Influence of Italian orthography on pronunciation of phonemes in Regional Italian

There is evidence today that, depending on different regions, Italians pronounce certain phonemes in different way. For example, the grapheme <z> in zio ‘uncle’ and zitto ‘quiet’ in Northern Italy is always pronounced as [dz], while in the South we hear [ts]. The situation is very similar with the grapheme <s>, which has two different outcomes, [s] and [z], depending on diatopic variable and distributional context (Berruto; Benincà; Sobrero; Telmon; Canepari). These particular types of pronunciation and articulation can be explained by several factors: dialectal influence on Regional Italian, incoherence of the orthographic system for certain phonemes (with some of the phonemic distinctions being graphemically undistinguished) (Maraschio; Cignetti and Demartini), phonemic (non)functionality and problematic phonemes in Italian language (Muljačić). In this paper we will try to examine the significance of each of these factors that have influenced the variation in pronunciation of Regional Italian as well as the usage of the problematic phonemes in Italian language in the future. On one hand, this paper will compare studies on different articulation and perception of Italian phonemes depending on regional variability (Regional Italian), especially for the problematic phonemes in Italian language. On the other hand, we will analyse the possible influence that the incoherent orthographic system has on the loss of distinctive functions in pronouncing those phonemes (today and in the past). In this way we will try to understand what is going to happen in the future with the (non)articulation/perception of certain phonemes of Standard Italian and how this will affect the phonemic system of Italian language.

1. Regional Italian vs. Standard language

Before we start with the discussion regarding the problem of discrepancy between the phonetic reality and the language system we should mention several basic concepts. Standard Italian is just one variety in the language repertory of the Italian community. It is prescribed by grammar manuals and historically it is related to Tuscany or Florence. Its realisation occurs merely in written language, while its oral form is significantly rare. Very few Italians use a variety that is not marked ei-
ther socially, geographically, or situationally (Graffi, Sobrero, Telmon 1997: 162). As the actual standardisation in Italy started with the spreading of literacy among wider population we can conclude that only in the second half of the 20th century adults were capable of transferring the standard language to the new generation. This brought about a new situation. After a certain period, the Neostandard arose. The Neostandard was characterized by the extension of the use of some forms (that in the past had been considered erroneous or colloquial by standard grammars) onto the middle and high classes and onto the highly educated population, e.g. the substitution of *il quale* with *che* or *il cui*, the substitution of *egli*, *ella*, *essi*, *esse* with *lui*, *lei*, *loro*, and many others (Graffi, Sobrero, Telmon 1997: 163). It is very important to bear in mind the actual continuum of one variety to the other. This means that the categorisations are relative and never absolute, and that the flow from the dialect to the standard language is of different grade depending on the structural level of the language, and on individual, situational and social contexts.

We are highlighting this fact because in this paper we are talking about Regional Italian, a diatopic variety that has regional characteristics. The concept of Regional Italian is a rather broad one. On the structural level it stretches from the phonological layer, all the way through morphology and syntax, lexicon, and semantics. It can be very close to dialectal forms or can be closer to the standard language. Considering the reciprocal interferences of Italian language and dialects, it has been confirmed that some structural levels have been only unidirectional (from dialect into Italian language) such as phonetics/pronunciation and intonation or phraseology (where dialects have influenced the standard language and no other way around). At the same time, other levels such as the lexicon and syntax have bidirectional influence. In morphology the dialect is mainly influenced by Italian. From the point of view of sociolinguistics, it can be considered a low variety (Regional Italian), usually with stronger dialectal characteristics or considered less prestigious, or it can be considered as a high variety (Regional Italian), usually with less dialectal characteristics or with regional characteristics that are considered more prestigious. No matter how we define it, even Sobrero (1988: 732) agreed that in order to define the variety it is not enough to say that it is all the varieties of the Italian language diversified in relation to the origin and geographical distribution of the speakers.

These regional characteristics are usually connected to the macrodialical areas of Italian territory (*italiano regionale meridionale*, *meridionale estremo*, *sardo*, *settentrionale*, *romano* etc.) (Graffi, Sobrero, Telmon 1997: 164) which brings us to the actual regional/dialectal influence on the Italian language. As the Neostandard

1 Maraschio (1993: 142) indicates that 62% of Italians (based on the 1981 census) are either semi–literate or of limited literacy. As we already mentioned how the orthographic system can easily reflect the phonemic/phonetic system and the other way around, this situation is not marginal from the sociolinguistic point of view. Regarding the social stratification of different pronunciation of phonemes Canepari (1980: 46–47) states that Italian Standard (based on Florentino type) has not been introduced spontaneously in none of the Italian regions, specifying how even actors have not completely acquired the correct pronunciation.

2 See also Cerruti (2011: 13) that talks about “folk” Regional Italians and “educated” Regional Italians.
includes varieties that are more or less uniformly present in the entire Apennine peninsula, Regional Italian should have specific characteristics only for a specific area, and usually is caused by the dialectal substrate. However, even though the pronunciation outcomes in usage differ from region to region, we do have different characteristics that can be joining the same phenomena (throughout the whole peninsula). In regard to the phonetic/phonemic level in this respect we mention the following:

- inconsistency in pronouncing geminate consonants throughout Italy (North: ‘be:lo’, ‘ro:ba’ vs. South ‘bello’, ‘robba’);
- different articulation of some minimal pairs, for example the intervocalic <-s->, depending on the region, can be pronounced [-s–] or [-z–], as ‘ka:za’ vs. ‘ka:sa’;
- very similar situation with the grapheme <z> with two different outcomes [ts] or [dz], depending on the context and region, as ‘dzio’ vs. ‘tsio’;
- pairs [e] ~ [ɛ] and [o] ~ [ɔ] can be found in many varieties, but outside of Tuscany they rarely present phonological oppositions, while sometimes they do present a different distribution or a vast area of neutralization (Sobrero 1988: 735).  

2. Aim and hypothesis

Many languages that have gone through the period of diffusion of the standard language in the last century have very similar variation outcomes. However, what is notably different (see Slavic languages) is the orthography in Italian language. Even though the features of Regional Italian (especially in pronunciation) were derived from dialect speakers who had learned the language imperfectly – this situation was very similar even to second language acquisition – the second and third generations could have crystallized certain features, but they have not done so.

Therefore, the aim of this study is to understand the main factors that cause different pronunciation of Italian in different regions. We present the possible reasons and we analyse the possible outcome (either in realisation–pronunciation or in the system (Standard Italian language)).

Even though we understand that the initial pronunciation of regional varieties of Italian language was caused by the dialect substrate we primarily assume that a vast amount of today’s variety in the pronunciation of Regional Italian is also caused by the graphemic non–distinction of certain phonemes. In this respect we present graphemes that are problematic (in pronunciation as well) and compare them with non–problematic graphemes (and their outcomes) to see if there is any difference between them. The second hypothesis relates directly to the first one, because the non–distinction in the orthographic system is supported by the non–
functionality of certain phonemes as minimal pairs which reciprocally support each other. In the end, we will present some possible solutions for today’s articulatory and orthographic systems of Regional Italian.

3. Methodology

We have started with the list of Telmon’s (2011: 104–116) phonological/phonetic variations of Regional Italian. The list presents most of the specific studies done across Italy. From 113 features collected by Telmon we agree that a vast majority is in relation with the dialectal substrate, e.g. the vocalic system of Sardegna presents only five vowels (in stressed context with only open–mid vowels). In their regional pronunciation of Italian, we also find only open–mid vowels.

As we stated in the introduction we have found general tendencies that we would call common for Regional Italian that occur in most of the regions, which means that even if the features are fragmented and differ from one another, many features can be joining the same general phenomena (across the whole peninsula).

We have checked the varieties found and compared them with the orthographic discrepancies, as these could be accounted for in the differentiation in pronunciation. We have also compared the above varieties with those phonemes that are considered problematic and vulnerable in the Italian phonological system. Furthermore, we have analysed opposite cases as well (orthographic transparency, but still variety in pronunciation) to check which varieties occur without orthographic influence.

4. Basic concepts

4.1. Phonology of Italian language and problematic phonemes

The phonological system of Italian language consists of 30 phonemes. This number actually varies depending on authors and depending on the chronological evolution of phonology as a science. However, different linguists agree (and we follow the line of Muljačić 1972) that there are certain phonemes that are considered vulnerable, even though they are still considered part of a system. The vocalic system consists of 7 vowels /i, e, ɛ, a, ɔ, o, u/, while there are 21 consonants /p, t, k, b, d, g, ts, dz, tʃ, dʒ, n, l, r, f, v, s, z, ʃ/ and 2 approximants /j, w/. In this phonological inventory eight phonemes are generally considered vulnerable by different authors and usually with different argumentation, but always related to the phonological non–functionality or lack of phonological distinction. These are /ɛ, ɔ, z, dz, tʃ, dʒ, j, w/. Here we can add some more characteristics that separate linguists in how they view geminate consonants, or vocalic quantity, again as being phonologically relevant or not.
4.1.1. Problematic phonemes and other features

Vowels

We can start from the first phonologically distinctive pairs that are at the same time undistinguished in orthography: <o> > /o/ vs. /ɔ/ and <e> > /e/ vs. /ɛ/.[4] In general, the open–mid vowels /ɛ/ and /ɔ/ lack some distributional contexts in which they can appear in the Italian standard language. As we know, they can only appear in the stressed syllable, while in the atonal context only close–mid vowels can occur. This atypical situation makes these two open vowels very vulnerable (especially to linguistic intuition and perception), and by adding the orthographic gap for those two vowels their existence is put to the test.

While we can only speculate about the future, at present we know that there are few minimal pairs that can confirm their functionalities in the phonological system. On the other hand, we witness the lack of difference in pronouncing, for example, /o/ and /ɔ/ in different parts of Italy – what were once maybe functional/dialectal vowels, today they do not serve their cause. The reason for this can be found in the lack of a systematic differentiation in writing, in different pronunciation in Northern and Southern Italy, and in the lack of minimal pairs like /peska/ vs. /pɛskə/.[5] Across the whole peninsula we can find either penta–vowel systems or epta–vowel systems. However, the epta–vowel systems, present in regional pronunciations of Italian, present a different distribution of close–mid and open–mid vowels in contrast with Standard Italian system, as found in Central Italy, Campania, Basilicata, Molise, Veneto and Trentino, e.g. [ˈloro] o [ˈumo]. On the other hand, the penta–vowel systems, as found in Lombardia, Piemonte, Emilia Romagna, Liguria or Puglia, can consist of seven phonetic vowels. However, its timber is exclusively dependent on the phonetic context. The penta–vowel systems are usually collocated in the furthest Southern Italy, e.g. in Sicilia, Calabria, Salento or Sardegna. We can agree with Devoto (1964: 150) who predicted the penta–vowel system of Italian language in the future.

Sibilants /s/ and /z/

Even today there is disagreement among linguists regarding the phonological status of pairs /s/ and /z/ in the Italian phonological system, with some (Porru’, Di Pietro, Saltarelli) considering [z] the allophone of /s/, and others (Muljačić, Fiorelli, Franceschi, Valesio) considering it a pure phoneme. If we look into the criteria that a phone has to meet in order to be considered a phoneme we understand why there is disagreement. The only possible context in which we find minimal pairs is the intervocalic context (which is very vulnerable, presenting usually archaisms

---


5 Muljačić (1972: 45–46) states that linguists usually prefer the standard variety of Rome where [z] does not occur or onto common Italian (considering the situation of majority factor in regional varieties).
or proper names). There are few minimal pairs, and they can only be found in the Florence variety and some central Italian varieties (Muljačić 1972: 46). We can go further and check the distributional context of [s] (never present in the intervocalic context) and [z] (never present after n, l, r; present at the beginning of a word only as a reflex of sonorization due to the voiced occlusive that follow [zdenta:to]). The situation with these two sounds is even more complex when we look into the regional variation in pronunciation – in very large areas of Northern Italy there is always [–z–] and in Southern Italy there is [–s–].

**Affricates /ts/ and /dz/**

The phonemic situation is slightly better with /ts/ and /dz/, even though they also show very few minimal pairs /radzda/ vs. /ratstsa/. However, in distributional context they are equally as strong. With this in mind, we agree that the graphemic transparency is again a problem, with only one grapheme <z> for both phonemes. Therefore, there is a variation of /ts/ – in the Southern Regional Italian it is pronounced [dz], especially after consonants, e.g. <calza> as [kaldza], while at the beginning of the word we have [ts] <zitto> as [tsitto].

**Geminates**

Geminate consonants tend to be simplified in Northern Italy. However, where supported by the graphemic system they tend to be pronounced, so the non-transparent geminates /ts/, /dz/, /ʃ/, /ʎ/, /ɲ/ as <azione> are pronounced as one [ats'jo:ne], whereas in the South, we find the gemination where it should not occur by standard practice, e.g. [abbile], [adʒdʒile].

**4.2. Dialect substrate and education**

As Sobrero (1988: 739) explains, not all phenomena of regional diversification in pronunciation can be explained by using the notion of the dialect substrate, even though the weight of it is significant. We can find other factors and explanations for those varieties, such as inherent dynamics of system, adstratum influences, divulging specific prestigious varieties, usage of “common Italian” etc.

Before the early twentieth century, Italian language was used usually in writing and formal styles and by a minority of the population. Its use has progressively increased during the twentieth century. This increase in use was supported by social factors such as the spread of education, the introduction of compulsory military service (where speakers of different dialects were brought together), the transition from an agrarian society to an industrial society, and the development of mass communication. Italian has then gradually spread into domains that had formerly been

---

reserved for dialects and among dialectal monolingual speakers (Cerruti 2011: 11). This means that dialect speakers were in a situation that can be considered as second language acquisition. By imperfect learning of Italian, dialect speakers have introduced dialectal features into varieties of Italian: “The regional varieties of Italian have therefore derived from this process.” (Cerruti 2011: 11).

In relation to the aforementioned subject of dialect and education, what we find particularly interesting is how certain skills have been immaculately taught at school, while others have been easily ignored. Here we are talking about the intensity and detailed teaching of the orthographic system to the children, while the pronunciation and prosody (usually of regional or dialectal origin, founded in our early life, and part of our deep sense of identity) are generally overlooked, without much criticism or evaluation by teachers. We are pointing this out not because we wish to judge or evaluate the actual fact of doing so, but because, interestingly enough, this could be another factor that has contributed to the lesser distinctiveness of some phonemes/phones mentioned earlier in this paper. In his suggestions for the pronunciation of Italian, Canepari (2006: 15–16) specifically points out that the use of orthography in order to acquire the correct pronunciation of Standard Italian should be the last didactic component used, because it confuses and restricts the more perceptive capacities.

4.3. Italian orthography

Italian orthography depends on the centuries–old tradition of written language. If we look into some examples, we will find many homographic phonemes (/e/ and /ɛ/, /o/ and /ɔ/, /i/ and /j/, /u/ and /w/, /ts/ and /dz/, /s/ and /z/), e.g. ‘ragazzo’ is pronounced with the phoneme [ts], and ‘mezzo’ with [dz]. Furthermore, there are heterographic phonemes (/k/, /ɡ/, /ʃ/, /dʒ/, /ʃ/, /ʎ/, /ɲ/) in which only one phoneme/sound can be written by three different graphemes, e.g. [k]: <casa>, <chiesa>, <quanto>. We mention this problem because it was presented how orthography can influence pronunciation, e.g. visual memory can be accounted for 80% of memory (Canepari 2006: 23–24), and the graphemes that do not correspond to only one sound can significantly influence the regional diversification in pronunciation.

Even though Italian orthography is said to be phonetic/phonological (Marschello 1993: 139), meaning that the discrepancy between phonemic systems and the graphemes is mostly reduced, we can agree with this concept only if we look at other more complicated languages such as English or French, which have a high frequency of graphemes based on etymology, presenting even more distance between levels of speech and writing. However, if we compare Italian orthography with other orthographies, for instance with orthographies in the Slavic languages, we can see that Italian orthography and its rules depend on the centuries–old tradi-
tion of written Italian.\(^8\) This incoherence we can easily demonstrate with the high number of homographic phonemes (/e/ and /ɛ/, /o/ and /ɔ/, /i/ and /j/, /u/ and /w/, /ts/ and /dz/, /s/ and /z/) and then the heterographic phonemes (/k/, /g/, /tʃ/, /dʒ/, /ʃ/, /ʎ/, /ɲ/). We know today that the influence of orthography on the pronunciation of phonemes and the other way around is very much present, and we know this because in contemporary Italian language we can encounter certain pronunciations that have been strictly induced by orthography. For example, we can hear certain pronunciations for the diacritical i that should be mute in *cielo* [tʃe:lo] and *scienza* [ʃentsa], or in the opposite case where it should be pronounced *sciare* [ʃiare] (Maraschio 1993: 141).

5. Results

The majority of variations are caused by the dialect substrate. Some variations are specific to only one or two regions and some are macroareally distributed. Those that are macroareally distributed have a high number of pronunciation variations supported by non-transparent orthography, usually homophones. For example, we find the opposition of mid vowels /o/ and /ɔ/ and /e/ and /ɛ/ only in Tuscany, while in Rome and in Campania there is a particular distribution of these vowels. Other parts of Italy tend either to lower their high–mid vowel (Lombardia, Piemonte, Ticino, Alto Adige, but also Puglia and Sicilia) or close their open–mid vowel (Veneto, Venezia Giulia, parts of Sicilia, Calabria, Salento, Abruzzo). We have a very similar situation with the sounds [s] and [z]. Mainly in Northern Italy speakers pronounce the intervocalic <s> as [z] with prestigious expansion in other areas, while in the South (Abruzzo Molise, Lazio, Roma, Campania, Puglia, Sicilia Umbria we find [s] as in [ˈviːso], [ˈaːnalis]. The functional phonological opposition of s and z we find again only in Tuscany. Phonemes ts and dz have similar outcomes – in initial position, most northern Regional Italians pronounce [dzitto] while Abruzzo, Molise, Campania, Sicilia, Puglia and Umbria have [tʃitto], but at the same time they pronounce *sen*[ts]a > as *sen*[dz]a. With geminates we have found very contradictory results. Other macroareal variations do not have varieties across Italy, or the transparent writing system does not pose a problem.

Those sounds that are microareally distributed are usually not connected to the problematic phonemes or orthography. We mention here only several examples: <cervello> is pronounced as [ʃ]ervello, scervellato > [ʃ]ervelato in Roma o Toscana; in Basilicata [ʃ]atola instead of scatola; also, Roma and Lazio present the rhotacism of l>r in contexts like <coltello> became co[r]tello, or the very well known gorgia Toscana <casa> in [ˈh]asa. These types of microareally distributed variations (and there are many more) are also usually diastratically marked. Even though

---

\(^8\) There is very little discussion in Italy today regarding the existence of a univocal norm and its application of Italian graphemic system, in the way it is done for other levels of Italian language, for example, phonetics and morpho–syntax (Maraschio 1993: 139).
some of them are represented by heterographic symbols they do not pose a phonological problem.

5.1. The opposite cases

The opposite cases show how the graphemic system can be accountable only for some variations. For example, the geminates such as *notte and *botte that are completely transparent in writing are reduced across Northern Italy in note, bote, bela, and at the same time gemination occurs across the South in the intervocalic [b and dʒ] robba and adʒdʒile. On the other hand, it confirms how the phonological function of geminates is redundant. We have also found a vast area where transparent and non-problematic phonemes are voiced, such as [ˈkambo] instead of [ˈkampo] and [ˈaŋge] instead of [ˈaŋke], or [vɪnˈdʒendzo] instead of [vɪnˈtʃentso] as found in Campania, Puglia, Sicilia, Molise, Abruzzo, Basilicata, Umbria and Lazio. Another interesting variation across Southern Italy is the one in which *caldo becomes [callo], or *biondo [bionno].

6. Discussion and conclusion

Throughout centuries, from Latin to present moment, phonetic changes in spoken language have also induced the phonological system to change. This means that some distinctive features became redundant and not relevant, and some others have taken their phonological functions. In Italian language this is the case with some phonemes (the ones mentioned earlier in Regional Italian) and as they lose their functionality the orthographic system pushes the phonological system to be even more vulnerable.

The phonemes that are considered problematic in Italian language, supported by the non-transparent orthographic system, usually homophones, present the highest variation geographically in the pronunciation of Regional Italian.

It may seem easy for linguists (in this rigid world of structures) to accept certain linguistic realities that do not correspond to grammatical rules (Grassi, Sobrero, Telmon 1997: 162–163) e.g. some uses, such as the use of present instead of future tense (*Quest’estate vado al mare) are nowadays more than accepted, while several decades ago this would have been criticized by grammar experts.⁹ We leave to the handbooks to explain the linguistic reality and to make those forms that were once deemed unacceptable in language acceptable.

On the other hand, the orthography (incoherence in graphemes) and the Italian phonological system (non-functionality of certain phones) still remain rigid, even though the articulatory reality of Italian language has been deeply changed by its own speakers. There is no need to predict simplification as an outcome of the language system in general if we accept the non-distinctiveness in phonemes and

⁹ This flexibility is present on all structural levels except phonology and orthography.
unify what has no function in reality, because even though the diversity of some pronunciations is diminished, some distinctive losses are being compensated for by other variation in pronunciation (all microareal pronunciations).

We agree with the idea that a standard language should remain as unified as possible, with few and rare changes, as a solid ground for literacy within one nation. However, we cannot close our eyes to real changes and situations that are happening in the real world with the language system. This means that the pronunciation level of the Italian standard should be taken into account, as well as the possible graphemic changes, especially for the phonemes that should have a valid phonemic distinction in the graphic system as well. As much as it seems complex to divulge what these new changes can bring through education system and mass media, they could bring about long term unification with the standard language that is fresh, transparent, and not a museum relict.

In reality, we must ask ourselves how we could resolve the destiny of dysfunctional phonological pairs when their regional outcomes are so different. How to impose a grapheme or phoneme as a standard form with different variations only as variables or allophones? These types of obstacles are present not only in Italian, but in many other languages as well. If there is an archaism or a neologism in the lexicon, these should be visible in phonology/phonetics and pronunciation as well. In addition to problems in changing the phonological system of Italian language, there is the problem of quantity – what should be the minimum percentage of speakers that should be taken as the criteria for formal changes? The next problem is then how to adapt the changes in writing to the new phonological system. Any changes to the system can cause resistance with linguists, and even more so with the speaker.

To start with, we can at least begin to discuss this matter. Some general propositions would be to solve the etymologic inconsistencies in orthography in the first place. Secondly, all phonological problems could be solved by one variety only, and this could be updated to the most frequently used one. Thirdly, there could be a different system parallel to the standard one, which takes into consideration all regional varieties, and which should be accepted by grammars and institutions as the accepted forms of spoken variety.10

Bibliography:


10 For the latest solutions on pronunciation see Canepari (2009) and (2012).

**Utjecaj talijanske ortografije na izgovor fonema u talijanskom regionalnom jeziku**

U radu se objašnjavaju razlike u izgovoru talijanskog regionalnog jezika (*italiano regionale*), koje su primarno prouzročene dijalektalnim supstratom. Većina dijalektalnih govornika prvih generacija difuzije talijanskog standarda usvojila ga je kao drugi jezik. Stoga su u izgovoru bile jače prisutne njegove regionalne razlike. Danas je to ipak drugačije, no varijacije u izgovoru i dalje prirodno postoje. Uzimajući velik broj (113) fonetičkih karakteristika regionalnoga talijanskog (iz Telmona), došli smo do zaključka da su neke od tih varijacija sekundarno prouzročene i netransparentnom grafijom i vrlo lošim fonološkim statusom ponekih fonema koji stvaraju veće, odnosno makroarealne razlike na cijelom talijanskom teritoriju. Kako bi se standardni jezik ponajprije osuvremenio te donekle unificirao u izgovoru, predlažemo rješavanje nekih postojećih problema: (1) osvremeniti i ujednačiti nekoherentni grafički sustav, često temeljen na etimologiji; (2) riješiti fonološke probleme (nefunkcionalne parove prilagoditi prema najnovijem modelu (većinskom) izgovora); (3) predložiti paralelni sustav koji uzima u obzir regionalne varijetete koji bi bili prihvaćeni u gramatikama kao ispravni oblici usmenog jezika.

**Keywords**: Italian, dialectology, pronunciation, phonemes, orthographic system

**Ključne riječi**: talijanski jezik, dijalektologija, izgovor, fonemi, grafički sustav