Language Variation, Change and Standardisation

Language Variation, Change and Standardisation

Edited by

Gintarė Judžentytė-Šinkūnienė

Cambridge Scholars Publishing

Language Variation, Change and Standardisation

Edited by Gintarė Judžentytė-Šinkūnienė

This book first published 2025

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-0364-5281-0
ISBN (Ebook): 978-1-0364-5282-7

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am deeply grateful to all the authors who have contributed to this book with great enthusiasm. I would also like to extend my sincere thanks to Liudmila Mockienė for her thorough work on language editing. Additionally, I am thankful to Cambridge Scholars Publishing for their initiative in endorsing the book proposal—particularly to Adam Rummens and Amanda Millar for their invaluable guidance throughout the entire process.

Gintarė Judžentytė-Šinkūnienė

INTRODUCTION: LANGUAGE VARIATION, CHANGE AND STANDARDISATION

GINTARĖ JUDŽENTYTĖ-ŠINKŪNIENĖ

This book aims to continue the dialogue with scholars interested in language change, variation, and standardisation, as well as to stimulate interest in language and activate dialogue between academics and the public on current language issues.

It is uncommon to encounter these three terms together in one sequence. Numerous linguistic works focus on pairs of these concepts, such as language variation and change (Labov 2010; Chambers and Schilling 2013), language variation and language standardisation (Stoeckle and Svenstrup 2011; McLelland and Zhao 2022), and language standardisation and linguistic change (Beaman 1980; Deumert 2004).

Such combinations are quite regular for a reason. For example, variation very often leads to linguistic change, or, in other words, linguistic change is the result of variation. When it comes to standardisation and variation, the focus is usually on eliminating variation or trying to minimise variation in the standard language. However, they do occur, especially in today's multilingual reality. The relationship between language change and standardisation rests on the well-known fact that standardisation aims to limit change (and variability), but change still occurs in standardised varieties. They are slowly accepted in case of success, but some of them can become standard or normative variants.

It is clear that these terms—“standardisation”, “variation”, and “change” familiar from sociolinguistic literature (Labov 1972; 2001; Haugen 1972; Trudgill 1974; 1986; Milroy et al. 1987)—are interconnected, overlapping, and influence each other at various stages in a language's life cycle. Therefore, it was decided that the book's title should reflect this interrelation.

It is also worth noting that some of the binary combinations of these terms involve more complex relationships. This volume addresses such complexities in several theoretical papers, for instance, those examining the relationship between language change and standardisation.

This volume addresses both theoretical and practical issues and is organised into two parts: (1) theoretical approaches to language variation, change, and standardisation; and (2) case studies focusing on language variation, change, and standardisation.

The first three papers, constituting the first part, feature researchers with expertise in language variation, change, and standardisation who argue that these terms are often ambiguous or too narrowly defined. The six papers in the second part focus on individual language cases and discuss one of these concepts in detail. Nonetheless, discussions of language standards frequently address issues of language change and variation, and analyses of particular languages and their variants must consider their interactions with the standard language.

This book primarily draws on data from Indo-European languages. One theoretical paper also discusses languages outside this group, such as Turkish. The case studies cover Baltic, Germanic, Italic, and Slavic languages—specifically English, Latvian, Polish, and Lithuanian. Maltese, a Semitic language derived from late medieval Sicilian Arabic with Romance superstrata, is also examined. Additionally, one case study compares Lithuanian and English data.

The contributors employ empirical data collected through their own research, as well as data from corpora such as *The Balanced Corpus of Modern Latvian* and *The Latvian Speech Recognition Corpus*, alongside fictional samples.

The studies presented employ a variety of methodological approaches, including experimental, dialectological, statistical, comparative, and qualitative methods of analysis.

Below is a brief overview of each contribution.

In the opening paper, **Pierluigi Cuzzolin** examines why the relationship between language change and standardisation is neither straightforward nor easily established. Starting from the principle that linguistic change differs from linguistic variation—since the former involves a process unfolding diachronically, while the latter concerns differences occurring synchronically—two conclusions follow: (1) linguistic change entails a transition from stage A to stage B, whereas variation occurs within a single chronological stage; (2) linguistic change presupposes the existence of variation within a language system, though variation does not necessarily lead to linguistic change.

Nicola McLelland proposes a four-dimensional framework for analysing language standardisation, urging a broader perspective that accounts for multilingual realities. While many studies focus on standardisation within a monolingual context, her approach emphasises the importance of recognising multilingual environments. She builds on recent research addressing multilingualism in standardisation and introduces a fourth dimension alongside the traditional diachronic, diatopic, and diastratic/diaphasic aspects. This additional multilingual dimension highlights the coexistence and interaction of multiple languages in the process of standardisation, offering a more comprehensive framework for understanding contemporary language practices in diverse societies.

In the third paper, **Albinas Drukteinis** offers a comprehensive review of the principles and criteria of codification as they have been articulated by Lithuanian applied linguists, alongside an examination of how these theoretical foundations have been put into practice. His analysis delves deeply into the complex relationships and hierarchical structures that exist between various codification principles and criteria, shedding light on areas where the existing codification theory remains underdeveloped or insufficiently emphasised. Furthermore, Drukteinis undertakes a critical evaluation of the specific criteria used for codifying particular linguistic phenomena, assessing their effectiveness and applicability. Importantly, he extends this evaluation by applying the discussed principles and criteria to emerging linguistic elements that have not yet undergone codification, thereby revealing potential directions for the further development and refinement of Lithuanian language norms and standardisation efforts.

The second part brings together scholars with major expertise in variation and standardisation. The difference between this part and the first one is that it deals with individual language cases, i.e., it discusses these terms in some Indo-European languages, namely, Latvian, Lithuanian, Maltese, Polish, and partly English.

Joseph M. Brincat focuses on bilingual Malta (Maltese and English are the official languages of the country), and more specifically on the Maltese variant used by its inhabitants. The article describes the long process of standardising the local language paving the way for its official status in Malta in 1934 and in Europe in 2003. At the end, the author concludes that the future of the Maltese language does not appear to be in danger and the community can be safely defined as bilingual with both languages (Maltese and English) being used freely in all areas. Its coexistence with widely used and prestigious languages, such as Sicilian and Italian, for almost a thousand years and with English during the last century has enriched its vocabulary and allowed it to keep pace with social and material progress. The author also adds that as long as the population finds the Maltese language useful in everyday activities, it will survive, just as it has for the last millennium, but social change is rapid these days.

In the next chapter, **Anna Mlekodaj** introduces the Podhale dialect spoken in the southern region of Poland and tries to explain why it cannot be standardised, although the process of standardisation of this dialect began at the beginning of the 20th century and is still going on. According to the author, the hidden reason is the importance of this dialect. The Podhale dialect played a special role during the Partitions as a living treasure of archaic forms of the Polish language. It was used by famous Polish poets at the turn of the 19th and 20th centuries, and at the same time local Podhale poets began to write in the dialect. The widespread recognition of the importance of dialect and its ennobling qualities leads the highlanders to perceive it as a deeply personal part of their identity. Consequently, they firmly believe that no one has the right to dictate how they should use it.

In the next paper, **Gintarė Judžentytė-Šinkūnienė** analyses the demonstrative pronouns in the Skuodiškiai subdialect of Samogitian. The study focuses on how the system of demonstrative pronouns in this subdialect relates to standard Lithuanian. It reveals notable generational differences: the younger speakers exhibit strong influence from the standard language, resulting in mixed deictic systems that combine elements of both the subdialect and standard Lithuanian. In contrast, the older generation predominantly retains the traditional deictic system, largely uninfluenced by the standard language. However, questions remain regarding how many speakers continue to use this original system. This intergenerational variation highlights important processes of dialect contact, change, and potential convergence, demonstrating how standard languages impact regional dialects, especially amid shifting sociolinguistic dynamics.

Andra Kalnača's paper delves into the normative aspects of direct address in Latvian, emphasising its crucial role in daily communication. She argues that the ways in which individuals are addressed—whether by first name, family name, kinship terms, profession, job title, or social status—are of great importance not only in spoken interaction but also in written communication. This normative dimension shapes social relationships and reflects cultural norms, politeness strategies, and hierarchy in Latvian society. Kalnača's study first offers a comprehensive overview of how forms of address have been traditionally described in Latvian grammars, highlighting any prescriptive recommendations or inconsistencies. Building upon this foundation, she then conducts an empirical analysis of actual language use, comparing observed communicative practices with the established norms to identify both preferred and prescriptive forms of address. The focus on personal names and common nouns referring to people underscores their pivotal function in facilitating effective and appropriate human interaction, making this investigation essential for understanding the interplay between language norms, social roles, and communication strategies in contemporary Latvian.

Dalius Jarmalavičius, **Danguolė Straižytė**, and **Virginija Jūratė Pukevičiūtė** explore the word-formation of place nouns in Lithuanian, drawing on data from two 20th-century novels, Sruoga's *Dievų miškas* and Kunčinas's *Tūla*. By comparing the original texts with their English translations, the study aims to identify common word-formation patterns of place nouns and examine how these are rendered in English. The analysis focuses on morphological features such as suffixes, prefixes, compounds, and derivations, highlighting predominant types within this category and addressing challenges in translation. This research contributes to a better understanding of Lithuanian morphological processes and cross-linguistic patterns in translating spatial terms.

The next paper focuses on the Latvian language consulting services provided by language consultants at the Latvian Language Agency. According to **Evelīna Zilgalve**, the agency receives about 800 phone calls and 1200 e-mails a month. Although most of the questions concern rendering of personal names in Latvian, the current study is based on requests from language users about orthography, punctuation, lexical units, and language style. The study concludes that a quarter of the questions are about word inflections, followed by lexical units and language style, orthography and punctuation. There are also occasional questions about capitalisation, terminology, text organisation, use of abbreviations, pronunciation, word order, and even grammar theory.

References

Beaman, Karen V. 1980. "Language Standardisation and Linguistic Change: A Pilot Study of Swabian Palatalization." *Proceedings from Southeastern Conference on Linguistics* (SECOL). Memphis: Memphis State University.

Chambers, Jack K., and Schilling-Estes, Natalie, eds. 2013. *The Handbook of Language Variation and Change*. Oxford: Wiley-Blackwell.

Deumert, Ana. 2004. *Language Standardisation and Language Change: The Dynamics of Cape Dutch*. Philadelphia: John Benjamins Publishing.

Haugen, Einar. 1972. *The Ecology of Language*. Stanford: Stanford University Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 2001. *Principles of Linguistic Change*, Vol. 2. Malden, MA: Blackwell.

Labov, William. 2010. *Principles of Linguistic Change*, Vol. 3. Chichester, West Sussex; Malden, MA: Blackwell.

McLelland, Nicola, and Zhao, Hui. 2022. *Language Standardisation and Language Variation in Multilingual Contexts: Asian Perspectives*. Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781800411562.

Milroy, Lesley. 2008. "Social networks." In *The Handbook of Language Variation and Change*, edited by Jack K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 549–572. Oxford: Wiley-Blackwell.

Stoeckle, Philipp, and Svenstrup, Christoph Hare. 2011. "Language variation and (de-)standardisation in Germany." In *Standard Languages and Language Standards in a Changing Europe*, edited by Tore Kristiansen and Niklas Coupland, 83–90. Oslo: Novus.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford and New York: Basil Blackwell.

PART 1:

THEORETICAL APPROACHES TO LANGUAGE VARIATION, CHANGE AND STANDARDISATION

THE PROBLEMATIC RELATION BETWEEN LINGUISTIC CHANGE AND STANDARDISATION[1]

PIERLUIGI CUZZOLIN

Abstract

As in any scientific discipline, the terminology employed in linguistics must be unambiguous. This means that whenever technical terms are used, they must always be clear beyond any doubt. From the present contribution, in which the relation between linguistic change and standardisation is discussed, it should be evident that even these labels, at first sight unambiguous and easy to understand, are not completely free from some ambiguity. What will be stressed is that by linguistic change is meant a change that takes place in diachrony and affects the structure of the grammar at some level. The linguistic change must be distinguished from the linguistic variation, a phenomenon that takes place in synchrony.

Standardisation is the process through which the linguistic variety becomes the reference variety, i.e., the standard, in a society, mainly due to social and cultural reasons. The standard is a socio-cultural construal, independent of the grammatical structures of a language.

The relation between these two notions will be illustrated by two examples: the former is taken from Latin and shows how the complement clause, expressed by the *Accusativus cum Infinitivo* (AcI) was successfully replaced by a more common pattern *Conjunction plus a clause with a finite verb* in a bottom-up process. The latter is taken from Turkish, namely Ottoman Turkish, and shows, on the contrary, an unsuccessful attempt to replace the original pattern of the complement clause through the pattern *ki* plus a clause with a finite verb of Persian origin, in a top-down process.

[1] I want to thank the organisers of the 27th Jonas Jablonskis' conference and especially prof. Gintarė Judžentytė-Šinkūnienė for their kind invitation. I also thank Nicola McLelland and Pietro Umberto Dini for their comments and Rosanna Sornicola that discussed with me some points of the contribution. All errors that remain are mine.

Apparently, Latin and Ottoman Turkish seem to show a similar cline of development, but a deep analysis of data shows that the two changes are very different.

Key words: linguistic change, variation, complement clause, standard, external factors.

Distingue semper (from the motto of the Jesuits)

1. Introduction

For any discipline that aims to be qualified as epistemologically correct there is a crucial tenet: notions and categories[2] employed in an investigation whatsoever must be clearly, i.e., unambiguously, defined. This is one of the fundamental principles founding the methodology of any discipline, and one of its corollaries is that identical labels must identify identical contents: this implies that any difference or variation, even the slightest one, must be enucleated, illustrated in its form and explained in its function(s). This is the way in which the content of any claim can be crystal-clear beyond any doubt. Any time a new linguistic object or entity is envisaged, a new label should be coined as to identify it in a precise unambiguous way and classify it. The methodological process just sketched corresponds to what has frequently happened in the history of all scientific disciplines, not only linguistics.

However, what could apparently be a truism—a scholar of any discipline always ought to follow the methodology just described—in fact turns out to be a procedure that is far from being regularly and constantly applied as such and, no matter how unexpected this observation can be, it also quite frequently receives no attention or is even disregarded by linguists. Practitioners of historical linguistics, and, more generally, linguists not always adequately reflect on the epistemological basis of their own discipline, but rather concentrate on certain aims and purposes, preferably adopting a problem-solving perspective within the theoretical framework they use.

One of the most remarkable exceptions is represented by Haspelmath, a typologist who has been working for years on the revision of the labels used in linguistics, i.e., what they designate, their content, and the way they have

[2] On the ambiguous notion of category, I have made my viewpoint clear in the paper *Del buon uso delle categorie in linguistica*, presented at the XLVI Annual Meeting of the Società Italiana di Glottologia, Cagliari 26–28 October 2022. An article based on the conference paper is now in press.

been used in the past decades by other linguists, independently of their theoretical orientation. Haspelmath has contributed to this desirable revision of the terminology employed in linguistics[3] with several papers in which he has proposed to use the "comparative concepts", whose purpose

> is to argue that crosslinguistic comparison should be based on comparative concepts created by the typologist, rather than on crosslinguistic categories that are instantiated in different languages, and to show how comparative concepts differ from language-specific descriptive categories. Although in practice typologists generally work with such special comparative concepts, this distinction between comparative concepts and descriptive categories has not been articulated clearly before (Haspelmath 2010, 663).

As this quotation suggests, in Haspelmath's view a clear terminology would clearly and unambiguously reflect the content of the descriptive categories. Haspelmath primarily refers to the typologists for two reasons: first, because he himself is a leading figure in the field of the typological linguistics; second, because typologists usually work with a huge number of languages of different types and frequently describe unwritten languages for the first time. Therefore, the fact that they very often need new labels to adequately capture the phenomena they illustrate should not be surprising. For instance, if in a language never described before the typologist identifies a verbal form used to give orders and commands, it is not unexpected if he/she labels that verbal form as "imperative", even though the label "imperative" was coined by Roman grammarians. However, the overextension of an identical traditional label to similar, but not totally identical, phenomena occurring in typologically different languages almost inevitably creates new problems[4].

Nonetheless, Haspelmath's claim implicitly concerns practitioners working in any field of linguistics such as, for instance, historical linguistics or general linguistics because the labels used are frequently the same, unfortunately not always conveying the same meaning.

Recently, Haspelmath has reaffirmed his point of view in a paper from which I draw some lines where a couple of points relevant to the reasoning presented here are very clearly formulated:

---

[3] "Linguists' traditional terminology is often based on stereotypes that arose through complex historical processes, and these do not always result in very useful concepts" (Haspelmath 2022, 210).

[4] Obviously, there would be more to say on this point, but it would go beyond the aim of the present paper. I will devote a specific contribution to this topic elsewhere.

> In a view of general linguistics that emphasizes the contrast between general (language-independent) concepts and language-particular categories, what exactly is the value of clearly defined terminology? I see it as twofold: On the one hand, technical terms that are widely used anyway should have a clear meaning, regardless of how useful they might end up being, and the terms in (1) are widely used. On the other hand, in order to test universal claims about the world's languages, we need comparative definitions that allow us to identify phenomena uniformly (Haspelmath 2022, 210; see also 2021).

From this quotation, it should be clear that the problem also concerns not only labels that refer to grammatical categories such as "nominative" or "accusative" but also terms or labels that do not refer to any linguistic category and, therefore, could be perceived as more neutral with respect to any theoretical framework and independent of any model. For instance, the label "linguistic change" apparently does not raise any problem because it is simply used to describe changes involving the structure of a language whatsoever. In Haspelmath's view it would not belong to any language-particular category and could not even be strictly considered a technical term. In the last decades linguists have described changes that mainly and more frequently have occurred at morphological and syntactic level, but phonology or semantics show changes as well. At any rate and broadly speaking, linguistic change basically means that a linguistic element, be it a single item or a pattern, has undergone a change and has become something different from what it was before. Therefore, the label "linguistic change" would simply generalise the observation that X has become Y.

Apparently, this case is simpler than others illustrated in literature because it seems to be easily applicable to any change in any language regardless of the linguistic type of the language in which such a change occurs. Furthermore, according to Haspelmath's perspective, there would be no need for a comparative concept to highlight it adequately. However, even a simple label can hide difficulties not immediately evident, and this is what precisely happens in the case of "linguistic change".

Very often this label has been interpreted as a simple variation in the language. However, this interpretation is simplistic and inaccurate because it does not take into account one crucial point, i.e., that linguistic change is a label improperly used whenever it is applied to describe variation in synchrony. I will address this crucial point below.

2. On the notion of "standard"

In linguistics, the notion of "standard" is also far from being crystal-clear, since this label has been employed in literature to describe different linguistic varieties from different viewpoints. Nonetheless, what is worth observing and crucial for the claim put forward in the present paper, is that the notion of standard has been typically worked out within the sociolinguistic framework. This means that standard cannot be defined on the basis of intrinsically structural properties of the linguistic system: standard can only be a social construal, a concept based on function and context, but not on form (Berruto 2007, 13). Therefore, the necessary consequence of this claim is that linguistic change and phenomena of standardisation involve two different levels of analysis that only partially overlap, the former being intrinsically structural and the latter—intrinsically sociolinguistic. The story of the relation between these two notions must always be treated with caution because the fundamental point in question is how to relate the structural phenomena of a language and their development to their inherent sociolinguistic aspects.

Obviously, as already observed, the questions raised by Haspelmath in the second quotation above are also legitimate and decisive in the field of historical sociolinguistics, and specifically at a macro-sociolinguistic level of analysis. At the same time, however, another question is in order, strictly tied with Haspelmath's proposals: Do the labels "standard" and "standardisation" have the same meaning for all languages? In other words, do these labels belong to the universals of linguistics in Coseriu's sense (Coseriu 1975, 392), or not?

In the following parts, I will try to provide an answer to these questions.

3. Defining the notion of standard

In my opinion, the statement number 1, i.e., "technical terms that are widely used anyway should have a clear meaning, regardless of how useful they might end up being" is accurate. By the label standard, scholars generally agree that it is a linguistic variety selected among other varieties due to some external factors that have made it most prestigious. The reasons for that choice can be different but I am aware that in any culture the decisive factors to select a linguistic variety and to determine its prestige as to make it the reference variety are always the same. The standard becomes:

- the variety endowed with the richest literary results or, broadly speaking, cultural products, and, therefore, the most prestigious one;

- the variety that has a symbolic value in so far as it is identitarian of a nation;
- the variety used by the cultural elite and taught in school. In this case one could also be entitled to claim that ideology is strongly involved.

It would be possible to illustrate each of the cases just listed above by means of several examples.

3.1. Sociolinguistic viewpoints on standard

There is a consensus among linguists on the notion of standard, but those sociolinguists who have proposed definitions of the concept have underlined different characteristics that are supposed to capture what is, or should be, peculiar to it. Obviously, not all the definitions proposed are equally complex. To the best of my knowledge, the simplest one was put forward by Finegan-Besnier, which stresses the fact that standard represents the official variety, the one on which all reference texts are based and in which they are codified:

> A standard language variety is a variety that has been designated as such and for which a set of forms has been identified and codified in dictionaries and grammars (Finegan-Besnier 1989, 496).

However, it is also interesting to observe that this definition employs the word "designated" which can only mean that standard is the result of a choice of a set of forms made by somebody based on non-linguistic factors. Usually, the people in charge to designate are members of a committee of linguists and sociolinguists invested with authority by some cultural or political institution.

A more refined and complex definition is the one proposed by Dittmar:

> "Standard" wird geschrieben, er ist schriftlich kodifiziert (System von Vorschriften), besitzt überregionale Rechtweite und Gültigkeit, wird vorzugsweise in institutionellen Kontexten und offiziellen Kommunikationssituationen benutzt und erscheint in der Alltagssprache (= Summe der Varietäten in einem bestimmten Varietätenraum) niemals in ihrer idealtypisch kodifizierten Norm … Der Standard wird in den Schulen gelehrt, sein Gebrauch verschafft in der Regel Prestige und *begünstigt den Erwerb sozialer Privilegien* (Dittmar 1997, 201; my emphasis).

In this definition several aspects essential to the concept of standard are listed:

- Standard is a written variety of language. With respect to Finegan-Besnier's definition given above, Dittmar's definition

underlines one crucial fact: there are many languages for which linguists have provided a grammar and a dictionary, but they still remain exclusively spoken. Since, according to Dittmar's viewpoint, standard always presupposes a written code, the notion of standard cannot be applied to exclusively spoken languages.
- Standard also possesses validity that goes beyond regional borders: its general validity makes it the variety used in institutional contexts and in official communicative situations.
- When standard elements occur in everyday speech they never occur in their ideal codified form.
- Finally, since standard is the variety that is taught in school, it is also the variety of prestige, the one that favours the acquisition of social privileges.

Even in Dittmar's definition none of the peculiar features of standard refer to the linguistic structure: standard is completely independent of form. However, since it is also the most refined and structurally complex variety employed by writers and poets, formal considerations are not completely out of the picture. Even in this case, formally complex structures are selected based on external, i.e., cultural, factors, rather than created by the standard.

3.2. Negative viewpoints on standard

It must be added that some linguists, instead of providing a definition of standard, have rather preferred to express negative comments. On the one hand, for instance, Hudson underlines that the variety represented by standard is, among other things, lacking in diversity, which makes it close to something pathological and, therefore, not interesting, because diversity is proper to any human language:

> The *unusual* character of standard languages, which are perhaps the *least* interesting kind of language for anyone interested in the nature of human language… For instance, one might almost describe standard languages as pathological in their lack of diversity (Hudson 1996, 34; original emphasis).

On the other hand, Romaine has expressed a highly negative evaluation of the notion of standard because it is a product deeply rooted in the ideology of the ruling elite (Romaine 1994). The inevitable consequence of the viewpoint upheld by Romaine is that, since standard is seen as a powerful instrument for inequality, standardisation, as an unnatural process, in Romaine's opinion, crucially ends up being the result of an artificial process in which ideological factors prevail, producing social injustice because accessibility to the standard also means to have the possibility for better conditions of life:

> The process of standardisation… is one of the main agents of *inequality*… Standard languages do not arise via a 'natural' course of linguistic evolution (Romaine 1994, 84; my emphasis).

If a conclusion is to be drawn from the above definitions and their different perspectives, standardisation could be defined as a process by which a linguistic variety endowed with a written code[5] is established among several others as the reference variety, i.e., *the standard*. The choice is made on the basis of cultural, social, and even economic factors. Standard becomes the instrument against which correctness of written codes in use is tested and evaluated. In any case, it is to be stressed that none of the definitions cited above is based on, or simply make reference to, the structure of the language, but they all exclusively employ sociolinguistic arguments. Obviously, linguistic structures are implicitly involved whenever the linguistic complexity, patterns or rules are mentioned in order to distinguish the features proper to written codes with respect to oral codes.

However, these remarks rather lie in the background without any real contribution to the discussion concerning the notion of standard.

I would like to conclude this section with a quotation from one of the most important contributions to the debate on the concept of standard:

> Certo, lo standard è, linguisticamente, complicato e artificiale. Di solito lo standard è elaborato, appunto, il che significa strutturalmente e lessicalmente più complesso del non-standard e del sub-standard, appunto perché si trova a dover rispondere a una gamma di esigenze molto ampia, globale, in particolare per quanto riguarda i domini elevati di impiego (Berruto 2007, 20).

Berruto also insists on the importance and usefulness of standard despite the negative evaluations and criticisms mentioned above:

> Lo standard serve a molto. Non obbedisce solo a motivazioni pratiche, in primo luogo avere una norma di riferimento fissata, stabile e condivisa nel comportamento linguistico, e permettere di coprire e soddisfare la gamma più vasta possibile di impieghi; ma rappresenta anche la spinta all'unitarietà che è in perenne dialettica […] con l'opposta tendenza alla differenziazione.

[5] To best of my knowledge, the issue of the relation between the features that qualify the standard as such and the peculiar features of the spoken varieties of language still remains underinvestigated. But this topic is beyond the aim of the present paper.

> Ha utilità pratica, utilità sociale e funzionale, e utilità simbolica (Berruto 2007, 21).

4. Linguistic change and linguistic variation

It would be relevant now to go back to the crucial point of distinguishing a real linguistic change from a variation in language. Variability, i.e., the possibility for a language to show variation at any level, is intrinsic to any definition of language. The basic difference between variability and its instantiation, i.e., actual variation, is that variability corresponds to various synchronic options in the language, whereas linguistic change always implies diachrony.

The point refers to a distinction not always easy to draw, since variation and variability are peculiar to any language. Obviously, the fact that a language shows variation does not imply that there is also change. More than one century ago Meillet (Meillet 1912) had convincingly showed that in French the verb *être* ('to be') possessed different meanings associated with specific contexts. This case perfectly illustrates what linguistic variation is. Here is the quotation from Meillet's paper:

> Par example *suis* est un mot autonome dans la phrase, du reste très artificielle, *je suis celui qui suis*, et a encore une certaine autonomie dans une phrase telle que: *je suis chez moi*; mais il n'est presque plus qu'un élément grammatical dans: *je suis malade*, *je suis maudit*, et il n'est tout à fait qu'un élément grammatical dans: *je suis parti*, *je suis allé*, *je me suis promené*, où personne ne pense ni ne peut penser à la valeur propre de *suis*, et où ce que l'on appelle improprement l'auxiliaire n'est qu'une partie d'une forme grammaticale complexe exprimant le passé. Il est pourtant claire—et l'histoire de la langue montre de manière évidente—que *suis* est dans *je me suis promené* le même mot que dans: *je suis ici*; mais il est devenu une partie constituante d'une forme grammaticale (Meillet 1912, 131).

According to Meillet, the case when an autonomous and meaningful word acquires a grammatical character (*caractère grammatical* in Meillet's words) represents the typical case of grammaticalisation. However, distribution of the verb *être* and its grammaticalisation show the typical situation of variation, i.e., the situation in which the function and the semantics of the verbal form are strictly tied with the context in which they occur.

The viewpoint adopted here is that any linguistic change always implies diachrony, i.e., any real change implies that X, whatever may it be and independently of the level of the analysis it belongs to, be it a syntactic pattern, a morpheme or even a phonological trait, in a reasonable span of time is replaced by Y. Implicitly, this replacement corresponds to the transition from a stage A with X to a stage B with Y. The most common situation is the one in which Y ousts out X from the system of the language, but it is not the only one that can be thought of and that is documented in the history of language.

In the second part of this contribution, I will illustrate both situations with examples taken from two genealogically unrelated languages but with some typological similarities, namely, Latin and Ottoman Turkish.

5. Standard and sub-standard

In my opinion, the statement number 2, i.e., "in order to test universal claims about the world's languages, we need comparative definitions that allow us to identify phenomena uniformly" is correct. As repeatedly claimed above, language change is a structural phenomenon implying a diachronic dimension. Causes of the change *can coincide* with some external factors, but the external factors *are not responsible* for shaping different forms among which the change takes place. In other words, they can only be crucial for selection of forms, which, in addition, can also have a negative side effect of narrowing the input of the speakers, with a consequent reduction of their linguistic competence.

Moreover, the label of "standard" is currently employed alongside the label of "non-standard" in literature, and even in this case, its meaning is not unambiguous. As illustrated by Berruto (2007, 36–37) the label of "sub-standard" is used in literature in at least four different meanings:

- A: sub-standard = non-standard;
- B: sub-standard = just below the standard;
- C: sub-standard = a variety between standard and dialect(s);
- D: sub-standard = close to the standard (≈ *Umgangssprache*).

It is interesting to observe that point A is preferably used within the European tradition.

5.1. On the process of standardisation

As already stated above, standard is a notion that typically belongs to sociolinguistics, a social construal based on function, not on structural properties of the language[6].

[6] Obviously, there is another crucial issue concerning the reason why standard is characterised by certain features occurring in any culture in which that standard is acknowledged. This point has also been mentioned by Berruto: "[…] la questione di che cosa viene selezionato e perché nello standard e per lo standard, di quali sono le

Since linguistic change and standardisation involve two different levels of analysis that only partially overlap, the former being intrinsically structural and the latter—intrinsically sociolinguistic, it should be clear that some conclusions must be drawn by the practitioner of historical linguistics. In my opinion, the first and most important conclusion is that the process of standardisation simply registers that a linguistic change makes a new element or a new pattern part of the reference variety, although not any linguistic change necessarily undergoes such a process.

It is convenient to illustrate this point through a couple of examples, taken from the history of two languages belonging to two different linguistic families, i.e., Latin (Italic, Indo-European), and Turkish (Turkic; Altaic?), whose histories clearly show how the interaction between linguistic change and standardisation could develop differently and how problematic this interaction can be for an accurate description of the scenario. At first sight, their histories have an identical linguistic change in common: an explicit complement clause occurring after a verb of saying or thinking replaces an implicit complement clause that precedes the verb of saying or thinking. However, a more accurate comparison and a deeper analysis of the change involved lead to the conclusion that the claim that Latin and Turkish underwent the same phenomenon is incorrect.

6. A case study from the history of Latin

As well know, in Latin the regular way to express the complement clause depending on a verb of saying or thinking was the pattern traditionally called *Accusativus cum Infinitivo* (from now on indicated as AcI): the subject of the complement clause regularly occurred in the accusative case and the verb was in the infinitive form, more precisely, in the present tense, if the action, or the state, took place simultaneously with the tense of the governing clause; in the past tense, if the action had already taken place before the tense of the governing sentence; in the future tense, if it would take place after it. A marginal but interesting remark is that the AcI is a very rare pattern among the world's languages, and both its origin and its structure are difficult to explain.

Just to give an example of the pattern, one could quote the following one, taken from Cicero's work *Pro Archia* 19:

1) *Homerum Colophonii civem esse dicunt suum.*
   'Colophon asserts that Homer is her citizen.'

---

scelte dello standard nella gamma di variazione di una lingua" (Berruto 2007, 20). I will treat this issue elsewhere.

In the history of Latin, the AcI was the exclusive way to express the complement clause until the 1st century BCE, when a new syntactic pattern started to occur in competition with it and, even though very slowly, started to spread. The new pattern was introduced by two conjunctions, *quod* and *quia*, both meaning 'that' and characterised by finite verbal forms. This new pattern was structurally identical to the other subordinate clauses, i.e., it consisted of an introductory conjunction with the subject occurring in the nominative case and a finite form of the verb.

In the syntax of Latin, the AcI, because of its antiquity traceable back to the Proto-Indo-European stage, represented a typological anomaly (see Lehmann 1989). The structural difference between the two patterns had also relevant consequences for pragmatics.

The first example recorded in which the new pattern occurs after a verb of saying belongs to a work traditionally but wrongly attributed to Julius Caesar, the *Bellum Hispaniense* ('On the Hispanic War'), composed around 46 BCE by Aulus Hirtius. Here is the example:

2) *Bell. Hisp.* 36, 1; tr. Way 1955
dum haec geruntur legati Carteienses renuntiauerunt quod Pompeium in potestatem haberent
'In the course of these proceedings envoys from Carteia duly reported that they had Pompeius in their hands.'

It is useful to compare how the sentence might have been constructed according to the 1st century BCE rules of Latin syntax:

3a) = 2) *dum haec geruntur legati Carteienses renuntiauerunt quod Pompeium in potestatem haberent*;
3b) *dum haec geruntur legati Carteienses renuntiauerunt se Pompeium in potestatem habere.*

From a semantic viewpoint, the complement clause in 3b) would have been ambiguous because it could mean both 'reported that they had Pompeius in their hands' and the contrary 'that Pompeius had them in his hands'. Moreover, given the fact that the semantic roles and features of the AcI were underspecified and, therefore, neutralised, the utterer of the complement clause was also presupposed to commit himself to the truth of its content. One of the most significant advantages of the *quod*-construction is that it allowed the utterer not to implicitly commit himself or herself to the truth of the complement clause, but to explicitly take a positive or negative position on its content.

A linguistic change always presupposes a development that can be captured in diachrony as far as the documents that register the process allow it.

7. From the *Accusativus cum Infinitivo* to the complement clause

The transition from AcI to Complementizer + Complement clause with a finite verb started as variation and for a long period of time the two patterns coexisted, but in the end the *quod*-construction ended up replacing the AcI almost completely (Da Milano and Cuzzolin 2019).

In Latin, many of the conjunctions introducing the subordinate clause such as *quod*, *cum*, *ut*, *qua* are old case forms of the relative pronoun, which developed over time the function of a general subordinator. From the point of view of typology, this development is closely related to the type of language Latin belongs to, where the syntactic relations tend to be marked on the word (Lehmann 1989, 7).

As already mentioned above, originally in Latin two complementizers introduced the complement clause: *quod* and *quia*. As I tried to show elsewhere (Cuzzolin 2013a, 2013b), *quod* is the original Italic form that was employed to introduce the complement clause; it was favoured by the antecedent *id* ('that') that formed the so-called correlative diptych. On the other hand, as I suggested, *quia* probably was a calque from a Greek model, witnessed by some dialect, possibly imported and spread into the Latin speaking areas by Greek slaves or people coming from Greece to Southern Italy.

Unfortunately, given the genre of texts delivered to us by the tradition through many centuries we do not possess any witness documenting how the construction with *quod/quia* plus indicative/subjunctive replacing the AcI was perceived by the speakers of Latin. However, a more fine-grained analysis is necessary. Obviously, we cannot employ a generic label such as "speakers of Latin" as if all speakers had the same competence and could use the same lexicon and exploit the same syntactic constructions. It is highly probable that readers of the *Bellum Hispaniense*, from which the example quoted above as (2) is taken, reacted to the *quod*-construction differently from readers of Petronius' *Satyricon*, in which the *quod*-construction occurs four times. This brings us to the second crucial difference between the two patterns, i.e., the one involving the sociolinguistic aspect. This linguistic change spread bottom-up: it started from speakers of low social level and, at a rather slow pace, was accepted in literary texts. The first work in which the *quod*-construction does not seem to occur sporadically is Petronius' *Satyricon*. In this novel the author supplies an extraordinarily vivid picture of life in Campania during the 1st century CE. It represents a melting pot that comprises numerous linguistic varieties and a wide sociolinguistic spectrum of speakers.

It is interesting to notice that in the four occurrences of the explicit complement clause instead of the AcI the subordinator is twice the expected *quod*, but also twice its synonym *quia*, regularly used to introduce a causal subordinate. Here are the examples.

4) Petr. 46, 4
Ego iam tres cardeles occidi, et dixi quia mustella comedit.
'I've had to kill three of his linnets already. I told him that a weasel had gotten them.'

5) Petr. 45, 10
Sed subolfacio quia nobis epulum daturus est Mammaea, binos denarios et meis.
'My nose tells me that Mammaea will set out a spread: two bits apiece for me and mine!'

6) Petr. 131, 7
At illa gaudio exultans: 'Vide, inquit, Chrysis mea, uides quod aliis leporem excitaui?'
'Skipping for joy, "Look, Chrysis, look," she cried out, "see what a hare I've started, for someone else to course!"'

7) Petr. 71, 9
Scis enim quod epulum dedi binos denarius.
'You know that I gave a dinner and two dinars for each guest.'

Comparing the example 2 with the examples 4 to 7 some differences are clear. In the former case, the readers could probably appreciate the attempt of the author to use a syntactic device to distinguish the author's interpretation of the facts. By employing the *quod*-construction the author avoided to convey his implicit commitment to the truth of what had been reported. From a pragmatic viewpoint, the pattern of the finite subordinate turned out to be clear and explicit whereas the AcI frequently was inevitably ambiguous.

In the case of Petronius' examples, the general picture is more complicated and the sociolinguistic aspects are decisive for a thorough analysis of the data. In Petronius' novel the examples quoted are always uttered by lower-class individuals, uneducated, with Greek names. What these documents make indisputably clear is that the pattern with the conjunction plus the finite verb replaced the AcI starting from the linguistic varieties typical of low social classes: it was a successful bottom-up process that gained ground even in literary texts as to become a common pattern in the Romance languages.

Especially in past decades this replacement was accounted for as a Graecism because the Greek culture and its language were the model *par excellence* for the Latin writers and poets: the political and cultural élite of Rome was regularly educated in Athens and was frequently bilingual. However, the claim of the *quod*-construction as a Graecism has been almost abandoned. What should rather be stressed is that probably many people with Greek as their mother tongue favoured the spread of the *quod*-construction. Therefore, if this hypothesis is correct, the *quod*-construction may have gained ground at the expense of the AcI not on structural, but on sociolinguistic ground.

Another factor that favoured the explicit subordination was the spread of the Christian religion, whose main goal was to fight paganism and convert people to the word of Jesus. Preaching was addressed in particular to humble, poor, uneducated people, who should be very similar and close to the characters illustrated in Petronius' novel. It is very likely that those people already used the *quod*-construction in their everyday speech instead of the AcI. The process took centuries, but, in the end, the AcI was almost completely ousted out of the Latin language and was only preserved by late authors that wanted to give stylistic prestige to their prose. However, it is interesting to observe that a pattern reminding the AcI has survived until nowadays in contemporary Italian (Da Milano and Cuzzolin 2019).

8. A case study form the history of Turkish

The case of Turkish. The history of the Turkish language, and particularly the variety that is called Ottoman Turkish, apparently presents a striking parallel to Latin. However, as it should be clear at the end of this illustration, the parallel is superficially similar, but in fact it shows crucial differences.

Let me briefly remind the reader that Ottoman Turkish is subdivided in three periods:

- Old Ottoman: 13th to 15th century;
- Middle Ottoman: 16th to 18th century;
- New Ottoman: 19th century to 1928.

Conventionally, Ottoman Turkish ended in 1928 when Mustafa Kemal Atatürk promulgated the reform of both the language and the alphabet, implicitly creating a new variety of contemporary Turkish.