Transition of Language Structure in Multicultural Society: Grammatical Gender, Quantification and Negation in Languages across Eurasia[†]

Chingduang Yurayong

Department of Languages, Faculty of Arts, University of Helsinki, 00014 Helsinki, Finland

(Corresponding author's e-mail: chingduang.yurayong@helsinki.fi)

Abstract

This study explores the transition of language structure as a manifestation of multilingual and multicultural society. It examines the dynamic nature of language by considering both internal and external factors which contribute to language change. The approach is rooted in language typology which aims to explore recurring patterns across languages and utilise them as criteria for classifying languages into various structural types. The objective is to use the internal composition and alignment of patterning models between languages in geographical adjacency as evidence to explain contact-induced changes resulting from interactions between multilingual speaker communities across Eurasia. The investigation includes a critical discussion of the extent to which language change is influenced by contact on the one hand. On the other hand, the question concerns how language change is motivated by communicative needs related to increasing structural complexity to accommodate a wider range of sociocultural contexts where language is used. Three case studies on grammatical gender, quantification and negation demonstrate how language contact can drive the convergence of structural patterns among languages, in addition to inherent language-internal tendencies to reorganise structure and enhance performance. At a methodological level, the study offers a framework for understanding transitions within language structure and their reflection of broader cultural adaptation and shifts in human civilisation.

Keywords: Language change, Language typology, Language contact, Multilingualism, Eurasia

Abbreviations

1	1st person		IMP	imperative
2	2nd person		INSTR	instrumental
3	3rd person		IPFV	imperfective
ACC	accusative		HAB	habitual
AOR	aorist		NEG	negative
CLF	classifier		PERF	perfect
CNG	connegative		PL	plural
COND	conditional		PROG	progressive
COP	copula		PTCL	particle
DAT	dative		PTCP	participle
DUR	durative		Q	question
EXIST	existential		SG	singular
EXT	extension		SUBJ	subject
FUT	future		TOP	topic
GEN	genitive		VIM	valence increasing marker

[†]Presented at the International Conference on Communicating in Multicultural Society and Transitional World (March 28, 2024 at Walailak University, Thailand)

1. Introduction

Language is a dynamic component of human civilisation which undergoes continuous evolution. While some elements and constructions are sustained over time, others change at varying paces across different structural domains. In this view, language can be considered as a living symbiosis which often comes across various necessities for internal reorganisation of its structure (Paul, 1886). A major motivation for recurring adjustment is the need to achieve optimal performance and balance between explicitness and efficiency of communication while considering the capacity and constrained complexity inherent in human language (Grice, 1989; Hawkins, 2004). This principle aligns with Aristotle's rhetoric and Grice's (1989) maxims of conversations rooted in his concept of the cooperative principle, both which emphasise the optimal relation between explicit and economic language use.

'If it is prolix, it will not be clear, nor if it is too brief. It is plain that the middle way is appropriate ..., saying just enough to make the facts plain.'
(Aristotle, Rhetoric, 3.12 - 3.16, cited in Horn, 2006: 14)

The maxim of quantity:
1) Make your contribution as informative as is required (for the current purposes of the exchange).
2) Do not make your contribution more informative than is required.
The maxim of manner:
1) Avoid obscurity of expression.
2) Avoid ambiguity.
3) Be brief (avoid unnecessary prolixity).
4) Be orderly.
(Grice, 1989)

From a vitality perspective, a speaker community which deliberately resists changes may risk language obsoletion as cultural practices evolve rapidly. The resistance to change may leave the language unable to keep pace with the speakers' demands for expressing emerging concepts and communicative needs in various registers. This corresponds to a view that the maintenances of structural complexity and contexts of use are correlative (Grieve, 2023). Considering both internal and external aspects of language, as introduced above, the present study views transition as a natural process of language evolution. Disrupting this mechanism through overly prescriptive measures, as exemplified by criticisms towards language authorities in various political states, likely increases the tendency of a language becoming obsolete.

The current study introduces a theory and method for researching transition of language structure through the lens of language typology. By applying this framework to examine three structural domains closely connected to human culture, the analysis explores how their transitions have unfolded in various languages and linguistic areas across Eurasia. The focus is given to language-external factors pertaining to the sociology and ecology of individual speakers and speech communities over millennia, which are examined to understand their contributions to these transitions. The theoretical discussion and interpretation of case studies shed light on how we can tackle structural aspects relevant for the studies of linguistic and cultural transition. Throughout the article, any language examples given without sources are constructed based on the author's personal knowledge.

The framework utilised in the present study essentially demonstrates how to move beyond the overreliance on lexical analysis and ethnographic description of speaker populations within the realm of studies intersecting language, culture and society.

2. A typological approach to transition of language structure

2.1 What is language typology?

Language typology is a subfield of linguistics which operates on two major tasks as defined by Croft (2003). The first explores what is possible in human language through comparison of languages across families and geographical areas. In other words, it asks a question what kind of patterns and constraints occur systematically in and across specific domains of language structure. This includes sound system, word formation, clause structure or context of use. From this cross-linguistic comparison, language typologists can generalise what is universal across languages, i.e. 'linguistic universals', a framework which was remarkably put forward by Greenberg (1963).

The notion of linguistic universals leads to the second task of classifying languages into types. It is assumed that the global complexity of the world's languages is considered equivalent in an absolute sense, i.e. every language is equally complex. What varies between languages is the internal composition and distribution of complex and simple features across various structural domains (Shosted, 2006; Miestamo, 2017). From a communicative perspective, there is also a correlative increase in structural complexity as speakers acquire a wider range of contexts where their language can be used, thereby increasing in parallel 'situational complexity' (Grieve, 2023). The relationship between structure and usage in this sense offers a useful tool for discussing language structure from a cultural perspective.

Complexity in one domain is often compensated by lower complexity in another (Fenk-Oczlon & Fenk, 2008). This can be observed in the inverse relationship between the number of phonological inventories (sounds) vs. morphological processes (word modifications). For instance, most Tai-Kadai languages are known for their large phonological inventories but have few morphological processes. Meanwhile, Austronesian languages prototypically have a simple phonological system but complex morphological structures. Constrained by the overall complexity limitation of human language, speakers find their own solutions to organise these components within their language system. This process results in various possible patterns which are characteristic for individual speaker groups. These patterns, in other words, are identified as occurring types to which a language can be classified.

Cross-linguistic comparison and typological classification inherently have a functional orientation. This stems from the understanding that languages adopt individually unique solutions for formally organising their internal structure, leading to significant variation in how they select existing resources to produce specific meanings (Haspelmath, 2010). Therefore, it is more reasonable to start exploring types of the world's languages through functions, asking what elements are used to express specific meanings. In contrast, a form-based approach in generative linguistics focuses on answering to a question what meanings specific linguistic elements can express. This might lead to situations where categorically equivalent forms in different languages are no longer comparable due to functional extensions within individual elements. Consider, for example, the semantic map of directional prepositions *to* in English, *à* in French and *zu* in German, whose functional ranges only partially overlap (see **Figure 1** based on Haspelmath, 1998). Note that English *to* and German *zu* share a common etymology, going back to Proto-Germanic *tō*.
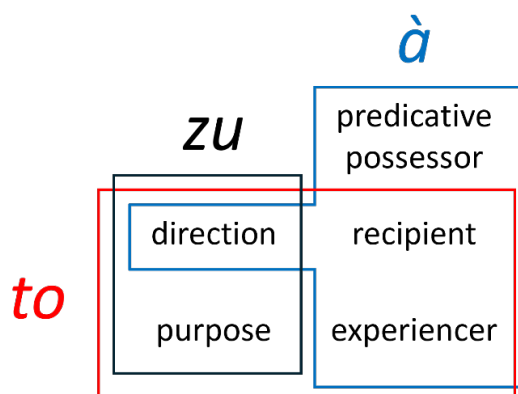

Figure 1 Semantic map and functional ranges of English *to*, French *à* and German *zu* prepositions.

Similar patterning models observed across individual languages, especially when they are spoken in adjacent areas, can be suggestive of interaction between their speakers. This evidence is frequently used in studies of structural changes due to language contact, which is discussed next.

2.2 How do we study transition of language structure typologically?

Language typology can serve as a good framework for diachronic studies of language for two major reasons. First, this connection emerges from the implicational characteristic of typological argumentation which predicts the presence or absence of one specific language feature through another, based on the inherent constraints of human language (Greenberg, 1957).

> '[The typology of languages] clearly adds to our understanding of linguistic historical change and our predictive power since from a given synchronic system certain developments will be highly likely, others have less probability, and still others may be practically excluded.' (Greenberg, 1957)

This methodological power can explain that, for instance, the presence of nasal vowels implies the presence of oral vowels in a language because a language cannot meaningfully distinguish nasality without a corresponding orality in vowel quality. The typological analysis can support the comparative method in historical linguistics as stated by Jakobson (1958), but it could also override the capacity of the comparative method which is restricted to 8,000 years before present (Nichols, 1992; Song, 2014).

> '[Thus o]ur 'predictive power' in reconstruction gains support from typological studies.' (Jakobson, 1958)

The second advantage of implicational analysis in language typology derives from the fact that a language always belongs to one possible structural type. Therefore, we are not dealing with infinite but restricted numbers of possibilities. This aligns with the concept of 'dynamicisation of typology' (Greenberg, 1978), which essentially means that language change generally follows predictable patterns. For instance, word order patterns in a language typically operate on a binary basis because words can either precede or follow another, and there is no third possibility to this. This inherent constraint underpins the principle that 'languages can only change from one occurring type into another occurring type' (Song, 2014). This principle is particularly useful for studies of interactions between languages. By considering the repertoire of speakers involved in contact and structural patterns available in their languages, it is possible to predict the directions of language change.

Given that languages belong to specific structural types and speakers develop solutions within these constraints, it would not be a surprise if similar types might coincidentally occur across languages spoken in different parts of the world. However, a general assumption in language contact studies is that 'there is no evidence that any languages have developed in total isolation from other languages' (Thomason, 2001). This is particularly relevant in multilingual societies where speakers engage with multiple languages within the same cultural setting. Such a continuous exposure to multilingualism can lead speakers to adapt their language structures to become formally and functionally comparable to those of the other languages in their repertoire. This phenomenon often relates to a broader cultural assimilation process (Sinnemäki, 2020). The outcome of such adaptation often leads to converging structural patterns among languages in contact. Performance-wise, the linguistic adaptation also serves as a strategy to reduce the cognitive effort required when switching between different structural patterns.

Multilingual speakers may cognitively apply a process of 'polysemy copying' (Weinreich, 1974; Heine & Kuteva, 2005; Wiemer & Wälchli, 2012), which extends or readjusts the functional range of a linguistic element in one language to match the function of an equivalent form in another language. This mechanism aligns with the concept of correlative increase of structural complexity as speakers encounter through extension of situational complexity (Grieve 2023, as discussed in Section 2.1 What is language typology?). In an intense contact situation, multiple layers of polysemy copying may lead to languages

sharing common structural templates, while retaining distinct vocabularies. Such a contact scenario reflects Kopitar's (1829) observation of multilingual speakers using one grammar for multiple languages in the Balkans.

> *nur eine Sprachform herrscht, aber mit dreyerley Sprachmaterie*
> 'only one structure is produced, but with three language materials [Albanian, Bulgarian, Romanian]' (Kopitar, 1829)

Such a high degree of convergence is termed 'intertranslatability' (Romaine, 1988). The phenomenon has been observed across adjacent languages in various regions in Eurasia, such as India (Masica, 1976), Mainland Southeast Asia (Enfield, 2005), Amdo-Qinghai areas (Janhunen, 2007) and Japan-Korea (Yurayong & Szeto, 2020). Consider, for instance, the following examples illustrating morpheme-by-morpheme correspondences between Japanese and Korean in (**1**), and Thai and Khmer in (**2**). Importantly, these language pairs do not belong to the same language families, as evidenced by their distinct vocabularies, unless it involves a shared loanword from the third language, such as the Sanskrit word प्रदेश *pradeśa* 'country, region' in (**2**), or the common Sinitic loanwords in Japanese and Korean.

(1)	a.	*kano*	*jyo*	*no*	*uta*	*goe*	*wa*	*koyoba*	*wo*	*ushina-u*	*kurai*	*da.*	**Japanese**
	b.	*kŭ*	*nyŏ*	*ŭi*	*norae.s*	*sori*	*nŭn*	*mal*	*ŭl*	*ilh-ŭl*	*chŏngdo*	*da.*	**Korean**
		that	woman	GEN	song	voice	TOP	word	ACC	lose-PTCP	degree	COP	

'The singing voice of hers is so beautiful that I lose my words.' (Uchiyama, 2011)

(2)	a.	*prateh*	*lo:k*	*miən*	*tɨkkak*	*tʰliə*	*te:ʔ*	**Khmer**
	b.	*pràtʰê:t*	*kʰun*	*mi:*	*hìʔmáʔ*	*tòk*	*măi?*	**Thai**
		country	you	exist	snow	fall	Q	

'Does it snow in your country?'

Examining structural similarities and differences between languages in adjacent areas methodologically serves as an effective tool for investigating their shared history, both in the pre-contact and post-contact stages. This approach allows for establishing a relative chronology of various restructuring stages in a language, as will be demonstrated in the following case studies of grammatical gender, quantification and negation.

3. Case studies

3.1 Grammatical gender

Grammatical gender serves as a strategy to differentiate and classify entities into semantically associated groups of words. It forms a system where a noun is assigned to a specific class, which is reflected in the forms taken by other elements syntactically related to it (Matthews, 1997). Consider different types of grammatical elements involved in the gender distinction of entities, as in (**3**) to (**5**). Functionally, grammatical gender is comparable to classifiers found in Asian languages and noun class markers in African languages, all of which are devices for referentially individuating entities (Corbett, 2007). Similarly to grammatical genders, classifiers are traditionally used to categorise nouns and nominals. The major difference is that classifiers are distinguished on the basis of natural properties rather than grammatical properties of entities referred to. The criteria can be shape, type or other semantic properties such as humanness, animacy, and natural gender of quantified objects (Aikhenvald, 2000; Hellinger & Bußmann, 2001).

(In)definite articles

(3) a. **un** *ragazzo* 'a boy'   c. **una** *ragazza* 'a girl'   **Italian**

    b. **il** *ragazzo* 'the boy'   d. **la** *ragazza* 'the girl'

Numerals

(4) a. **edin** *măž* 'one man'   c. **edna** *žena* 'one woman'   **Bulgarian**

    b. **dva** *măže* 'two men'   d. **dve** *ženi* 'two women'

Verb endings

(5) a. (*anta*) *ta-ktub-**u*** 'he writes'   b. (*anti*) *ta-ktub-**īna*** 'she writes'   **Arabic**

Cross-linguistically, it is noticeable that individuating entities by grammatical genders is not a rare phenomenon, as it is present in all continents (see the global distribution in **Figure 2**). This may highlight the linguistically universal role of such coreferential mechanism in enhancing the efficiency and success of interpersonal communication. The availability of such grammatical elements in the grammatical system namely facilitates the identification of referential relations which are established between verbs describing actions or events, and nouns representing participants.
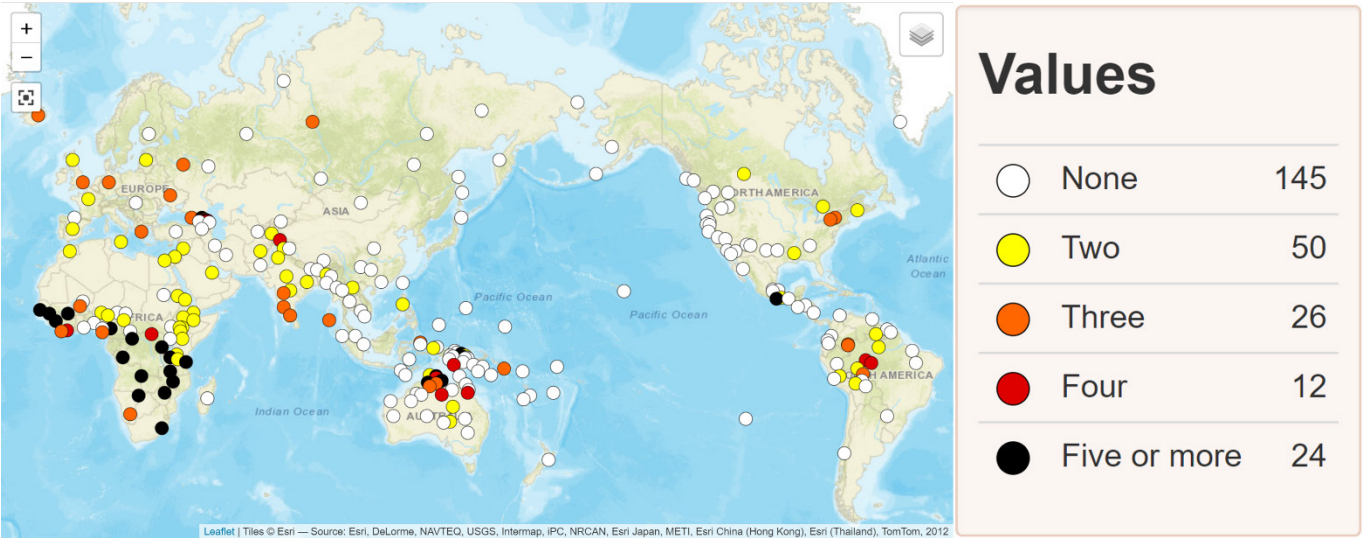

Figure 2 Geographical distribution of languages with and without grammatical genders (Corbett, 2013).

Diachronically, as human civilisation advances, language acquires and acoustically represents an increasing number of concepts. This leads to a transition which necessitates a more economical means of memorising words as easily recognisable groups to facilitate their use, thereby optimising performance through complexity (a principle introduced in Section 1 Introduction). In response to this internally emerging need, various speech communities have chosen specific grammatical resources for such classification purposes. Crucially, a decision is often influenced by their internal cognition and environment. One well-known example of grammatical gender is found in Indo-European languages, which currently exhibit diverse types of gender classification, ranging from three (masculine vs. neuter vs. feminine) to two (masculine vs. feminine or animate vs. inanimate) or zero genders (Luraghi, 2011). Consider the gender marking by definite articles in examples (**6**) to (**9**), which illustrate variation across Indo-European languages.

Three sex-based genders

(6)	a.	Masculine	*o*	***o** patéras*	'the father'	**Greek**
	b.	Neuter	*to*	***to** scholeío*	'the school'	
	c.	Feminine	*i*	***i** thálassa*	'the sea'	

Two sex-based genders

(7)	a.	Masculine	*o*	***o** senhor*	'the gentleman'	**Portuguese**
	b.	Feminine	*a*	***a** senhora*	'the lady'	

Two animacy-based genders

(8)	a.	Animate	*de*	***de** leider*	'the leader'	**Dutch**
	b.	Inanimate	*het*	***het** huis*	'the house'	

Zero gender distinction

(9)	a.	Masculine	*the*	***the** groom*		**English**
	b.	Neuter	*the*	***the** house*		
	c.	Feminine	*the*	***the** bride*		

It is also worth considering a more complex gender system beyond Indo-European languages, as that can reduce our bias towards the Indo-European models. For example, in Tsez, a Northeast Caucasian language spoken in the Southern Dagestan area of Russia, both sex and animacy intertwine and create a four-gender system: 1) male humans, 2) female humans and inanimates, 3) animals and inanimates, and 4) inanimates. Example (**10**) illustrates the differential marking for each gender on adjectives *exora* 'tall, high.'

(10)	a.	Male humans	zero	*exora žek'u*	'a tall man'	**Tsez**
	b.	Female humans and inanimates	*y-*	***y**-exora ɣˤanabi*	'a tall woman'	
	c.	Animals and inanimates	*b-*	***b**-exora got'i*	'a high haystack'	
	d.	Inanimates	*r-*	***r**-exora ɣun*	'a tall tree'	

(Polinsky 2015: 4)

Of the aforementioned types, the transition towards a zero-gender system in some modern Indo-European languages can be explained internally by phonological changes which render gender-specifying elements indistinguishable. For instance, grammatical genders can still be identified through the ending of words, such as the feminine ending *-a* in Italic languages (**3**) and (**7**), Slavic languages (**4**) and Greek (**6**), but that element is no longer visible in Germanic languages in (**8**) and (**9**) due to a phonological erosion in the word-final environment. Consider, for example, the distinguishable masculine *an*-stem and feminine *i*-stem endings in the Proto-Germanic forms **gum-**an**-* 'groom' (ultimately cognate with *human*) vs. **brūd-**i**-* 'bride.' On the one hand, languages may still maintain gender distinction on the specifying elements, as in the case of Dutch (**8**). On the other hand, languages may have undergone further simplification, as in the case of English (**9**), where grammatical gender is no longer distinguishable in nominal endings or articles. This follows because the English article paradigm has been simplified from three distinct Old English articles (*sē* for masculine, *þæt* for neuter, *sēo* for feminine, and *þā* for plural) to a single Modern English article *the*.

Apart from the reorganisation of internal structure discussed above, the borrowing of a grammatical or conceptual model from neighbouring languages, which lack grammatical gender, should not be underestimated as another reinforcing factor alongside internal change. This is probably the case for Turkic influence on Iranian languages, which have undergone a similar process as the English case described above. Namely, grammatical gender is not a meaningful grammatical category in Turkic languages. This structural

model could have influenced a nominal paradigm of Iranian languages, such as Tajik in (**11**) where modifiers *yak* 'one, a(n)' and *hub* 'good' have identical forms regardless of the genders of references.

(11)	a.	Masculine	*yak šavhar-i hub*	'a good husband'	**Tajik**
	b.	Neuter	*yak daraht-i hub*	'a good tree'	
	c.	Feminine	*yak zan-i hub*	'a good wife'	

One accelerating factor has been obviously multilingualism, which often leads to structural convergence (as discussed in Section 2.2 How do we study transition of language structure typologically?). This is particularly prominent along the Silk Road network in Central Asia, where the interaction and language shift have been active among Turkic and Iranian speaking populations over a millennium (see also numerous studies in a collective volume on Turkic-Iranian contacts edited by Johanson and Bulut 2006).

Nevertheless, the structural approach adopted in the current study maintains that grammatical gender assignment is not motivated by the speakers' perceptual association between individual referents and a specific biological sex, as is often romanticised by some groups of language enthusiasts (Eckert & McConnell-Ginet, 2013). There have also been studies which hypothesise the lack or presence of genders in official languages of political states as a measurement of gender equality and related parameters (Prewitt-Freilino, Caswell & Laakso, 2012; Jakiela & Ozier, 2020). This approach is, however, problematic, and it does not make sense to associate a lower degree of Eurocentrically evaluated gender equality with a community using language with large inventories of classifiers or noun class markers but no sex-based gender. In the typological sense, those languages should be considered to have as many genders as their existing classifiers or noun class markers. Instead, the process of gender assignment in grammar is better seen as being determined by language structure. In other words, it is more reasonable to seek an explanation from ambiguity caused by phonological erosion like the case of English, or from the purpose of enhancing the efficiency of communication where interlocutors can identify entities involved in specific actions or events more explicitly (see also the discussion on gender-specific classifiers in Section 3.2 Quantification).

3.2 Quantification

Quantification functionally resembles grammatical gender in the sense that both are used to classify and identify words and involve various grammatical categories. Considering major parts of speech, nominals and verbs are two classes most frequently subject to quantification. Counting of entities and actions can be either indefinite (e.g. ***some/few/most/many*** *students/times*) or exact (e.g. ***one/two/ten/hundred*** *students/times*). In this study, the focus is on an adjustable context of exact quantification, which involves numerals. The data mainly concern a sortal type of quantifier phrases which categorise counted words based on their semantic features such as ***four*** *men*, excluding a mensural type which comprises units to measure words with low countability such as ***five bottles of*** *water* (see also the definition and complete typology of quantifying noun classification in Gil 2013). The sortal classifying devices are well known for languages spoken Eastern Eurasia under the term 'classifiers' (Bisang, 1999; Lu, 2012; Gerner, 2014).

From a cultural perspective, the presence of grammatical devices of quantification in a language may be taken for granted by people living in modern society, driven by numerical data. However, it is worth noting that not all human languages possess grammatical devices to count entities or events in a large scale. For instance, Aka-Jeru and Jarawa languages of the Andaman Islands originally lack lexical resources for counting beyond 'two' and 'three', respectively. Beyond these quantities, only an indefinite quantifier 'many' is used, as illustrated in example (**12**).

(12)	a.	*ṭɔplɔ*	*nɛrtapʰul*	-	*narakamu*	**Aka-Jeru**
	b.	*wɑja*	*naja*	*kaŋitəǰile*	*mala*	**Jarawa**
		'one'	'two'	'three'	'many'	
		(Abbi, 2013; Kumar, 2012)				

This phenomenon could be attributed to the living environment of the Andaman Islands. Historically prior to contact with populations from India, the local society may not have required exact and adjustable counting methods to the extent seen in other regions where trade has necessitated accurate counting. This phenomenon, again, shows a correlative increase of grammatical and situational complexity (Grieve 2023, as discussed in Section 2.1 What is language typology?). One can also find a similar situation at the reconstructed protolanguage level, for example, in Proto-Hmong-Mien where only *ʔui* 'two' and *pjɔu* 'three' are of native origin, while the rest comprises borrowings from Sino-Tibetan languages (Ratliff, 2010).

      The ultimate origin of the numeral classifier system can be disputed, but Sinitic languages are often considered as the spreader on the one hand, maybe simply because of their long attestation. On the other hand, Tai-Kadai languages have also been proposed as the source of this grammatical system due to the robust geographical distribution and functional range observed in their classifier system, compared to other language families in the area (Jones, 1970; Lu, 2012; Honkasalo & Yurayong, 2024). When focusing on the aspect of multilingualism, however, it is less problematic to assume that two major civilisations, Chinese and Iranian, have significantly influenced the development of an exact and adjustable counting system with classifiers in their neighbouring languages. The subsequent discussion presents results from a survey of counting strategies across 543 distinct linguistic varieties spoken in Eurasia by Honkasalo and Yurayong (2024). **Figure 3** illustrates a distribution of numeral classifiers in the dataset to foreground further discussion regarding the multicultural society of Eastern Eurasia. From the map, it is noticeable that the classifier system is robust in East Asia, dominated by Sinitic languages, and disperses as one moves towards peripheral areas. Two contact scenarios in the Sinosphere and the Iranosphere are discussed next.
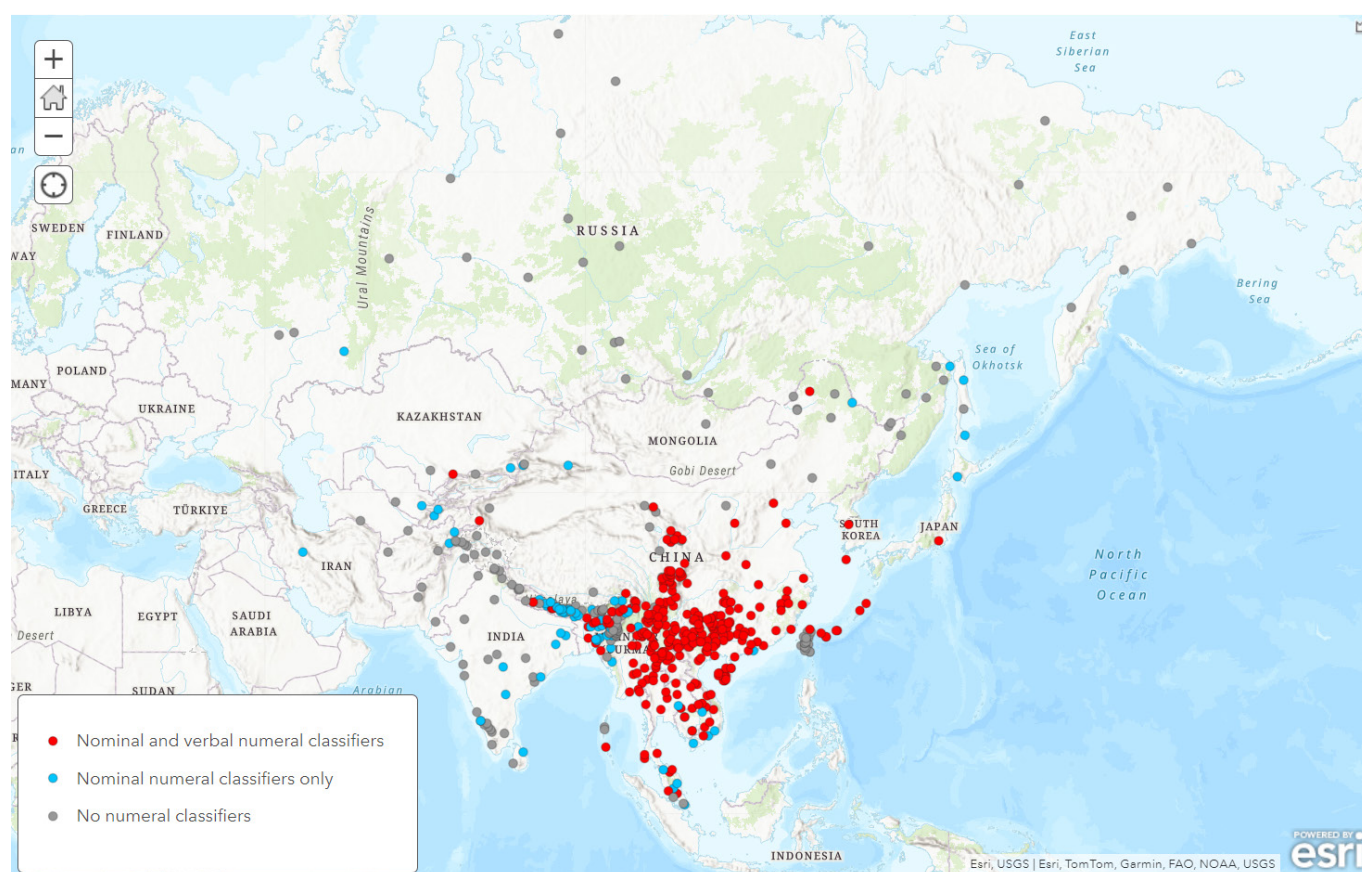

Figure 3 Geographical distribution of Eastern Eurasian languages with(out) numeral classifiers.

      As for Sinitic, which are a language group with the most robust classifier system in Northeast Asia, their contribution to the development of classifier systems in neighbouring languages is evident. In the northeast, classifiers already occurred in the Pre-Old-Japanese and Early Middle Korean stages, but their use and inventories noticeably became more robust after coming into contact with Sinitic languages in the

mid-1st millennium CE (Yurayong & Szeto, 2020). For instance, the Sinitic generic classifier 個 has entered and become the general counting devices for items, particles and things in Japanese 個 *ko* and Korean 개 *kae*. For Japanese, the Sinitic generic classifier adds to the native counting system with the Japanese generic classifier -つ *-tsu*, which goes back to Proto-Japonic **-tu*.

In the northwest, meanwhile, Sinitic languages have introduced a classifier system to Mongolic and Turkic languages, which prototypically had no dedicated devices for sortal classification (Sandman & Di Garbo, 2023). This is an obvious case of borrowing both lexical items and grammatical patterns from Sinitic. The following Mangghuer examples (**13**) illustrate that the quantifier phrases are etymologically Sinitic words: a nominal classifier *-tiao* from 条 *tiáo* for counting stick-like entities such as roads (**13a**), and a verbal classifier *-zhuan* from 转 *zhuàn* for counting rotations (**13b**). This goes beyond structural convergence (as discussed in Section 2.2 How do we study transition of language structure typologically?), as it involves the borrowing of both lexical items and patterns, the former of which possibly motivates the latter.

(13)　a.　*niker*　*jiura*　　　*ning-du*　　*mer*　**liang-tiao**　*bang, ...*　　　**Mangghuer**
　　　　　　this　　time　　　　this-DAT　　[road　**two-CLF.stick**]　COP
　　　　　　'Now here there are two roads, …'
　　　b.　*ni*　*zhaler*　　　**yi-zhuan**　*mergu*　*nuqi*　　*ri-ku, ...*
　　　　　　this　hired_farmhand　[**one-CLF.circle**　kowtow　pass　　come-IPFV]
　　　　　　'When this hired farmhand came back (after) kowtowing for one circuit around the temple), …'
　　　　　　(Slater, 2003)

At the same time, the grammatical system of multilingual Sinitic speakers has also been influenced by the lack of numeral classifiers in the neighbouring languages. Such a transition is observed in Northwestern Mandarin contact varieties, including Dungan, Tangwang and Wutun, where inventories of nominal numeral classifier systems have considerably reduced. The extreme case is illustrated by that of Dungan, where a generic nominal classifier derived from 个 is the only remaining item (Honkasalo, 2024). The lack of such paradigmatic contrasts no longer qualifies for prototypical nominal classifier systems and can be compared with a language with only one grammatical gender (as discussed in Section 3.1 Grammatical gender).

In Central Asia, Iranian languages have spread the use of numeral classifiers to many Turkic languages. For instance, Uzbek uses *ta* as a generic classifier as well as *dona* for inanimate items and *nafar* for humans, all of which commonly have Persian etymologies (Levy-Forsythe & Kagan, 2022). In the Altaic context concerning Mongolic, Tungusic and Turkic languages, this borrowing scenario has left its watersheds along the ancient Silk Road network across East and Central Asia, as illustrated in **Figure 4**. The lack of sortal classifying devices in Altaic languages outside the Silk Road network provides an implication supporting the idea presented above that trade is one motivating factor for a language community to develop a more exact and adjustable quantification system.

From the discussion above, the role of language contact in the emergence of counting strategies is especially pronounced along the Silk Road trade network since the early historical era, both over land and sea. This historical context has introduced a new context of use, thereby enriching the counting system and increasing both grammatical and situational complexity (as discussed in Section 2.1 What is language typology?). By cross-linguistically examining their origins and distributions, it becomes evident that counting devices are innovations which spread across languages from specific cultural centres with socioeconomical dominance in the regions concerned.
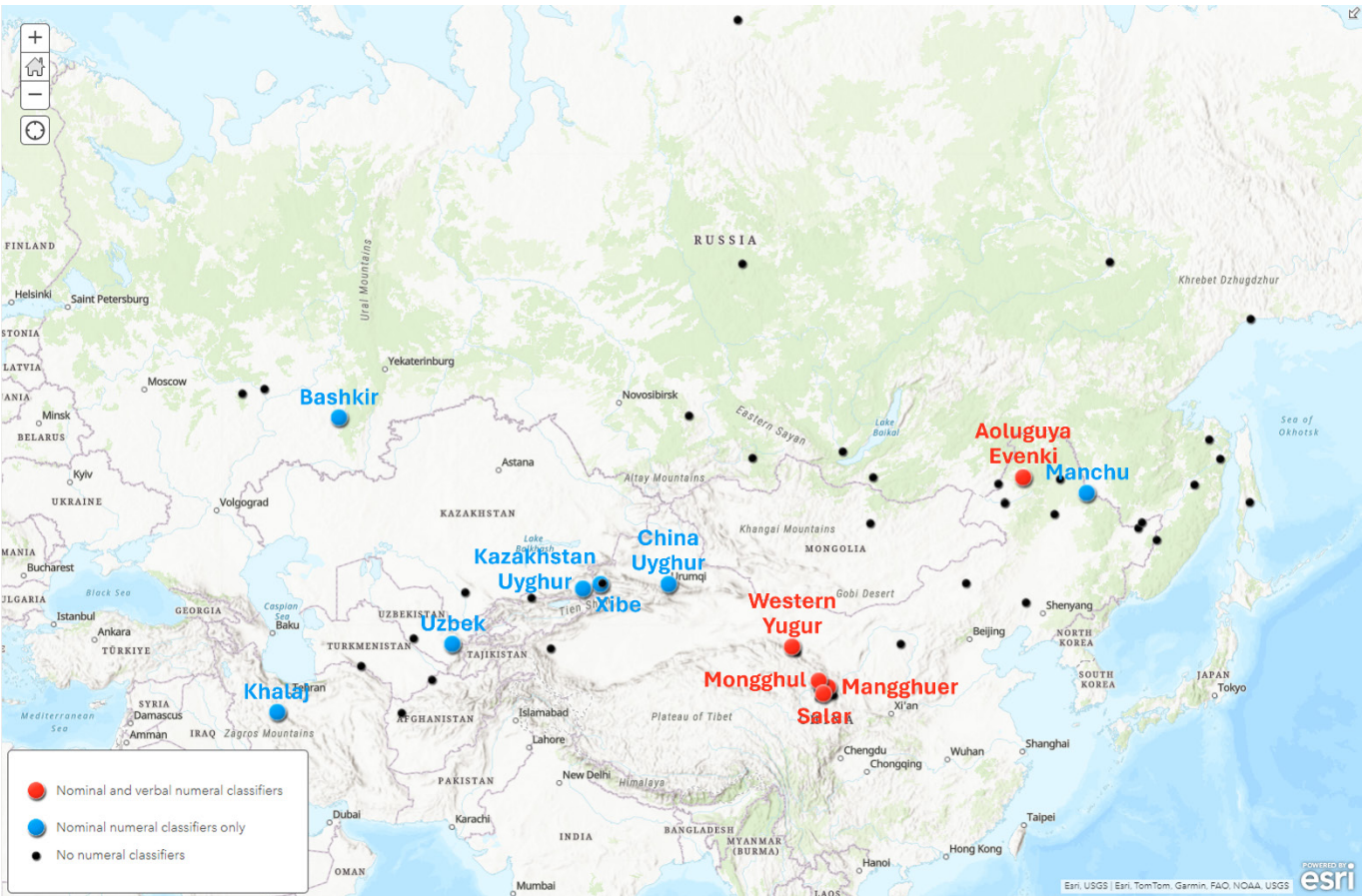
Figure 4 Geographical distribution of Mongolic, Tungusic and Turkic languages with(out) numeral classifiers.

In any case, there are also instances of classifier subsystems which develop independently of contact but are rather inherited from protolanguage or motivated by communicative needs. A study on gender-specific classifiers in 100 languages by Song and Yurayong (2019) has shown that Tai-Kadai languages have a robust gender-specific classifier system, as seen in Longchow Zhuang (**14**).

(14)	a.	*sa:m³³*	***po¹¹***	*luk²¹*	*ba:u⁵⁵*	**Longchow**
		three	**CLF**	child	male	**Zhuang**
		'three sons'				

	b.	*so:ŋ³³*	***me¹¹***	*pi¹¹*	*sa:u³³*	
		two	**CLF**	elder_sibling	female	
		'two sisters'				
		(Massupong, 1982)				

The prominent tendency of gender-distinctive classifiers in Tai-Kadai languages may have influenced Assamese whose multilingual speaker population comprises also those who previously shifted their language from Ahom and other Tai-Kadai languages of Northeast India. Consider the Assamese gender-distinctive classifier system: -জন *-zɔn* for male and -জনী *-zɔni* for female, both derived from Sanskrit जन- *jana-* 'person.'

Meanwhile, Korean has grammaticalised gender nouns 놈 *nom* 'jerk' and 년 *nyŏn* 'bitch' into classifiers, as shown in example (**15**). These classifiers are added to the pre-existing system which did not

distinguish between two natural genders of the counted persons and only possessed gender-neutral human classifiers 명 *myŏng* (from Sinitic 名 'name'), 인 *in* (from Sinitic 人 'person'), 사람 *saram* and 분 *pun*.

(15)	a.	*namja*	*ne*	**nom**	b.	*yŏja*	*se*	**nyŏn**	**Korean**
		man	four	**CLF**		woman	three	**CLF**	
		'four jerks/men'				'three bitches/women'			

There is no attested contact history between Koreanic and Tai-Kadai speaking populations but only lexical resemblances due to Sinitic borrowings and a similar realisation of onomatopoeic words (Yurayong, 2020). Therefore, the communicative explanation is more suitable for emerging gender-specifying classifiers in Korean. One possible motivation for distinguishing genders might have originated from an extension to a specific discourse context, increasing the range of context, i.e. situational complexity (Grieve 2023, as discussed in Section 2.1 What is language typology?). This concerns, for example, pejoration and aggression where the explicitness of gender distinction is preferred for individuation and identifiability of referents. This process can find a parallel in emerging gender-distinctive classifiers in Thai, หนุ่ม *num*[11] 'young man' for male and นาง *naːŋ*[33] 'lady' and สาว *saːw*[15] 'young lady' for female, which often appear in humour and sarcasm. One piece of structural evidence for this communication-based explanation is that these Korean and Thai gender-specifying classifiers occur mostly with small and non-compound numerals, not in the context of high quantity and long numerals where gender-neutral human classifiers are preferred.

3.3 Negation
Negation is a universal property and can be grammatically expressed in all human languages (Dahl, 1976). This characteristic of negation indicates that it has always been a fundamental feature of language which has lived various stages of language development. Given this principle, the historical interactions of individual speech communities can also be examined through their expressions of negation.

On the one hand, there are tendencies in negation which are cross-linguistically valid. For example, it has been proposed that there is less need for temporal specification in negative utterances compared to positive ones because the negated content is contextually available in discourse and needs not be specified in full detail (Miestamo, 2005). This general tendency has been empirically and quantitatively tested with spoken language corpora of English, Finnish and Korean (Miestamo, Silvennoinen & Yurayong, 2024). Despite the three languages possessing diverse typological profiles (English as a morphologically simplex SVO language, Finnish as an agglutinative language with free word order, and Korean as an agglutinative SOV language), they mutually confirm this hypothesis. The results support the validity of this tendency across languages and also align with the idea of indifferent absolute complexity across languages (Shosted, 2006; Miestamo, 2017, as discussed in Section 2.1 What is language typology?).

On the other hand, intense interactions between speech communities can reshape their strategies for encoding negation. For instance, Yurayong, Szeto and Honkasalo (2024) have demonstrated that the temperate climatic zone of Eastern Eurasia manifests a transition of the alignment patterns across various negative constructions under influences from three dominant cultural spheres: 1) Chinese, 2) Mongolian, and 3) Iranian. The following presentation discusses results from an investigation of 130 distinct linguistic varieties spoken across temperate Asia, highlighting the intersection between multicultural society and typology of negation.

Focusing on a phrasal level, negative constructions can be divided into four different morphosyntactic contexts, as illustrated through example (**16**) from Kilen, a Tungusic language. These contexts are: 1) nominal negation (**16a**), 2) non-finite verb negation (**16b**), 3) finite verb negation (**16c**), and 4) imperative negation a.k.a. prohibitive (**16d**). This ranking follows the rationale of verbiness degree from least to most, according to which only imperative can be considered as a true verbal context, while finite and non-finite verbs can appear in nominalised forms, as has been discussed in the context of Altaic-type languages (Malchukov & Czerwinski, 2021; Janhunen, 2023).

(16)   a.   *bu*   *atilə*   **atɕi**-*dɛi*   *imaxa-wə*   *wa-mbi-mu.*   **Kilen**
          1PL.INCL   [net   **NEG**-INSTR]   fish-ACC   catch-HAB-1PL
          'We are fishing without a net.'

     b.   *bi*   *ti*   *tab-***rtɕə**   *mɔri-mə*   *jalu-jə.*
          1SG   [that   be_fat-**NEG**   horse-ACC]   ride-IPFV
          'I ride the horse that is not fat.'

     c.   *bi*   *nikan*   *gisu-mə*   *tatimi-***rtɕə**-*mi.*
          1SG   [Chinese   language-ACC   teach-**NEG**-1SG]
          'I do not teach Chinese.'

     d.   **atɕi**   *xəsu!*
          NEG   speak
          'We are fishing without a net.'
          (Zhang, 2013)

The four morphosyntactic contexts may be marked in four distinct manners in some languages, while other languages may encode two or more contexts with an identical construction. **Table 1** provides an overview of systems attested or reconstructed for representative languages of several language families spoken in Eastern Eurasia.

Table 1 Negative morphemes and constructions in different morphosyntactic contexts.

Language	Nominal	Non-finite	Finite	Imperative
Old Chinese	無 *ma* + NOUN	不 *pə* + VERB	不 *pə* + VERB	毋 *mo* + VERB
Proto-Tibetic	NOUN + ***myed**	***ma/myi** + VERB	***ma/myi** + VERB	***ma** + VERB
Middle Korean	NOUN + *ani-/eps-*	*ani* + VERB	*ani* + VERB	VERB + *mal-*
Proto-Tungusic	(NOUN + ?)	***e-** + VERB	***e-** + VERB	***e-ji** + VERB
Proto-Mongolic	NOUN + ***busu/ügei**	VERB + ***ügei**	***ülü/ese** + VERB	***bU** + VERB
Proto-Turkic	NOUN-***sIz**	VERB-***MA**	VERB-***MA**	VERB-***MA**
Old Iranian	***a(n)**-NOUN	***na** + VERB	***na** + VERB	***ma** + VERB

From a diachronic perspective, many transitions have taken place as we arrive at modern languages. Namely, the alignment patterns may have drastically changed from the baselines given in **Table 1**. The distribution of patterns among modern languages is provided in **Table 2**, with areal trends observed in **Figure 5**.

Table 2 Alignment patterns across negative constructions.

Pattern	Linguistic varieties
Nom = NFin = Fin ≠ Imp	Kinnauri and Eastern Tamang, Spoken Manchu and Xibe, Eastern Mongolic
Nom = NFin ≠ Fin ≠ Imp	Ordos and Dagur
Nom ≠ **NFin = Fin = Imp**	Kazakhstani Gansu Dungan, Tibetan, Northern Tungusic, Turkic, Central Asian Iranian
Nom ≠ **NFin = Fin** ≠ Imp	Sinitic, Koreanic, Nanaic Tungusic, Southern Mongolic, Sarikoli and Pashto

*Abbreviations: Fin = Finite verb construction; Imp = Imperative construction; NFin = Non-finite verb construction; Nom = Nominal construction
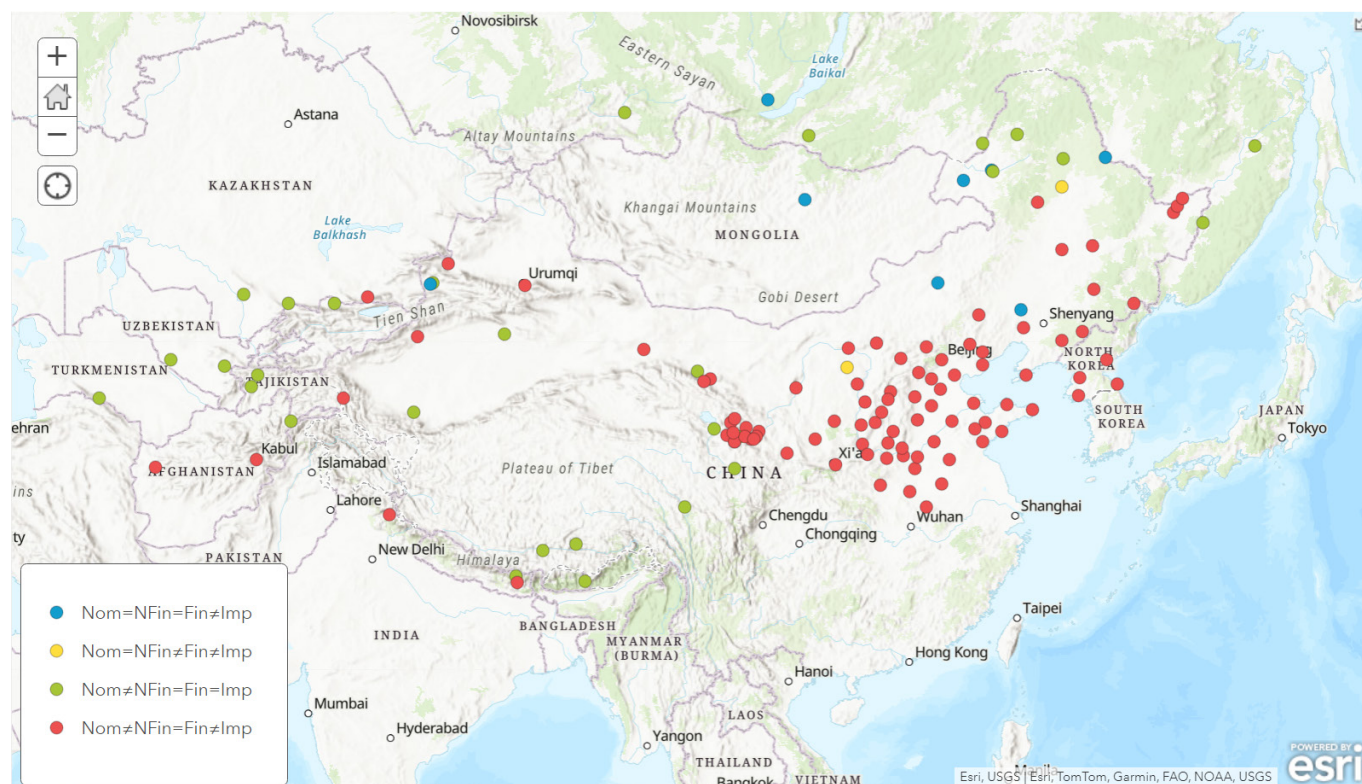
Figure 5 Geographical distribution of alignment patterns across negative constructions.

The subsequent discussion focuses on three scenarios which investigate the contact influences of Chinese, Mongolian and Iranian on their neighbouring languages, illustrating polysemy copying across languages in contact (a scenario discussed in Section 2.2 How do we study transition of language structure typologically?).

The Chinese culture has exerted significant influence, providing up to 74 % of lexical resources in the Modern Korean language, a phenomenon sometimes characterised as the 'Chinese Language Empire: ca 2nd BC to 19th century' (Bailblé, 2015). Alongside this heavy lexical borrowing, the Sinitic negative construction pattern Nom ≠ NFin = Fin ≠ Imp has introduced new possibilities to the pre-existing Koreanic system with the same pattern. Most obvious are prenominal negators 불- *pul-* and 무- *mu-* borrowed with Sinitic loanwords, such as 불안 **pul-an** (不安) 'unrest' and 무조건 **mu-jogŏn** (無條件) 'without condition, unconditional.' Nevertheless, the Koreanic pattern of employing negative predicates 아니 *ani* (negative copula 'not be') or 없- *ŏps-* (negative existential 'not exist/have') after a negated nominal is still actively used, such as 다름 아닌 *tarŭm ani-n* 'none other than' and 소리 없이 *sori ŏps-i* 'without noise, silently.' The Sinitic borrowing thus reinforces the asymmetry against the link between nominal and verbal negative constructions inherited from Middle Korean (see **Table 1**). It is worth clarifying that this study does not view Sinitic influence as causing a change in the distinction between nominal negation and negation in other contexts in Koreanic. Rather, the Sinitic contact has reinforced the pre-existing distinction which is inherited from Middle Korean. Furthermore, the borrowed nominal negators appear to be restricted in use, primarily limited to negating borrowed vocabularies and not typically used with native Koreanic lexicon.

The Chinese model Nom ≠ NFin = Fin ≠ Imp can also contribute to retention of the system inherited from a protolanguage. This is the case of Mongolic languages spoken in Gansu and Qinghai areas. Reinforced by a tendency of consistently using preverbal negators in neighbouring Sinitic languages as in (**17**) to (**19**), these Mongolic languages, Bonan and Eastern Yugur, have continued using Proto-Mongolic preverbal negators **ülü/ese*, as shown in (**20**) and (**21**).

(17) *fã*　　*imã*　*tʂhi = uã*　*sa,*　　**pɔ**　　　*ʂr̃ = kɨ,*　　*lãfɨ = tʂɤ.*　　**Zhoutun**

　　noodle　all　eat = COMP　PTCL　[**NEG.IMP**　leave = VIM]　waste = PROG

　　'Eat all the noodles! Do not leave them behind! It is a waste.' (Zhou, 2022)

(18) *tɕin¹¹ zəʔ²*　**pəʔ²**　*xa³⁵*　　*y⁴²*　　*lie¹¹,*　　*tʼiæ¹¹*　　*tɕ ʻ¹¹*　　*lie¹¹⁻²¹!*　　**Taiyuan**

　　today　　[**NEG**　down　rain　PERF]　sky　　　sunny　　PERF　　**Jin**

　　'Today it does not rain; it is sunny.' (Wang, 2007)

(19) **pu**　*tshĩʂə̃*　　*tɕɪ*　　*ŋixa*　　　　　　　　　　　　　　　　　　　**Tangwang**

　　[**NEG**　biological]　GEN　girl

　　'non-biological daughter' = 'not (their) own daughter' (Xu, 2017)

(20) *kozə*　*əsə*　*tʰəŋgətɕə*　*dappʰa*　**əsə**　*gə-sa*　　*pʰamba*　*ki.*　　**Bonan**

　　self　NEG　in_that_way　[faith　**NEG**　do-COND]　benefit　　NEG.COP1.SUBJ

　　'If a person doesn't do that, and doesn't have faith, (he) doesn't get the benefits.' (Fried, 2010)

(21) *tere*　*ja*　*arik-i*　　**l'**　　*uu-qi*　　　　*kün*　　*bai.*　　**Eastern**

　　3SG　TOP　[liquor-ACC　**NEG**　drink-PTCP.FUT]　person　COP　　**Yugur**

　　'He is a person who will not drink liquor.' (Nugteren, 2003)

Meanwhile, this Proto-Mongolic construction has been replaced by the nominal negation with the existential negator **ügei* in Eastern Mongolic languages spoken in Mongolia and Northeastern China (Janhunen, 2003; Brosig, 2015). Consider example (**22**) from Chakhar Mongolian spoken in Inner Mongolia, where a semantic verb *yawaa-* 'to go' is negated by the negative existential suffix *-gůů* plus an auxiliary finite copular verb *bää-* 'to be'.

(22) *nartïïn*　*tereg*　*bas*　*yawaa-***gůů**　　*bää-n.*　　　　　　　**Chakhar**

　　Bart　　bus　　yet　[go-**NEG.EXIST**]　COP-DUR　　　　　　**Mongolian**

　　'The bus to Nart has not left yet.' (Sechenbaatar 2003: 185)

　　The Mongolian culture has, in turn, provided a model for the Manchus in many respects, including the writing system, in particular. Interestingly, Mongolian influence is also found in negative constructions. Firstly, modern spoken Manchu has dropped the use of the Proto-Tungusic negative verb **e-* in finite verb constructions, as retained in Oroqen (**23**), a related Tungusic language spoken in Northeastern China. In place of the Proto-Tungusic pattern, modern spoken Manchu has replaced it with a construction using non-finite participles and a postverbal existential verb *-(a)kU* (**24**). The spoken Manchu construction is interestingly identical to that of Eastern Mongolic languages in the adjacent area with a pattern Nom = NFin = Fin ≠ Imp, such as in example (**22**) above, pointing to a contact-induced change in spoken Manchu.

(23) *noonɪn*　　**e**-*ki-n*　　　　*janda-ra.*　　　　　　　　　　　　**Oroqen**

　　3SG　　　[**NEG**-AOR-3SG　sing-CNG]

　　'S/he does not sing.' (Whaley, 2023)

(24) *bi*　　*sim-be*　　*je-ce-le-***ko**.　　　　　　　　　　　　　　**Spoken**

　　1SG　[2SG-ACC　eat-EXT-PTCP.AOR-**NEG.EXIST**]　　　　　　　**Manchu**

　　'S/he does not sing.' (Zikmundová & Wa, 2023)

The Iranian languages with the pattern Nom ≠ NFin = Fin = Imp, meanwhile, have also affected nominal negation in Turkic languages whose speakers have adopted the Iranian culture since the early historical period (see the discussion on Turkic-Iranian contact history in Johanson & Bulut 2006). Specifically, Iranian negative prefixes *na-/no-*, *bi-/be-* and *bät-/bet-* adopted with Iranian loanwords have introduced a model of preverbal negators in nominal negation (**25a**). These negative prefixes are used alongside the Proto-Turkic privative suffix **-sIz* 'without' which mainly occurs with native Turkic lexicon (**25b**). Diachronically, this phenomenon dates back as far as Middle Turkic, as the prefixes were productively used in Chagatai texts (Schlussel, 2018).

(25)   a.   *shuekuer* 'thanks'   > ***na**-shuekuer* 'ungrateful'   **Uyghur**

         ümid 'hope'   > ***na**-'ümid* 'hopeless'

         namaz 'prayer, not-praying'   > ***bi**-namaz* 'not-praying'

         täläy 'luck, chance'   > ***bi**-täläy* 'unlucky'

         nam 'reputation'   > ***bät**-nam* 'bad reputation'

         qiliq 'habit'   > ***bät**-qiliq* 'bad habit'

       b.   *čegra* 'border'   > *čegri-**siz*** 'without borders'

         täm 'taste'   > *täm-**siz*** 'tasteless'

         šekär 'sugar'   > *šekär-**siz*** 'sugar-free'

         mädäniyät 'culture'   > *mädäniyät-**siz*** 'uncivilised'

         (de Jong 2007: 26, 27, 32)

The profound impact of Iranian languages is strikingly evident in Moghol, a Mongolic language spoken in Afghanistan. Drawing from Pashto as a model where the non-imperative negator نه *ná* consistently precedes the verb, the Proto-Mongolic negation system (as given in **Table 1**) is significantly simplified. Similarly to the Iranian model from Pashto, Moghol speakers use of the preverbal imperative negator *bi* (< **bU*) across all morphosyntactic contexts of negation alongside the two preverbal negators *la* (< **ülü*) and *sa* (< **ese*) (see also a description and examples in Ramstedt 1905; Weiers 2003). This transition streamlines the negation paradigm, in which all negators occur consistently in the preverbal position, including the Proto-Mongolic postverbal negative existential **ügei*, as illustrated in **Table 3**.

Table 3 Negation patterns in Moghol (Mongolic) and Pashto (Iranian).

Context	Moghol	Pashto
Nominal negation	*(u)gai/**bi*** + VERB	***ná*** + VERB
Non-finite verb negation	*la/sa/**bi*** + VERB	***ná*** + VERB
Finite verb negation	*la/sa/**bi*** + VERB	***ná*** + VERB
Negative imperative	***bi*** + VERB	*má* + VERB

One might speculate that the Iranian negative prefix *bi-* (as previously discussed and illustrated in example **25a**) could have indirectly reinforced the generalisation of the Moghol imperative negator *bi* due to their phonological resemblance. This accidental similarity may have well facilitated convergence towards intertranslatability (Romaine 1988, as discussed in Section 2.2 How do we study transition of language structure typologically?). In essence, this example illustrates the outcome of intense language contact, involving a language on the verge of becoming extinct and its speaker population shifting to speaking another language.

4. Conclusional remarks

The present study has demonstrated the utility of language structure as a tool for researching multiculturalism and transition of human civilisation. It underscores that in addition to lexical items and social aspects of language often discussed in cultural studies, the inherent structural properties of human language can be equally, if not more, informative about human past. This follows from the nature of structural changes, which are deeply rooted in human cognition and thus less susceptible to conscious manipulation during cultural transitions. Compared to material elements in culture, for instance, music and its associated concepts can undergo rapid shifts as communities adopt new cultural practices and aesthetics (Yurayong, 2019).

At a methodological level, the case studies have shown that traces of intercultural interaction extend beyond the lexical level and can influence also structural transitions within recipient languages. This transition occurs through the internal reorganisation and negotiation between occurring types of language structure available as models in the multilingual repertoire. However, the study has also highlighted that communicative needs in human interaction cannot be neglected when seeking explanation for the emergence or restructuring of individual language structures. It thus highlights an essential methodological point that one must not overuse the language-external explanations based on increasing situational complexity to describe all phenomena observed in a multilingual society. Understanding the internal composition of language structure, including structural complexity and inherent constraints, is also crucial for a successful application of this method (Thomason, 2010).

The framework practiced in the current study offers a methodological model which could potentially be applied analogously to other anthropological aspects of human civilisation, such as architecture, arts, belief systems, and the development of consumption and culinary cultures.

Author statement

References

Abbi, A. (2013). *A Grammar of the Great Andamanese Language: An Ethnolinguistic Study*. Leiden, the Netherlands: Brill.

Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Bailblé, O. (2015). A history of the Korean lexicon. *International Journal of Teaching, Education and Language Learning, 2*(3), 34-66.

Bisang, W. (1999). Classifiers in East and Southeast Asian languages: Counting and beyond (pp. 113-186). In Gvozdanović, J. (Ed.). *Numeral types and changes worldwide*. Berlin, Germany: Mouton de Gruyter.

Brosig, B. (2015). Negation in Mongolic. *Journal de la Société Finno-Ougrienne, 95*, 67-136.

Corbett, G. G. (2007). Gender and noun classes. In Shopen, T. (Ed.). *Language Typology and Syntactic Description* (pp. 241-279). Cambridge: Cambridge University Press.

Corbett, G. G. (2013). Number of genders. In Dryer, M. S., & Haspelmath, M. (Eds.). *The World Atlas of Language Structures Online*. Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.

Croft, W. (2003). *Typology and Universals* (2nd ed.). Cambridge: Cambridge University Press.

Dahl, Ö. (1976). Typology of sentence negation. *Linguistics, 17*(1-2), 79-106.

Eckert, P., & McConnell-Ginet, S. (2013). *Language and gender*. Cambridge: Cambridge University Press.

Enfield, N. J. (2005). Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology, 34*, 181-206.

Fenk-Oczlon, G., & Fenk, A. (2008). Complexity trade-offs between the sub-systems of language (pp. 43-65). In Miestamo, M., Sinnemäki, K., & Karlsson, F. (Eds.). *Language complexity: Typology, contact, change*. Amsterdam, the Netherlands: John Benjamins.

Fried, R. W. (2010). *A grammar of Bao'an Tu, a Mongolic language of Northwest China* (Doctoral dissertation). United States: The University of Buffalo.

Gerner, M. (2014). Verb classifiers in East Asia. *Functions of Languages, 21*(3), 267-296.

Gil, D. (2013). Numeral classifiers. In Dryer, M. S., & Haspelmath, M. (Eds.). *The World Atlas of Language Structures Online*. Leipzig, Germany: Max Planck Institute for Evolutionary Anthropology.

Greenberg, J. H. (1957). The nature and uses of linguistic typologies. *International Journal of American Linguistics, 23*(2), 68-77.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements (pp. 73-113). In Greenberg, J. (Ed.). *Universals of Language*. Cambridge: MIT Press.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Grieve, J. (2023). Situational diversity and linguistic complexity. *Linguistics Vanguard, 9*(s1), 73-81.

Haspelmath, M. (1998). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison (pp. 1-30). In Tomasello, M. (Ed.). *The New Psychology of Language Vol. 2: Cognitive and Functional Approaches to Language Structure*. New York: Routledge.

Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language, 86*(3), 663-687.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Heine, B., & Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Hellinger, M., & Bußmann, H. (2001). *Gender across languages: The linguistic representation of women and men*. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Honkasalo, S. (2024). Kazakhstani Gansu Dungan as a contact language: An analysis of Russian influence. *Languages, 9*(2), 59.

Honkasalo, S., & Yurayong, C. (2024). Verbal numeral classifiers in languages of Eastern Eurasia: A typological survey. *Open Linguistics*, *10*.

Horn, L. (2006). Implicature (pp. 2-28). In Horn, L. R., & Ward, G. (Eds.). *The Handbook of Pragmatics*. Oxford: Blackwell Publishing.

Jakiela, P., & Ozier, O. (2020). *IZA DP No. 13126 Gendered language*. Bonn, Germany: IZA - Institute of Labor Economics.

Jakobson, R. (1958). Typological studies and their contribution to historical comparative linguistics (pp. 17-35). In *Proceedings of the 8th International Congress of Linguists*. Oslo, Norway: Oslo University Press.

Janhunen, J. (2003). Proto-Mongolic (pp. 1-29). In Janhunen, J. (Ed.). *The Mongolic Languages*. London: Routledge.

Janhunen, J. (2007). Typological interaction in the Qinghai linguistic complex. *Studia Orientalia, 101*, 85-102.

Janhunen, J. (2023). The unity and diversity of Altaic. *Annual Review of Linguistics, 9*, 135-154.

Johanson, L., & Bulut, C. (2006). *Turkic-Iranian contact areas: Historical and linguistic aspects*. Wiesbaden, Germany: Harrassowitz Verlag.

Jones, R. (1970). Classifier constructions in Southeast Asia. *Journal of the American Oriental Society, 90*(1), 1-12.

Jong, F. D. (2007). *A grammar of modern Uyghur: Morphology*. Utrecht, the Netherlands: Houtsma.

Kopitar, J. B. (1829). Albanische, walachische und bulgarische Sprache. *Jahrbücher der Literatur* (*Wien*)*, 46*, 59-106.

Kumar, P. (2012). *Descriptive and typological study of Jarawa* (Doctoral dissertation). India: Jawaharlal Nehru University.

Levy-Forsythe, Z., & Kagan, O. (2022). Classifiers and the mass-count distinction in Uzbek (pp. 685-704). In *Proceedings of the 32nd Semantics and Linguistic Theory Conference*. Mexico City, Mexico: El Colegio de Mexico and Universidad Nacional Autónoma de México

Lu, T. Q. (2012). *Classifiers in Kam-Tai Languages: A Cognitive and Cultural Perspective*. Boca Raton: Universal Publishers.

Luraghi, S. (2011). The origin of the Proto-Indo-European gender system: Typological considerations. *Folia Linguistica, 45*(2), 435-464.

Malchukov, A., & Czerwinski, P. (2021). Verbalization/insubordination: A diachronic syntactic isogloss in Northeast Asia. *International Journal of Eurasian Linguistics, 3*(1), 83-104.

Masica, C. P. (1976). *Defining a linguistic area: South Asia*. Chicago: Chicago University Press.

Massupong, C. (1982). *A grammar of the Tai dialect of Lung-Chow* (Master's thesis). Thailand: Silpakorn University.

Matthews, P. (1997). *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.

Miestamo, M. (2005). *Standard negation: The negation of declarative verbal main clauses in a typological perspective*. Berlin, Germany: Mouton de Gruyter.

Miestamo, M. (2017). Linguistic diversity and complexity. *Lingue e Linguaggio, XVI*(2), 227-253.

Miestamo, M., Silvennoinen, O. O., & Yurayong, C. (2024). Asymmetry in temporal specification between affirmation and negation: Adverbials and tense-aspect neutralization. *Studies in Language*. https://doi.org/10.1075/sl.23036.mie

Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

Nugteren, H. (2003). Shira Yughur (pp. 265-285). In Janhunen, J. (Ed.). *The Mongolic Languages*. London: Routledge.

Paul, H. (1886). *Prinzipien der Sprachgeschichte* (2nd eds.). Halle, Germany: Max Niemeyer.

Polinsky, M. (2015). *Tsez Syntax: A description*. Cambridge: Harvard University.

Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles 66*(3), 268-281.

Ramstedt, G. J. (1905). Mogholica: Beiträge zur Kenntnis der Moghol-Sprache in Afghanistan. *Journal de la Société Finno-Ougrienne 22*(4), 1-60.

Ratliff, M. (2010). *Hmong-Mien language history*. Canberra, Australia: Pacific Linguistics, College of Asia and the Pacific, The Australian National University.

Romaine, S. (1988). *Pidgin and creole Languages*. London: Longman.

Sandman, E., & Di Garbo, F. (2023). Contact-induced reduction, loss, and emergence of numeral classifiers (pp. 161-199). In Allassonnière-Tang, M., & Kilarski, M. (Eds.). *Nominal classification in Asia and Oceania: Functional and diachronic perspectives*. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Schlussel, E. (2018). *An Introduction to Chaghatay: A Graded Textbook for Reading Central Asian Sources*. Ann Arbor: Maize Books.

Sechenbaatar, B. (2003). *The Chakhar dialect of Mongol: A morphological description*. Helsinki, Finland: Société Finno-Ougrienne.

Shosted, R. (2006). Correlating complexity: A typological approach. *Linguistic Typology, 10*(1), 1-40.

Sinnemäki, K. (2020). Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach. *Journal of Historical Sociolinguistics, 6*(2), 20191010.

Slater, K. (2003). *A grammar of Mangghuer: A Mongolic language of China's Qinghai-Gansu Sprachbund*. London: Routledge.

Song, H., & Yurayong, C. (2019). Gender distinction in the East Asian human numeral classifiers: A case study of Korean and Thai. Poster presentation at the 13th Association of Linguistic Typology Conference, held at Università di Pavia, September 4–6, 2019.

Song, J. J. (2014). *Linguistic typology: Morphology and syntax*. London: Routledge.

Thomason, S. G. (2001). *Language Contact: An Introduction*. Edinburgh, Scotland: Edinburgh University Press.

Thomason, S. G. (2010). Contact explanations in linguistics (pp. 31-47). In Hickey, R. (Ed.). *The Handbook of Language Contact*. New Jersey: Blackwell Publishing.

Uchiyama, M. (2011). *Ilbonŏ-nŭn ppŏnhan p'aet'ŏn-ŭi panbokida advanced* [일본어는 뻔한 패턴의 반복이다]. Gyeonggi, South Korea: See&Talk.

Wang, W. (2007). *Jìnyuán fāngyán yánjiū* [晋源方言研究]. Beijing, China: Chinese Publishing House.

Weiers, M. (2003). Moghol (pp. 248-264). In Janhunen, J. (Ed.). *The Mongolic Languages*. London: Routledge.

Weinreich, U. (1974). *Language Contact: Findings and Problems*. The Hague, the Netherlands: Mouton de Gruyter.

Whaley, L. J. (2023). Orochen (pp. 184-205). In Vovin, A., Fuente, J. A. A. D. L., & Janhunen, J. (Eds.). *The Tungusic Languages*. London: Routledge.

Wiemer, B., & Wälchli, B. (2012). Contact-induced grammatical change: Diverse phenomena, diverse perspectives (pp. 3-64)*. *In Wiemer, B., Wälchli, B., & Hansen, B. (Eds.). *Grammatical Replication and Borrowability in Language Contact*. Göttingen, Germany: Mouton de Gruyter.

Xu, D. (2017). *The Tangwang Language: An Interdisciplinary Case Study in Northwest China*. Berlin, Germany: Springer.

Yurayong, C. (2019). Areal-linguistic approach to the diversity of the generic term for 'music' in the world's languages. *Rasprave Instituta za hrvatski jezik i jezikoslovlje, 44*(2), 675-691.

Yurayong, C. (2020). Cognate words in Korean and Thai from a historical-comparative perspective. *Korean Language and Cultural Studies Journal, 2*, 29-62.

Yurayong, C., & Szeto, P. Y. (2020). Altaicization and de-Altaicization of Japonic and Koreanic. International. *Journal of Eurasian Linguistics, 2*(1), 108-148.

Yurayong, C., Szeto, P. Y., & Honkasalo, S. 2024. Areal typology of Sinitic in the temperate climate zone: Focusing on adjustable quantification and negation asymmetry. In Xu, D., & Wang, C. (Eds.). *Chinese Language Contact and Typology*. Journal of Chinese Linguistics Monograph Series. Hong Kong: The Chinese University of Hong Kong Press.

Zhang, P. (2013). *The Kilen language of Manchuria: Grammar of a moribund Tungusic language* (Doctoral dissertation). Hong Kong: The University of Hong Kong.

Zhou, C. (2022). *Zhoutun*. London: Routledge.

Zikmundová, V., & Wa, G. (2023). Spoken Manchu (pp. 463-482). In Vovin, A., Fuente, J. A. A. D. L., & Janhunen, J. (Eds.). *The Tungusic Languages*. London: Routledge.