you are still in front of the pc doing your email, you think
	ENGLISH:	a. <i>I'll write two emails, and I'll be done</i> b. <i>I'll write the other two emails, and I'll be done</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

ITEM (10)

AIM - exact quantity (>2)	Neutral CONTEXT:	After a lesson, you prepared five questions to ask, but you only managed to ask two of them... you say
	ENGLISH:	a. <i>I have three ... questions to ask</i> b. <i>I only asked two out of five questions I had, I would like to ask you the... three, can we stay a little longer?</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

ITEM (11)

AIM – indefinite quantity (e.g. <i>many</i> , <i>a few</i>)	Neutral CONTEXT:	It is very late in the evening, you are still in front of the pc doing your email, you think
	ENGLISH:	a. <i>I'll write a few other emails, and I'll be done</i> b. <i>I cannot go to sleep, until I'm done with the many other messages I left behind</i>
	tell us how you would say the same thing in YOUR LANGUAGE:	a. b.

3.2.3 Dual sets

Special attention is deserved by contexts with dual sets, based on the diachronic evidence provided by Gianollo and Mauri (2020: 136-149) on the development of dual Latin *alter* to (non-dual) Italian *altro*. We adopt a broad definition of duality and consider as dual every set

that is characterized by a twofold internal configuration. This may mean that the set is composed by two elements, in line with a strict definition of duality, or it may mean that the set is composed by a larger quantity of elements, which are however organized into two groups for specific discourse aims (obtaining an effect similar to complement anaphora). In these cases, it was not possible to conceive biased contexts, being duality already a highly specific condition that neutralizes incrementality, as argued in Section 2.2.

Item (12) provides a context that is strictly dual and is aimed at eliciting the ‘other’ expression that is likely to be employed to indicate the second of two elements. As will be discussed in Section 4, it is in these contexts that we may observe a convergence between the meaning of ‘second’ and the meaning of ‘other’, which constitutes a cross-linguistically frequent polysemy.

ITEM (12)

AIM – Dual set (two individual entities)	Neutral CONTEXT:	Frank has two brothers
	ENGLISH:	<i>One is a teacher, ... is a doctor</i>
	tell us how you would say the same thing in YOUR LANGUAGE:

Item (13) offers a context where a larger set is divided into two groups for the discourse aims. The speaker refers to a party and further identifies a twofold configuration: a group of persons who are still dancing, and another group, consisting of the remaining elements of the set, who have gone home. Duality is thus not an inherent property of the set, in the sense that the individuals in the set are more than two, but rather a property regarding its internal bipartite organization, which is in turn functional for the point that the speaker wants to make.

ITEM (13)

AIM – Dual set (two groups)	Neutral CONTEXT:	You organized a very nice party, now it is getting late
	ENGLISH:	<i>My brother and his girlfriend are still dancing, the others have already gone home</i>
	tell us how you would say the same thing in YOUR LANGUAGE:

Finally, Item (14), shows an example of reciprocal context, expressing a binary symmetric relation between entities, which involves the use of ‘other’ expressions in several languages (cf. Section 2.2 and 4).

ITEM (14)

AIM – Dual set (reciprocal)	Neutral CONTEXT:	It is important to keep good relations...
	ENGLISH:	<i>and be nice to one another</i>
	tell us how you would say the same thing in YOUR LANGUAGE:

3.2.4 Ordinal series

The last type of context we aim to explore through the questionnaire is the context of a series, within which more than two elements are organized according to an ordinal sequence. In the English utterance of Item (15) two of three slots are left empty, to leave the options between ‘one/the first’ and ‘third/the last’ open, while the second slot is overtly indicated. This decision is aimed to elicit the expression employed for ‘second’, which is known to have synchronic and diachronic connections to ‘other’ expressions in many languages (cf. Section 2.2 and 4). At the same time, it leaves open the possibility to use the ‘other’ expression to refer to the last element of a series.

ITEM (15)

AIM – Ordinal series	Neutral CONTEXT:	Frank has three pets
	ENGLISH:	<i>... is a dog, the second is a cat, ... is a parrot</i>
	tell us how you would say the same thing in YOUR LANGUAGE:

4. Cross-linguistic variation: main insights and challenges

The questionnaire was administered to at least one native speaker for 12 different languages. This first-stage sample includes Romance languages (Italian, French, and Portuguese), Germanic languages (English, German, Dutch), a Slavic language (Polish), an Iranian language (Persian), Semitic languages (Arabic, Hebrew), a Turkic language (Turkish) and a Finno-Ugric language (Finnish). A larger survey based on a more representative language sample is planned for a successive research stage, following the empirical validation of the questionnaire.

The goals with which the questionnaire was designed have been achieved, in that we were able to confirm the cross-linguistic significance of the non-identity / additive ambiguity, to identify different strategies employed for the various contexts, and to detect polysemies pointing to recurrent diachronic paths. In what follows, we will provide some relevant examples of the attested variation, which of course have no ambition of exhaustivity or typological representativeness, but allow us to identify some of the cutoff points where different languages may lexicalize the semantic functions connected to ‘other’ differently. In this section, we will focus on the dimensions that emerged more clearly from our limited survey, and that are more directly relevant for the discussion introduced in the previous sections. Then, in Section 5, we will conclude by singling out the most promising research perspectives that these preliminary results open up.

Language	[+id +incr]	[+id -incr]	[-id +incr]	[-id -incr]	complement anaphora	dual individuals	‘second’
ENGLISH	other	other	other	other	other	other	second
ITALIAN	altro	altro	altro	altro	altro	altro	secondo
FRENCH	autre	autre	autre	autre	autre	autre	deuxième
PORTUGUESE	outro	outro	outro	outro	outro	outro	segundo
PERSIAN	digar	digar	digar	digar	digar	digar	dovvomi n
ARABIC	ākhar	ākhar	ākhar	ākhar	ākhar	ākhar	thānī

	(Moroccan akhor)	(Moroccan akhor)	(Moroccan akhor)	(Moroccca n akhor)	(Moroccan akhor)	(Moroccan akhor)	(Moroccca n tani)
GERMAN	noch ein- 'still a' weiter- 'further'	ander-	ander-	ander-	ander-	ander-	zweit-
DUTCH	nog een 'yet a'	ander-	ander-	ander-	ander-	ander-	tweed-
TURKISH	daha 'more'	başka	başka	başka	diğer	diğer	ikincisi
POLISH	kolejny 'next' jeszcze jedny 'still one'	inny	inny	inny	inny	drugi	drugie
HEBREW	od	axer-	axer-	axer-	axer-	sheni	sheni
FINNISH	vielä toinen 'yet another'	toinen	muu	muu	muu	toinen	toinen

Table 3. Lexicalizations for 'other' and related concepts in the language sample

Table 3 summarizes the lexicalizations found in the languages of the sample for (i) the indefinite contexts eliciting the various combinations of the $[\pm id]$ and the $[\pm incr]$ dimensions (Items 4, 5, 6, 7); (ii) the dual contexts represented by dual individuals and complement anaphora (Items 12, 13); (iii) the word for 'second' (Item 15). As mentioned, this represents only a selection of the dimensions of variations observed through the questionnaire, which allows us to focus on the first generalizations that can tentatively be drawn from our small language sample.

In Table 3, translations follow the glosses given in context by informants. For simplicity, the table does not show cases in which additive particles can optionally accompany the word for 'other' in the additive reading. More in general, it is simplified with respect to possible alternative strategies in the language and only records the strategy indicated by the informants as most natural in the context.⁶ Co-occurrence with determiners is not indicated (e.g. *other* is given in the table also when in fact *another* is used).

Table 3 displays quite clearly the most visible result from our survey, namely the fact that the biased contexts provided for indefinite referents and covered by the lexeme *other* in English lead to the choice of different expressions in certain languages, thus revealing a lower degree of ambiguity of the 'other' lexeme there. In such cases, the word employed for 'other' indeed appears to exclude the meaning of prototypical additivity ($[+id +incr]$: pure incremental addition of a referent of the same kind, cf. Item (4)), allowing instead for incremental and non-incremental contexts involving a referent of a different type ($[-id +incr]$ and $[-id -incr]$, respectively), but also for substitution of a referent of the same type ($[+id -incr]$). That is,

⁶ For instance, the table does not indicate that languages that have a lexicalization for 'other' in the incremental contexts can also use additive elements as an alternative: e.g. in English besides *another* we can find *one more*. See further in the section for comments on this fact.

besides the prototypical non-identity context [-id -incr], the word for ‘other’ can express (i) incrementality, but only in the case of type difference, or (ii) type identity, but only in the case of non-incrementality (substitution of token).

The reason why a form of identity ([+id -incr], i.e. in substitution) can be expressed by the same lexeme expressing non-identity (in incremental and non-incremental contexts) is likely due to the contrast that is inherent in substitution (the clearest realization of the [-incr] value): each time an item is replaced by another one, even if they belong to the same type, there must be some context-dependent reason that allows to conceive of the two items as having opposite values for a sub-property (e.g. in Item (5), broken and not broken), thus leading to compare them in the light of this opposition (cf. Gianollo & Mauri 2020: Section 4.2).

Turkish provides an interesting case in point: the lexeme *başka* is employed with the meaning of ‘other’ when a substitution (as in (12a)) or some diversity (as in 12b)) is implied, whereas if the entity being referred to belongs to the same type and the speaker simply aims to convey pure additivity, as in (12c), a different strategy is used, involving the use of the quantifier *daha* ‘more’.

12) Turkish (Turkic)

a.	[ITEM 5: You are in a supermarket and you realize that the bag that the shop assistant gave you is broken, you ask:]					
	<i>Lütfen</i>	<i>başka</i>	<i>bir</i>	<i>çanta</i>	<i>alabilir</i>	<i>miyim?</i>
	Please	different	a	bag	take	may.I?
	‘Please, can I take another bag?’					

b.	[ITEM 7: When a friend asks you what you like about travelling, you say:]			
	<i>Başka</i>	<i>kültürleri</i>	<i>tanımayı</i>	<i>severim</i>
	Other	cultures:ACC	knowing:ACC	like:1SG
	‘I like to open up to other cultures’			

c.	[ITEM 4: You are in a supermarket and you realize that the bag that the shop assistant gave you is not enough, you ask:]					
	<i>Lütfen</i>	<i>bir</i>	<i>çanta</i>	<i>daha</i>	<i>alabilir</i>	<i>miyim?</i>
	Please	a	bag	more	take	may:1SG?
	‘Please, can I take another bag?’					

A similar situation is attested in Hebrew, where the lexeme *axer* is used to refer to ‘other’ in those cases where the set construction implies the substitution of an element of identical type [+id -incr] (13a) or when some qualitative comparison leads to a diversity ([-id]) meaning (13b), while the form *od* ‘again, more’ is found to convey incrementality of a token of the same type [+id +incr] (13c).

13) Hebrew (Semitic)

a.	[ITEM 5: You are in a supermarket and you realize that the bag that the shop assistant gave you is broken, you ask:]		
	<i>efshar</i>	<i>sakit</i>	<i>axeret?</i>
	Possible	bag	other/different?
	‘Please, can I take another bag?’		

b.	[ITEM 7: When a friend asks you what you like about travelling, you say:]		
----	---	--	--

	<i>ani</i>	<i>ohevet</i>	<i>lehiptax</i>	<i>letarbuyot</i>	<i>axerot</i>
	I	like:1SG	open.up.to	cultures	other/different
	'I like to open up to other cultures'				

c.	[ITEM 4: You are in a supermarket and you realize that the bag that the shop assistant gave you is not enough, you ask:]				
	<i>efshar</i>	<i>od</i>	<i>sakit?</i>		
	Possible	again/more	bag?		
	'Please, can I take another bag?'				

These two languages show another interesting distinction: in both Turkish and Hebrew a separate expression was chosen for dual sets (as in (14a) and (15a)). In Turkish we find the lexeme *diğer*, borrowed from Persian *digar* and restricted to reference in sets composed of two individuals, while in Hebrew we find *sheni*, which also has the meaning of 'second'. The same form is employed in Turkish also for larger sets organized in a twofold way, that is, in contexts of complement anaphora (as in (14b)), where Hebrew instead recurs to the 'diversity' lexeme *axer*, as in (15b):

14) Turkish

a.	[ITEM 12: Frank has two brothers]				
	<i>Biri</i>	<i>öğretmen,</i>	<i>diğeri</i>	<i>doktor.</i>	
	One:ACC	teacher,	other:ACC	doctor	
	'One is a teacher, the other is a doctor.'				

b.	[ITEM 13: You organized a very nice party, now it is getting late. You say: "My brother and his girlfriend are still dancing..."]				
	<i>diğerleri</i>	<i>çoktan</i>	<i>eve</i>	<i>gittiler</i>	
	others	already	home:DAT	go:PST:PL.	
	'the others have already gone home.'				

15) Hebrew

a.	[ITEM 12: Frank has two brothers]				
	<i>exad</i>	<i>more,</i>	<i>ha.sheni</i>	<i>rofe</i>	
	One	teacher	DEF.second	doctor	
	'One is a teacher, the other is a doctor.'				

b.	[ITEM 13: You organized a very nice party, now it is getting late. You say: "My brother and his girlfriend are still dancing..."]				
	<i>ha.axerim</i>	<i>kvar</i>	<i>halxu</i>	<i>ha.bayta</i>	
	DEF.others	already	went	DEF.home	
	'the others have already gone home.'				

The three languages in our sample that lexically distinguish, within dual sets, between sets formed by two individuals and cases of complement anaphora, that is Hebrew, Polish, and Finnish, are also the three languages in which the lexeme used for 'other' in sets of two individuals also means 'second'. Possibly, the ordinal meaning prevails here, and inhibits the extension to contexts of complement anaphora.

Interestingly, also Persian *digar* has historical connections with the meaning 'second': it derives from the reconstructed Proto-Iranian form **dwiṭīyakah* 'second', which contains the

Proto-Indo-European root **dwóh₁* ‘two’. In Modern Persian we find a different word for ‘second’, namely *dovvo/dovvomin*, while *digar* became the general expression meaning ‘other’, which can be employed in a wide spectrum of contexts, like Italian *altro*, including pure additivity ([+id + incr]), as we can see in (16):

16) Persian (Indo-European)

a.	[ITEM 5: You are in a supermarket and you realize that the bag that the shop assistant gave you is broken, you ask:]						
	<i>Bebakhshid,</i>	<i>yek</i>	<i>kiseye</i>	<i>digar</i>	<i>be</i>	<i>man</i>	<i>midahid?</i>
	Excuse [me],	one	bag	other	to	me	give:SG?
	‘Please, can I have another bag?’						

b.	[ITEM 4: You are in a supermarket and you realize that the bag that the shop assistant gave you is not enough, you ask:]								
	<i>Bebakhshid,</i>	<i>bāz</i>	<i>ham</i>	<i>yek</i>	<i>kiseye</i>	<i>digar</i>	<i>be</i>	<i>man</i>	<i>midahid?</i>
	Excuse [me],	again	also	one	bag	other	to	me	give:SG?
	‘Please, can I have another bag?’								

The use of the form for ‘second’ to mean ‘other’ in dual contexts is not surprising, given that in cases of bipartite sets reference to ‘the other(s)’ equates to referring to ‘the second one/group’. A look at the Database of Cross-linguistic Co-lexifications (CLICS, Rzymski & Tresoldi 2020) reveals that the use of the same lexeme for ‘other’ and ‘second’ is extremely frequent across languages, as shown in Figure 1. This further confirms what we observed for Latin (cf. Section 2.2) and suggests that a dual set construal tends to neutralize the discourse prominence of non-identity vs. incrementality: reference made by ‘other’ in these cases leads to the anaphoric identification of the remaining (or complementing) part of a presupposed dual set. Complementarity indeed appears to be the core meaning being conveyed by ‘other’ in dual (or dually construed) sets.

The use of explicitly iterative ('again') and additive ('more') strategies in prototypical additive contexts is suggestive of how the two dimensions of qualitative comparison and set construction (see Table 1) are managed by speakers, who seem to foreground either ADDITIVITY, leaving the dimension of qualitative comparison in the background, or NON-IDENTITY, leaving the dimension of set construction in the background.

If the nominal phrase is definite, however, in languages like German we find the lexeme used for non-identity contexts (Ger. *ander-*) both in quantified (18.a) and dual contexts (18.b). This suggests that non-identity-'other' and complementarity-'other' are very close, probably due to the fact that the remaining, or complementary, part of a presupposed dual set can be easily conceived as being inherently different from the anaphoric antecedent, i.e. the previously mentioned referent, and to the fact that incrementality is neutralized with presuppositional alternatives. What we observe in a language like German is thus the use of the same form for both (i) indefinite referents qualified as [-id], and (ii) definite referents, independently of their being quantified or not.

18) German

a.	[ITEM 9: It is very late in the evening, you are still in front of the pc doing your email, you think:]										
	<i>Ich</i>	<i>schreibe</i>	<i>die</i>	<i>beiden</i>	<i>anderen</i>	<i>Mails,</i>	<i>und</i>	<i>dann</i>	<i>bin</i>	<i>ich</i>	<i>fertig.</i>
	I	write	DEF	two	other	emails,	and	then	am	I	finished.
	'I'll write the other two e-mails and I'll be done.'										

b.	[ITEM 12: Frank has two brothers]						
	<i>Der</i>	<i>eine</i>	<i>ist</i>	<i>Lehrer,</i>	<i>der</i>	<i>andere</i>	<i>Arzt.</i>
	The	one	is	teacher,	the	other	doctor.
	'One is a teacher, the other is a doctor.'						

Finnish has a unique status in our sample: in the questionnaire, *toinen* was used in [+id] contexts, whereas *muu* was used in [-id] contexts, independently of incrementality. The data for Finnish emerging from the questionnaire, thus, seem to suggest that the encoding of [\pm id] prevails in this language with respect to [\pm incr]; that is, the dimension of type (NON)-IDENTITY is more important for the lexical contrast than the dimension of additivity. In order to unambiguously express incrementality, *toinen* is accompanied by the additive particle *vielä* in [+id, +incr] contexts. Moreover, *toinen* remains strongly tied to its ordinal function and is constrained to sets composed of two individuals (19.a). In the case of complement anaphora, *muu* was used (19.b). However, the contrast between the two words for 'other', *toinen* and *muu* is made more complex by the interaction with number and by interpretive factors that are not captured by the questionnaire, on which see Kaiser, this volume.

19) Finnish (Finno-Ugric)

a.	[ITEM 12: Frank has two brothers]						
	<i>Toinen</i>	<i>on</i>	<i>opettaja</i>	<i>ja</i>	<i>toinen</i>	<i>on</i>	<i>lääkäri.</i>
	One-of-the-two	be-3SG	teacher	and	one-of-the-two	be-3SG	doctor.
b.	[ITEM 13: You organized a very nice party, now it is getting late. You say: "My brother and his girlfriend are still dancing..."]						
	<i>muut</i>	<i>ovat</i>	<i>jo</i>	<i>menneet</i>	<i>kotiin.</i>		
	other-PL	be-3SG	already	go-PPF.PL	home-ILL		

All in all, our data suggest that the ‘other’ domain is synchronically and diachronically connected to three poles of attraction:

- I. COMPLEMENTARITY, whereby reference being made through ‘other’ is relevant to the anaphoric identification of the remaining (or complementing) part of a presupposed dual set;
- II. NON-IDENTITY, whereby reference being made through ‘other’ is relevant to the identification of the set composition (which members?), in the light of some qualitative comparison;
- III. ADDITIVITY, whereby reference being made through ‘other’ is relevant to the determination of the set extension (how many members?), in the light of some underlying iterative addition.

We find languages distinguishing between the three poles, such as Turkish (cf. (12) and (14)); languages using the same strategy for complementarity and non-identity-‘other’, such as German (cf. (17) and (18)); and languages allowing the same strategy for all three, such as Italian (cf. Gianollo & Mauri 2020). No language is attested in this first-stage sample that uses the same form for complementarity and additivity, without using it also for non-identity ‘other’. Finnish does not represent an exception to this claim because, if it is true that the form for sets of two individuals is the same as the form for prototypical additivity and differs from the form used for non-identity, nonetheless the form used for complement anaphora is the same form used for non-identity. For reasons of space, we cannot discuss this finding further, but it is the object of Gianollo & Mauri (forthcoming).

These first-stage results thus point on the one hand to a cross-linguistic variation in the ‘other’ domain that is worth analyzing in detail, and on the other hand to regular patterns recurring in unrelated languages, motivated both diachronically and semantically. In the next section we will provide some concluding remarks and discuss the future steps of this project, crucially involving the use of the questionnaire described in Section 3.2.

5. Conclusions

The aim of the paper was twofold. On the one hand, we spelled out some theoretical and methodological considerations, based on which a systematic cross-linguistic investigation of the ‘other’ semantic domain becomes possible. On the other hand, we aimed to discuss in detail the rationale, the structure and the contents of the questionnaire that we designed to collect cross-linguistic data, together with the preliminary results achieved through its administration to informants from 12 different languages.

After describing the distinction between the non-identity and the additive reading of ‘other’, based on the corpus study on Italian conducted by Gianollo & Mauri (2020), we identified the factors that play a role in disambiguating between the two readings, with a focus on definiteness, quantification and dual set construction (Section 2). Our hypothesis was that a systematic analysis of the interaction between these factors and the use of ‘other’ lexemes could be revealing of some regular cross-linguistic variation, especially with respect to the additive vs. non-identity opposition, potentially pointing also to recurrent diachronic paths of change in the development of ‘other’ expressions.

To test this hypothesis, we built a questionnaire specifically designed to elicit the production of ‘other’ expressions across languages, which was the object of Section 3. The main challenge consisted in creating a relevant discourse background that could activate the natural use of ‘other’ lexemes in contexts characterized by incrementality (or the absence

thereof) and type identity (or the absence thereof), both with singular and plural, definite and indefinite, quantified and non-quantified nominal phrases, in dual and non-dual sets.

After a pilot trial, we finally employed this questionnaire to collect data from a sample of 12 Indo-European and non-Indo-European languages. In Section 4 we presented the first results of this cross-linguistic survey, which turned out to be revealing of both the language-internal and cross-linguistic relevance of the points of variation that we identified, confirming that the way in which languages organize the ‘other’ domain is sensitive to the dimensions of set incrementality, qualitative comparison, and dual set configuration.

In particular, our data confirm the cut-off point between the additive and the non-identity reading of ‘other’, which may correspond to separate expressions (in languages with a restricted ‘other’ lexeme, e.g. Hebrew and German) or to ambiguous ones (in languages with a more versatile form, e.g. Persian and Italian). This survey crucially pointed out the potential relevance of a third reading, namely complementarity, whereby ‘other’ anaphorically refers to the remaining, or complementary, part of a presupposed dual set. Complementary ‘other’ is indeed not compatible with an additive reading, and projects non-identity at the token level, not at the type level. Interestingly, we observed that it is possible for a language to use three separate expressions for complementarity ‘other’, additivity ‘other’ and non-identity ‘other’ (cf. Turkish), to use the same form for complementarity and non-identity ‘other’ and a different form for additivity, or to use one and the same form for all three cases.

These results are extremely promising in view of the next stage of research, which involves the analysis of a larger language sample that will be explored both synchronically, by means of the questionnaire, and diachronically, for those languages for which enough historical evidence is available (see Gianollo & Mauri, forthcoming). Understanding how the semantics of ‘other’ is mapped into the world’s languages and how ‘other’ lexemes evolve has indeed a potential impact on our understanding of the very concept of ‘otherness’, paving the way for further potential developments in experimental and sociological studies.

References

- Bar-Asher Siegal, Elitzur. 2015. What Is New in the NP-Strategy for Expressing Reciprocity in Modern Hebrew and What Are Its Origins? In Edit Doron (ed.), *Language Contact and the Development of Modern Hebrew*. Leiden: Brill, 245–257.
- Beck, Sigrid. 2000. The semantics of *different*: comparison operator and relational adjective. *Linguistics and Philosophy* 23, 101–139.
- Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In Anna Szabolcsi (ed.), *Ways of Scope Taking*. Dordrecht: Kluwer, 71–108.
- Bortolussi, Bernard. 2015. *Syntaxe des indéfinis latins. Quis, quisque, alius*. Paris: Presses de l’Université Paris-Sorbonne.
- Breban, Tine, 2003. The grammaticalization of adjectives of identity and difference in English and Dutch. *Languages in Contrast* 4.1, 165–199.
- Breban, Tine, Kristin Davidse, Lobke Ghesquière. 2011. A typology of anaphoric and cataphoric relations expressed by English complex determiners. *Journal of Pragmatics* 43, 2689–2703.
- Brugè, Laura. 2017. Otras reflexiones sobre *otro*. *Rassegna iberistica* 40.108, 225–252.
- Brugè, Laura. 2018. *Otro*: Consideraciones a favor de su no inclusión en la categoría determinante. *Borealis: An International Journal of Hispanic Linguistics*, 7.2, 109–134.
- Brugè, Laura & Giuliana Giusti. 2021. Some notes on *outro* in Portuguese. *Revista de Estudos Linguísticos da Universidade do Porto* N.o Especial 2021, 441–471.
- Carlson, Greg N. 1987. Same and Different: Some Consequences for Syntax and Semantics. *Linguistics and Philosophy* 10, 531–566.

- Charnavel, Isabelle. 2015. *Same, different and other* as comparative adjectives – A uniform analysis based on French. *Lingua* 156, 129–174.
- Cinque, Guglielmo. 2015. A note on ‘other’. In Cinque, Guglielmo (ed.), *Charting the Landscape of Linguistics. Webschrift for Josef Bayer*. Konstanz: Universität Konstanz, 22–27.
- Eguren, Luis & Cristina Sánchez. 2003. La gramática de *otro*. *Revista de la Sociedad Española de Lingüística*, 33.1, 69–123.
- Eguren, Luis & Cristina Sánchez. 2004. Contrast and addition in Romance: A case study in microvariation. In Auger, Julie, Clancy J. Clements, Barbara Vance (eds.), *Contemporary Approaches to Romance Linguistics*. Amsterdam: Benjamins, 159–176.
- Gianollo, Chiara & Caterina Mauri. 2020. *Uno, l’altro e un altro* ancora: ambiguità dell’alterità tra sincronia e diacronia. *Quaderni di Semantica Special issue*, 131–174.
- Gianollo, Chiara & Caterina Mauri. Forthcoming. *A diachronic typology of ‘other’*. Ms., University of Bologna.
- Gutiérrez Rodríguez, Edita. 2011. Adjetivos y determinantes: *otro y demás*. In Victoria Escandell Vidal, Manuel Leonetti & Cristina Sánchez López (eds.), *60 problemas de gramática: dedicados a Ignacio Bosque*. Madrid: Akal, 62–68.
- Kulikov, Leonid. 2014. Grammaticalization of reciprocal pronouns in Indo-Aryan: Evidence from Sanskrit and Indo-European for a diachronic typology of reciprocal constructions. *Journal of South Asian Languages and Linguistics* 1.2, 117–156.
- Mantenuto, Iara. 2020. *‘Other’ and ‘More’ in San Sebastián del Monte Mixtec and Beyond*. PhD dissertation, UCLA.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70, 369–415.
- Mauri, Caterina & Andrea Sansò. 2018. Linguistic strategies for ad hoc categorization: theoretical assessment and cross-linguistic variation. *Folia Linguistica Historica* 39: 1-35.
- Mauri, Caterina, Silvia Ballarè, Eugenio Gorla, Massimo Cerruti, Francesco Suriano. 2019. KIParla corpus: a new resource for spoken Italian’. In Raffaella Bernardi, Roberto Navigli, Giovanni Semeraro (eds.), *Proceedings of the 6th Italian Conference on Computational Linguistics CLiC-it*, <http://ceur-ws.org/Vol-2481>.
- Milner, Jean-Claude. 1984. *Syntaxe et sémantique du constituant réciproque “l’un...l’autre”*, in Jean-Claude Milner, *Recherches sur l’anaphore*, Paris: Université Paris 7, 35–68.
- Moltmann, Friederike. 1992. Reciprocals and Same/Different: Towards a Semantic Analysis. *Linguistics and Philosophy* 15, 411–462.
- Moxey, Linda & Anthony Sanford. 1993. *Communicating Quantities. A Psychological Perspective*, Mahwah, NJ, Laurence Erlbaum Associates.
- Muliačić, Žarko. 1970. Altérité et alieté, *Filologija* 6, 105–111.
- Nouwen, Rick. 2003. Complement anaphora and interpretation. *Journal of Semantics* 20, 73–113.
- Oxford, Will 2010. *Same, other, and different*: A first look at the microsyntax of identity adjectives. *Proceedings of the 2010 annual conference of the Canadian Linguistic Association*, 1–15.
- Rzyski, Christoph, Tiago Tresoldi et al. 2020. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data* 7: 13. [DOI: 10.1038/s41597-019-0341-x](https://doi.org/10.1038/s41597-019-0341-x)
- Striegnitz, Kristina. 2005. *Generating Anaphoric Expressions. Contextual Reasoning in Sentence Planning*. PhD dissertation, Universität des Saarlandes.
- Van Peteghem, Marleen. 1995. L’ ‘indéfini’ *autre*; analyse sémantico-référentielle ou en quoi *autre* est déictique. In Rika Van Deyck (éd.), *Diachronie et variation linguistique. La deixis temporelle, spatiale et temporelle*. Ghent: Communication and cognition, 87–114.

- Van Peteghem, Marleen. 1997. *Autre* et ses correspondants en roumain. *Revue Romane* 32.1, 27–50.
- Van Peteghem, Marleen. 2000. Les indéfinis corrélatifs *autre*, *même* et *tel*. In Léonie Boseveldt-De Smet, Marleen Van Peteghem, Danièle Van de Velde (eds.), *De l'indétermination à la qualification. Les indéfinis*. Arras: Artois Presses Université, 117–201.
- Tekavčić, Pavao. 1972. *Grammatica storica dell'italiano*, Volume II: *Morfosintassi*, Bologna: Il Mulino.
- Tovena, Lucia & Marleen Van Peteghem. 2003. Facets of 'different' in French: *différent* and *autre*. In Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr, Francis Corblin (eds.), *Empirical issues in formal syntax and semantics 4*. Paris: Presses Universitaires de la Sorbonne, 63–79.