# CONSTRAINT INTERACTION AND WRITING SYSTEMS TYPOLOGY

Antonio Baroni

# CONSTRAINT INTERACTION AND WRITING SYSTEMS TYPOLOGY

**Antonio Baroni**
(antonio.baroni@studenti.unipd.it)
Università degli Studi di Padova

RÉSUMÉ
L'objectif de cette contribution est d'analyser les différences entre les systèmes d'écritures du point de vue de la Théorie de l'Optimalité (TO), un modèle linguistique fondé sur l'interaction d'un ensemble de contraintes susceptibles d'être enfreintes, qui se basent, généralement, sur des facteurs perceptuels, articulatoires ou cognitifs. En particulier, cet article propose d'analyser les phénomènes suivants à l'aide de la TO : la représentation de la longueur vocalique et consonantique, l'effet Buben, le Dialecte Oculaire ou Eye Dialect. On démontre que cette approche est capable de rendre compte des différences entre les systèmes d'écriture, tout en soulignant leurs similarités.

MOTS-CLÉS: orthographe – système d'écriture – graphématique – Théorie de l'Optimalité – linguistique – phonologie – phonétique – scripta – allemand – anglais – italien – français – pendjabī – finnois – alphabet – syllabaire – abjad – abugida – chinois – coréen – longueur vocalique – longueur consonantique – effet Buben – Eye Dialect.

ABSTRACT
The aim of this contribution is to analyze differences between writing systems from the perspective of Optimality Theory, a linguistic framework based on the interaction of a set of violable constraints, generally grounded in perception, production or cognition. In particular, this article proposes optimal theoretic analyses of the following phenomena: vowel and consonant length representation, Spelling Pronunciation, Eye Dialect. It is shown that the current approach is able to account for differences across writing systems, yet highlighting their similarities as well.

KEYWORDS: orthography – writing systems – graphemics – Optimality Theory – linguistics – phonology – phonetics – script – German – English – Italian – French – Punjabi – Finnish – alphabet – syllabary – abjad – abugida – Chinese – Korean – vowel length – consonant length – Spelling Pronunciation – Eye Dialect.

Constraint interaction and writing systems typology

## 1. Introduction

Most linguists, when dealing with graphemics, written language, writing systems and orthography, feel the need to justify themselves. It is about time to change this attitude and to stop feeling guilty about treating graphemics as part of linguistics. The main reason why linguists claim that writing is not language and should not be considered by linguistics is its derivative nature. Such a position is shared by influential scholars such as Saussure (1916), Bloomfield (1933), Jakobson (1963), and implicitly, by most generative literature. For example, Pinker & Bloom (1990, p. 707) claim that “language is not like writing or the wheel”. At the same time, those who tried to integrate graphemics into linguistic theory, especially in the framework of structuralism, ended up with analyses that were nothing but a list of grapheme–phoneme and phoneme–grapheme correspondences, where the graphemes (or graphonemes, cf. Hořejši 1971, p. 186) consisted merely in the letter or group of letters transcribing a phoneme. These first attempts had the merit, nevertheless, to recognize a certain degree of independence of written language from speech (Artymovyć 1932, Vachek 1939). One of the most interesting arguments in favor of the study of writing is proposed by the Danish scholar Louis Hjelmslev (1943), who claims that language is form and is not limited to only one substance. Sign language or writing would then still be language, since the form is not affected by the substance (air, ink, etc.). However, as one of his students would point out later, even if one tries to ignore the physical reality of language, “on tient compte de la substance à toute étape de l’analyse” (Fischer-Jørgensen 1949). As a matter of fact, all phonological theories, even those that claim to be completely phonetics-free (e.g. CVCV theory, cf. Scheer 2004) have to deal with physical reality at some point of their analysis (at least in the labels they assign to distinctive features or elements (e.g., [+rounded], [-high], palatality, etc.). The fact that these labels are not applicable at all at sign language makes very clear that the medium through which language is expressed affects language itself or at least its analysis.

The main aim of this paper is not to deny the derivative nature of written language, its non-universality, or its ontogenetic or phylogenetic secondarity in relation to speech. Quite simply, I argue that once speakers live in an environment where they learn and experience reading and writing, all the arguments mentioned above become irrelevant. For literate people, speech and writing are just two different expressions of the same faculty. I intend to show that it is possible to integrate written language in linguistic theory, in particular, through the framework of Optimality Theory.

## 2. Contradictory Principles

Writing systems are normally described as a combination of different principles and these principles are often in conflict with one another (Coulmas 2009, Baroni 2011). Chinese writing, for example, is at the same time morphographic – its graphic units stand for morphemes – and syllabographic – its graphic units stand for syllables. Generally, Chinese characters are composed of an element that conveys the meaning – the radical – and of an element that gives a vague indication about pronunciation. Contrary to what most people might think, Chinese is a not an example of a semasiographic system, completely detached from speech. It is still dependent from spoken language, but works at a different level than alphabetic systems. On the other hand, all alphabetic orthographies display some morphological and semantic components, although not in a consistent way. In English some morphemes maintain their spelling unaltered in derived words even though pronunciation differs, e.g., <electric, electricity> [ɪlɛktɹɪk, iːlektɹɪsɪti], <sign, signature> [saɪn, sɪɡnətʃə]. In <electric>, the first <e> stands for [ɪ], the second <e> stands for [ɛ] and the second <c> stands for [k], but in <electrical>, the first <e> now corresponds to [i], the second <e> to [e] and the second <c> for [s]. Similarly, in <sign>, <i> corresponds to a diphthong and <g> is silent, whereas in <signature>, <i> stands for a monophtong and <g> is pronounced as [ɡ]. One might argue that in the case of <electric, electricity> an English native speaker might know by rule how to pronounce the two words (e.g., something like velar softening in Latinate words, in the spirit of Chomsky & Halle 1968), so he or she would not need orthography to be overly precise. In the case of <sign, signature>, a phonetic final sequence as [-gn] would violate English phonotactics, so it is likely that native speakers would not pronounce <gn> as such. However, which rule tells us that <g> must be silent and not <n>? The only suitable way to analyze cases like <electric, electricity> and <sign, signature> in synchrony is to invoke what I propose to call *Morpheme Identity Conservation Principle*: the graphic shape of a morpheme is conserved regardless of phonology. Another orthography that applies the *Morpheme Identity Conservation Principle* quite often is German orthography, e.g., the plural of <Haus> 'house' is <Häuser>; it might as well have been spelled *<Heuser> and it would have still been pronounced [hɔʏzɐ], but preserving the <a> of <Haus> indicates a morphological relationship between the two forms. Nevertheless, neither English nor German are consistent in applying the principle: <see> and <sight> and <collide> and <collision> are obviously morphologically related, but their spelling differs. In German, the plural of <Alt> 'old', is <Eltern> and not *<Ältern>. Another principle that seems active in many alphabetic orthographies is *Homophones Distinction*: words that sound the same but have different meanings are spelled differently. In English we have up to four different graphic words corresponding to the same pronunciation, i.e., <right> vs. <rite> vs. <write> vs. <Wright> for [ɹaɪt]. In French we have six: <sein> 'bosom', <saint> 'saint', <ceint> 'girded', <sain> 'sane, healthy', <seing> 'signature', and one could add <cinq> 'five' when the final <q> is silent (normally before another consonant); they all sound [sɛ̃]. In Danish [vɛˀɐ] can be spelled <vejr> 'weather, <vær> 'to maintain double', <hver> 'each', <værd> 'worth'. Even a shallow orthography as Italian distinguishes homophones, mostly employing diacritics, as in <da> 'from' vs. <dà> 'he/she/it gives', <di> 'of' vs. <dì> 'day', <la> 'the', <là> 'there', or with silent <h>, as in <anno> 'year' vs. <hanno> 'they have', <a> 'at, to' vs. <ha> 'he/she/it has', <o> 'or' vs. <ho> 'I have', <ai> 'to the' vs. <hai> 'you have'. In all these orthographies, though, there are examples of homophones that are also homographs (e.g., Danish <dør> [dœˀɐ] 'door' and 'die/dies') and even homographs that are not homophones (e.g. English [liːd] vs. [lɛd], both spelled <lead>; French *se* [fje] 'to trust' vs. [fjɛʁ] 'proud', both spelled <fier>; Italian [ˈaŋkora] 'anchor' vs. [aŋˈkora] 'again, still', both spelled <ancora>).

Leaving aside the relationship with pronunciation, it is interesting to note that different writing systems and orthographies also differ in the importance they give to visual salience. Chinese writing is undoubtedly very complex and quite cumbersome, but the shape of each character is very distinct from all the others and on the printed page they stand out much more than, say, Latin letters. Yet, the letters of the Latin alphabet possess lines that extend in upper and lower space, as in <b, d, f, g, j, k, l, p, q>, a characteristics that seems to enhance readability (Sampson 1985, p. 94; Pontecorvo 1994, p. 278). On the contrary, Hebrew letters normally consist of a horizontal line on top and a vertical line on the right. There appears to be experimental evidence that Hebrew and Arabic readers make longer eye fixations than European readers (Gray 1956, p. 59). Even within orthographies employing Latin alphabet there are notable differences with regards to visual salience. For example, English avoidance of final <i, u> is due to the fact that, during the Middle Ages, scribes did not use spaces between words so the end of a word had to be signaled by other means. <i, u> did not have very salient shapes so they were systematically replaced by <y, w> or <ie, ue> word-finally. Today the norm is to divide two words with a space and computer writing made letters much more distinct from each other. However, the rule still remains in English spelling, although final <i, u> appear to be now tolerated (e.g., in the non-standard spelling <thru> for <through> or in given names such as <Nikki>)[1].

[1] Social, cultural, political and even religious factors play a major role in orthographic issues (Smalley 1964, Sebba 2006). However, for space reasons, this paper will only focus on the linguistic aspects of writing.

## 3. Optimality Theory and Grounded Constraints

Given that writing systems and orthographies appear to be better described resorting to multiple coexistent conflicting principles rather than to inviolable rules, Optimality Theory (OT henceforth) seems to be suitable for the analysis of graphemic phenomena. In this section I will present the framework, discuss some issues that may rise in applying OT to graphemics and propose some ways to solve them. OT (Prince & Smolensky 1993) was originally conceived to account for phonological facts but has been applied more recently to other fields as well, such as morphology (Legendre 2000), syntax (Grimshaw 1997; Legendre, Grimshaw & Vilkner 2001), semantics (de Swart 2006), and graphemics (Primus 2004; Wiese 2004; Song & Wiese 2010, Baroni 2013). As a matter of fact, nothing in the architecture of OT excludes the possibility to deal with fields other than phonology. OT assumes the existence of quite a large set of universal constraints that are valid for language in general. What differentiates language X from language Y is the ranking of these constraints. Constraints are universal but their hierarchy is language-specific. Most constraints are based on generalizations about the most frequent phenomena in the world's languages and/or are grounded on perceptual, articulatory and psychological facts. Constraints are of two types: Faithfulness and Markedness. Faithfulness constraints preserve the input from undergoing changes and ensure lexical conservation. Markedness constraints avoid excessive effort in production and lack of salience in perception.

(1) Examples of proposed constraints in OT:

**Faithfulness**

Max — All the segments in the input must be preserved in the output (No deletion)

Dep — All the segments in the output must be present in the input (No insertion)

**Markedness**

Onset — Syllables have onsets.

NoCoda — Syllables do not have codas.

OT is articulated in GEN, the component that virtually gives every possible output, EVAL, the component which evaluates the best output among a set of candidates, and these candidates are evaluated following CON, the constraints set. This process of selection is typically shown by a tableau, where the winning candidate is indicated by a pointing hand. Given the constraints C1, C2, C3 and the following ranking, C1 > C2 > C3 (C1 is ranked higher than C2 that is ranked higher than C3), the result of the evaluation is the following:

Tableau 1: constraint interaction in OT

| **Input** | C1 | C2 | C3 |
|---|---|---|---|
| ☞Candidate (a) | | | ** |
| Candidate (b) | | *! | * |
| Candidate (c) | *! | | |

Candidate (a) is the winner because it does not violate C1 and C2, which are ranked higher than C3, even if it violates C3 twice. Candidate (c) is ruled out because it violates the highest ranked constraint (the exclamation point indicates a fatal violation) and candidate (b) is ruled out simply through comparison with candidate (a). Note that the cells of the tableau that turn out to be irrelevant are shaded.

An issue that immediately arises is the necessity to limit the overgeneration of possible constraints. Phonologists have come to the agreement that constraints must be grounded, either in phonetics or in cognition. At worst, seemingly unnatural constraints are assumed to exist given speakers' capacity to induce them from positive or negative evidence in the language (Baroni 2001). Grounding is quite straightforward in phonology, since it is relatively easy to find articulatory or acoustic counterparts of phonemes and phonological features. Therefore, a constraint like NoCoda is grounded in the fact that consonant place and manner cues are richer in pre-vocalic position, rather than post-vocalically. Similarly, NoVoicedObstruents is grounded in the fact that it is articulatorily costly to maintain voicing in obstruents, whereas voicing is spontaneous for sonorants. Constraints of the Align family, proposed, for example, by McCarthy & Prince (1993) for Tagalog, militate for a specific morpheme to appear at the left or at the right edge of a word (i.e, word-initially or word-finally). The beginning and the end of a word are normally more cognitively salient for the listener and it is therefore preferable to place relevant information there, rather than word-medially. The question is: if phonological constraints can be grounded in phonetics and morphological constraints in cognition, what about graphemic constraints? Firstly, I argue that most of them can be justified on cognitive grounds as well. For example, it is likely that one-to-one relationships, where one function corresponds to one form, are more easily computable than more complex relationships (cf. Dressler et al. 1987). Therefore, a relationship of the type <a> ↔ /a/ might be preferable to /s/ → <c, s, ss, sc, x, ç> (as in French orthography). Secondly, I also argue that phonetic properties of specific segments can have an impact on writing. From a survey conducted by Justeson (1976), it appears that some features and segments are more likely to have a corresponding grapheme in a given writing system compared to others, somehow more marked. For example, if in a writing system there is a straightforward way to note a long vowel (e.g., a letter standing for /aː/), there must also be a way to note the short version of that vowel[2]. Similarly, if voiced obstruents are noted, so are voiceless ones. Inherently weak phonemes or phonemes in weak positions are often omitted, such as nasals in pre-consonantal position, consonants in complex clusters, glides or /h/. It is also important to note that generally consonants carry more lexical information than vowels. As a matter of fact, there are writing systems, such as the Semitic abjads, that normally note only consonants, but there are no writing systems only noting vowels. That also depends on the structure of Semitic languages, but in other languages as well consonants seem to play a greater role than vowels for word recognition. Written abbreviations normally maintain only consonants, as in English <mmt> for *moment*, Italian <cmq> for *comunque* ‘however’, <tmb> for Spanish *también* ‘also’, and in casual and/or fast speech, vowels are among the first segments to be reduced or deleted, generating very long consonant clusters even in languages that normally do not allow them, e.g., European Portuguese *desprezar* ‘to despise’, careful speech [dɨʃpɾɨzaɾ], casual speech [dʃpɾzaɾ] (Heinz 2012), Italian *sempre* 'always', careful speech [sɛmpre], casual speech [smpɸ][3]. Consonant representation may therefore have a sort of primacy over vowel representation in writing.

Constraints on the graphic representation of speech may be grounded either in cognition (e.g., preference for simple relationships) or in phonetics (e.g., preference to represent - “make visible” - relatively more salient sounds). Song & Wiese (2010) propose two constraint families that they call Simplicity and Visibility. Simplicity reflects the preference for unmarked structures, whereas Visibility militates for certain segments or features to be represented. More specifically, Simplicity bans graphemes whose visual shape is complex (i.e. formed by a great number of strokes), favors continuous over discontinuous shapes and privileges straightforward speech-to-writing and writing-to-speech relationships (e.g., one phoneme – one grapheme, one morpheme – one grapheme, etc.). A type of Visibility constraint is Visibility-V-Length, which is active in languages like German and Finnish where vocalic length is distinctive and needs to be signaled in writing, e.g. Finnish /muːta/ → <muuta> 'more', German /zeː/ → <See> 'lake'. I propose to consider another constraint family, opposite to Visibility, that militate against the representation of weak phonemes and features, *Visibility. For example, in Semitic abjads a constraint like *Visibil-V (do not represent vowels) must be ranked higher than a Visibility constraint militating

[2]In Arabic and Hebrew, if vowels are noted at all, normally only long vowels have a corresponding sign, whereas short vowels are omitted. Justeson's predictions are not entirely wrong, though, since the signs that transcribe long vowels are not specific for that use, but are otherwise employed as consonant signs.

[3]Extracted from the spoken Italian corpus CLIPS, available at http://www.clips.unina.it/it/.

for the representation of all phonemes. In an optimal theoretic view, the differences between writing systems can be accounted for by VISIBILITY constraints re-ranking:

(2)

**Alphabet**: each grapheme (minimal graphic unit) represents a phoneme (Latin, Cyrillic).
VISIBIL-PHONEME > VISIBIL-SYLLABLE > VISIBIL-MORPHEME.
**Abjad**: each grapheme represents a consonant (Arabic, Hebrew), vowel representation is facultative and defective.
VISIBIL-CONSONANT, *VISIBIL-VOWEL > VISIBIL-PHONEME > VISIBIL-SYLLABLE, VISIBIL-MORPHEME.
**Abugida**: each grapheme represents a syllable but the graphic elements standing for the consonant and for the vowel are recognizable (Devanāgarī, Ge’ez).
VISIBIL-SYLLABLE, VISIBIL-CONSONANT, VISIBIL-PHONEME > VISIBIL-MORPHEME.
**SYLLABARY**: each grapheme represents a syllable (Cherokee, Japanese kana).
VISIBIL-SYLLABLE > VISIBIL-PHONEME > VISIBIL-MORPHEME.
**Morphosyllabary**: each grapheme corresponds to a syllable and/or a morpheme (Chinese).
VISIBIL-MORPHEME, VISIBIL-SYLLABLE > VISIBIL-PHONEME.
**Korean Hangŭl** is unique in the sense that it is an alphabet where the shape of consonant graphemes is articulatorily iconic and graphemes are grouped according to an abstract syllabic division.
VISIBIL-PHONEME, VISIBIL-CONSONANT-FEATURE, VISIBIL-SYLLABLE > VISIBILMORPHEME.

With regards to alphabetic orthographies, on the other hand, we can expect them to rank VISIBILITY constraints similarly, their diversity probably consists in the re-ranking of SIMPLICITY constraints. In shallow orthographies (e.g., Italian, Spanish, Finnish) SIMPLICITY constraints such as *COMPLEX (one-to-many and many-to-one mappings between phonemes and graphemes are banned) must be ranked higher than in opaque orthographies (e.g., English, French, Danish). In the next section I will show some examples of constraint interaction in alphabetic orthographies, both opaque and shallow ones. More specifically, I will deal with the representation of vowel and consonant length and the two phenomena generally known as Spelling Pronunciation (or *effet Buben* in French) and Eye Dialect.

## 4. VOWEL LENGTH

Vowel length is distinctive in a number of languages and it appears cross-linguistically more often than consonant length. Out of a sample of 60[4] languages, in 20 both vowel and consonant length are distinctive, in 29 only vowel length and in 11 only consonant length. If a language possessing a written form has distinctive vowel length, there are six possibile ways to represent it:

(3)

1. <$V^1V^1$>: if a vowel /V/ is represented by a grapheme <V>, the long vowel /V:/ is represented by the reduplication of the grapheme <V>, e.g., Dutch <beet> /be:t/ 'bite',

[4]The languages in the sample are the following, grouped per family: Indo-European (Italic: Latin; Hellenic: Ancient Greek; Slavic: Czech, Slovak, Polish, Russian, Ukrainian; Germanic: Dutch, English, German, Old Norse, Danish, Icelandic, Swedish; Romance: French, Italian, Lombard, Catalan; Celtic: Irish, Scottish Gaelic, Welsh; Baltic: Latvian, Lithuanian; Indo-Iranian: Hindi, Gujarati, Punjabi, Sanskrit, Oriya, Bengali; Anatolian: Hittite), Finno-Ugric (Finnish, Hungarian, Estonian), Semitic (Hebrew, Arabic), Aymaran (Aymara), Sino-Tibetan (Cantonese, Burmese, Limbu), Iroquoian (Cherokee), Austronesian (Hawaiian, Fijian, Maori, Pattani Malay, Samoan, Yapese), Eskimo-Aleut (Inuktikut), Dravidian (Kannada), Altaic (Korean, Japanese, Turkish), Uto-Aztecan (Luiseño), Algic (Mi'kmaq), Austroasiatic (Khmer, Vietnamese), Tai-Kadai (Lao), Niger-Congo (Ganda), Oto-Manguean (Trique), Wagiman.

Estonian <saada> /sa:da/ 'to get.[5]

2. <$V^1V^2$>: if a vowel /V/ is represented by a grapheme <V>, /V:/ is represented by <V> plus another vowel grapheme different from <V>, e.g., German <Liebe> /li:bə/ 'love', English <beat> /bi:t/.
3. <VC>: a grapheme <V> represents a long or a short vowel depending on whether it is followed or not by a <C>; or alternatively, whether it is followed by a single or a double <C>, e.g., German <zehn> /tse:n/ 'ten', Swedish <byta> /by:ta/ 'to switch' vs. <bytta> /bʏta/ 'bucket'[6].
4. Diacritics: a diacritic sign (macron, acute/grave/circumflex accent, etc.) is placed over <V> to represent /V:/, e.g., Czech <můj> /mu:j/ 'my, mine', Scottish Gaelic <bàta> /pa:ʰtə/ 'walking stick'.
5. <V:>: given /V/ and /V:/, they are represented respectively by <V> and <V:>, e.g., Hindi <अ> /a/ vs. <आ> /a:/ or <न> /na/ vs. <ना> /na:/, Ancient Greek <ε> /e/ vs. <η> /ε:/.
6. Vowel length is not represented graphically, e.g., in Korean Hangŭl (cf. Weingarten 2012).

German employs three different methods to represent vowel length, (1), as in <See> ‘lake’, (2) as in <Liebe> ‘love’ and (3) as in <Ihrer> ‘yours’. Song & Wiese (2010, p. 92) point out that only certain vowel graphemes can be doubled in German, namely <a, e, o>, whereas <ä, ö, i, u, ü> cannot. Sequences like <ii>, <üü> or <uu> are not acceptable in written German. According to them, the reason lies in the complexity of <ä, ö, i, ü>, whose shape is discontinuous (using their terminology, they bear the feature [disconnected]). They posit a constraint banning doubling of letters bearing this feature. Similarly, the impossibility to double <u> appears, in their analysis, to be caused by a feature, [open-up]. I propose that it is not necessary to consider <u> separately from the other letters that cannot be doubled. The doubling of <ä, ö, i, u, ü> engenders the adjacency of identical strokes: the umlaut in <ää, öö>, the dot and the minim in <ii>, the minims in <uu> and the umlaut and the minims in <üü>. The shape of <a, e, o> is such that these letters, if doubled, maintain the alternation of different strokes. Instead of positing two different constraints, I unify them in one: *SEQIDSTROKES, militating against sequences of identical strokes. This constraint is not hard to justify, since it appears to be grounded both in production (avoidance of identical gestures when writing) and perception (lack of visual distinctiveness when reading). *SEQIDSTROKES might be considered the written counterpart of the phonological constraint OCP[FEATURE], that bans adjacent segment sharing the same feature. In German *SEQIDSTROKES must be ranked below VISIBIL-V-LENGTH but higher than DEP-L, a constraint banning the use of letters that do not correspond to any specific phoneme.

(4) Ranking for German: VISIBIL-V-LENGTH > *SEQIDSTROKES > DEP-L.

| | |
|---|---|
| VISIBIL-V-LENGTH | Represent vowel length. |
| *SEQIDSTROKES | Avoid sequences of two adjacent identical strokes |
| DEP-L | Every letter in the output must have a corresponding phoneme in the input (no silent letters). |

Tableau 2: representation of /i:/ in German.

| /i:/ | VISIBIL-V-LENGTH | *SEQIDSTROKES | DEP-L |
|---|---|---|---|
| a) <ii> | | *! | |
| ☞b) <ie> | | | * |
| ☞c) <ih> | | | * |
| d) <i> | *! | | |

[5] Estonian vowels display three degrees of length, with short, long and extra-long vowels. However, long and extra-long vowels are not distinguished in writing.

[6] In Swedish, as in other Germanic languages, the difference in length normally involves a lax vs. tense opposition.

Tableau 2 shows how constraint interaction selects <ie> and <ih> as possible representations of /iː/ in German. The input is the phonological form /iː/ and I consider here four possible candidates for its graphic representations. Candidate (a) represents vowel length by letter doubling but violates *SeqIdStrokes. Candidate (d) is ruled out because it does not represent length by any means. (b) and (c) are both winning candidates because they only violate Dep-L, which is the lowest constraint in the hierarchy, but they do represent length and they do not display a sequence of identical strokes.

In a language like Finnish, whose orthography is very transparent, constraints militating for a 1:1 relationship between phonemes and graphemes must be ranked higher than in German. Dep-L dominates *SeqIdStrokes, and, as a matter of fact, in Finnish vowel length is consistently signaled by letter doubling, regardless of letter shape, e.g. Finnish <liina> 'cloth'[7].

Tableau 3: representation of /iː/ in Finnish

| /iː/ | Visibil-V-Length | Dep-L | *SeqIdStrokes |
|---|---|---|---|
| ☞a) <ii> | | | * |
| b) <ie> | | *! | |
| c) <ih> | | *! | |
| d) <i> | *! | | |

5. Consonant Length

Just like vowel length, consonant length can be noted following different methods:

(5)
1. <CC>: if there is a consonant grapheme <C> standing for /C/, then /Cː/ will be rendered by <CC>, e.g., Polish <lekki> /lɛkːi/ 'light, not heavy', Ganda <ŋŋenda> /ŋːeːnda/ 'I go'.
2. Diacritics: a diacritic sign is placed upon/before/after <C> to render /Cː/, e.g., Punjabi <ਦਸ> [d̪əs] 'ten' vs. ਦੱਸ [d̪əsː] 'to tell'.
3. <Cː>:
   1. a specific grapheme is employed for <C> and another one for <Cː>, e.g., in Wagiman <d> stands for /t/ and <t> stands for /tː/.
   2. a specific grapheme is created for <Cː> doubling <C> and then merging the two signs in a conjunct, e.g., Brahmanic scripts, such as Devanāgarī.
4. <VC><CV>: a VC¹ grapheme is followed by a C¹V grapheme, e.g. Hittite syllabary <as> + <su> + <us> + <sa> + <an> + <ni> for /asːusːanːi/ 'master horse trainer'.
5. Gemination is not noted by orthography (Amharic-Ethiopic abugida, cf. Weingarten 2011).

Most languages with phonemic consonant length use (1), i.e., letter doubling, whereas diacritics, which are quite common, cross-linguistically, to note vowel length, are seldom employed on consonant graphemes. Italian is typical in this sense, e.g., <coro> 'choir' vs. <corro> 'I run'. Punjabi, which employs a Brahmanic script, Gurmukhi, prefers to place a diacritic sign, called *áddak*, before the sign standing for the syllable containing the geminate consonant. In order to compare Italian and Punjabi, I propose to consider the following constraints:

(6)
Visibil-C-length — Represent consonant length.

[7] Note that the model for Finnish orthography, at its earlier stage, was based on Latin, High German and Swedish and the relationship with speech was therefore much less transparent.

| | |
|---|---|
| NODIACRITICS | Letters do not bear diacritics. |
| *LL | Avoid sequences of identical letters. |

VISIBIL-C-LENGTH militates for phonemic consonant length to be represented. NODIACRITICS may appear hard to consider a universal constraint, but I argue that it is grounded both in ease of production and ease of perception. When an orthography employs diacritic signs, they are very often abandoned in spontaneous writing (both handwriting and computer writing) because they hinder fluency (Coulmas 1989, p. 237). During the reading process, diacritics require a certain effort to be identified since they are normally not very salient. *LL is somehow similar to *SEQIDSTROKES, meaning that normally language does not like the repetition of identical elements. Moreover, a study conducted by Pontecorvo (1993) on Italian children proves that pre-literate or semi-literate children do not perceive letter doubling as linguistically significant. As a matter of fact, letter doubling to note geminates is one of the major sources of spelling mistakes in Italian. Interestingly, Finnish students seem to have similar issues (Lehtonen 2006).

Tableau 4: Italian *corro* ‘I run’

| /korːo/ | VISIBIL-C-LENGTH | NODIACRITICS | *LL |
|---|---|---|---|
| ☞a) <corro> | | | * |
| b) <coro> | *! | | |
| c) <coŕo>[8] | | *! | |

In Italian, the constraint against diacritics is ranked higher than the one banning the repetition of the same letter, so that <coŕo> is not a possible graphic representation of /korːo/. Since length must be noted somehow, (a) is the winning candidate, although it violates *LL, and (b) is ruled out because it violates VISIBIL-C-LENGTH.

Tableau 5: Punjabi <ਪੱਤਾ> /pətːɑ/ ‘leaf’

| /pətːɑ/ | VISIBIL-C-LENGTH | *LL | NODIACRITICS |
|---|---|---|---|
| ☞a) <ਪੱਤਾ> | | | * |
| b) <ਪਤਾ> | *! | | |
| c) <ਪਤਾਤਾ> | | *! | |

In Punjabi, unlike in Italian, *LL dominates NODIACRITICS, so (a) is the winning candidate, since letter doubling is banned. However, it would not be correct to claim that in Italian NODIACRITICS always dominates *LL. Vowel graphemes generally cannot be doubled but they can bear a graphic accent, e.g., to indicate that the vowel is stressed or to distinguish homophones, whereas consonant graphemes can be doubled but they cannot combine with any diacritic sign. This situation is quite common in many orthographies. I therefore argue that in languages like Italian, *LL has to be split in *<VV> and *<CC> and NODIACRITICS in NODIACRITICSC and NODIACRITICSV. *<VV> and NODIACRITICSC dominate *<CC> and NODIACRITICSV, as shown by the example in Tableau 6.

Tableau 6: Italian *scappò* ‘he/she/it fled’.

| /skaˈpːɔ/ | VISIBIL -C-LENGTH | VISIBIL-STRESSEDV | *<VV> | NO DIACRITICS C | *<CC> | NO DIACRITICSV |
|---|---|---|---|---|---|---|
| ☞a) <scappò> | | | | | * | * |

[8] Note that candidates (c) of Tableau 4 and candidates (b), (c) and (d) of Tableau (6) are unattested in the history of Italian orthography and highly unlikely to emerge. However, they are employed exactly to show why they are not possible forms (i.e., because of the constraint ranking). One of the underlying ideas of OT is the Richness of the Base hypothesis, i.e., the assumption that GEN generates any possible realizations of a given input without restriction.

| | | | | | | |
|---|---|---|---|---|---|---|
| b) <scaṕoo> | | | *! | * | | |
| c) <scappoo> | | | *! | | * | |
| d) <scaṕò> | | | | *! | | * |

In Tableau 6 not all the constraints are crucially ranked with each other. Tied constraints are indicated by a dotted line. VISIBIL-C-LENGTH ties with VISIBIL-STRESSEDV, a constraint militating for the stressed vowel to be indicated, they are both undominated and no candidate violates them. The two constraints dominate *<VV> and NODIACRITICSC, that ban candidates (b), (c) and (d) and dominate *<CC> and NODIACRITICSV.
Generally, in the world's writing systems, diacritics are much more likely to be placed upon vowel graphemes rather than consonant graphemes. This could be due to the fact that normally a long consonant is ambisyllabic whereas a long and/or stressed vowel belongs to one syllable. Letter doubling might render ambisyllabicity more iconically than a diacritic sign.

## 6. SPELLING PRONUNCIATION & EYE DIALECT

This section will deal with Spelling Pronunciation (Levitt 1978) and Eye Dialect (Bowdre 1964). The former consists in the modification of the pronunciation of certain words based on how they are spelled, whilst the latter leaves pronunciation unaffected but changes the spelling, normally to give the impression of non-standard speech or to result "cool" and rebellious. If the former can be described as a form of hypercorrection, normally considered acceptable and likely to eventually become part of the standard, the latter is a purposeful deviation from the norm. What the two phenomena have in common is that they bring the phonological and the orthographic forms closer. Both Spelling Pronunciation and Eye Dialect tend to occur more often in languages with opaque orthographies. In French, many final letters that used to be silent are now (optionally or compulsorily) pronounced, as in *but* [by ~ byt] 'aim', *cinq* [sɛ̃ ~ sɛ̃k] 'five', *août* [ut] (formerly [u]) 'August', *sens* [sɑ̃s] (formerly [sɑ̃]) 'sense, meaning'. In English, several words of classical origin now contain a [θ] that was never there, as in *author* [ɔːθə/ɚ], from Latin *auctor* [auktor] (Neuman 2009, p. 400). These are all examples of Spelling Pronunciation. Eye Dialect forms in English are, among others, <tonite>, <sed>, <thru>, <tho>, <woz> for *tonight, said, through, though, was.* French examples are <koi>, <jamè>, <z'yeux> for *quoi* 'what', *jamais* 'never', *(le)s yeux* 'the eyes'. I argue that both Spelling Pronunciation and Eye Dialect can be described through constraint re-ranking, in which complex relationships between phonemes and graphemes are penalized. Importantly, even if pronunciation is aligned to spelling, as in the case of Spelling Pronunciation, phonotactics cannot be violated. Therefore, in a word like <often>, <t> can acquire a phonetic value but in a word like <knight>, <k> cannot correspond to /k/, it must stay silent. This is because English phonotactics does not allow a sequence like /kn/ in the onset. I will consider then the following constraints:

(7)

| | |
|---|---|
| VISIBIL-PHONEME | All the phonemes in the input must have a corresponding grapheme in the output. |
| MINSON | The minimum sonority distance between two consonants before a nucleus must be ≥ 2, following this scale (Kenstowicz 1994): glides 4, liquids 3, nasals 2, obstruents 1. |
| DEP-L | No silent letters. |

Tableau 7: English <often>

| | VISIBIL-PHON | MINSON | DEP-L |
|---|---|---|---|

| | | | |
|---|---|---|---|
| ☞a) <often> /ɒftən/ | | | |
| b) <often> /ɒfən/ | | | *! |

In Tableau 7 there is no actual input, the candidates under evaluation are orthographic form – phonological form pairs. Under this ranking, (a) is a better candidate than (b) because if <t> corresponds to /t/ instead of being silent, none of the relevant constraints is violated (therefore the relevant tableau cells are left empty).

Tableau 8: English <knight>

| | VISIBIL-PHON | MINSON | DEP-L |
|---|---|---|---|
| ☞a) <knight> /naɪt/ | | | *** |
| b) <knight> /knaɪt/ | | *! kn =1 | ** |

Tableau 8 shows that <knight> cannot possibly be pronounced with initial /kn/ because this sequence would violate MINSON (/k/ = 1, /n/ = 2, 2 – 1 = 1, which is < 2), ranked higher than DEP-L. No candidate violates VISIBIL-PHON, since all the phonemes have a written correspondent, but both violate DEP-L, but (b) only twice (<g> and <h> are silent, but <k> is not). As a matter of fact, <igh> functions as a complex grapheme (a grapheme composed of more than one letter) but maintains nevertheless a regular 1:1 relationship with the diphthong /aɪ/. <igh> can only correspond to /aɪ/ in English, but /aɪ/ can also be spelled <i> before a consonant grapheme followed by <e>, like all other tense vowels (compare, for example, *mat* vs. *mate*, *ton* vs. *tone, hug* vs. *huge*, etc.). This kind of mappings can be expressed in English through bidirectional constraints (Baroni 2013):

(8)

| | |
|---|---|
| <VCe> ↔ TENSE V | A vowel grapheme followed by a consonant grapheme and <e> maps onto a tense vowel, and vice versa. |
| <igh> ↔ /aɪ/ | <igh> maps bidirectionally onto /aɪ/. |

The main difference between the two constraints is that the latter is very specific; it only applies to one sound and one grapheme, whereas the former applies to a class of sounds: all English tense vowels. It is therefore more general. Both constraints play a role in standard orthography, cf. *sight* vs. *site*, *might* vs. *mite*, etc. I argue that in Eye Dialect, more general constraints are ranked higher than more specific ones, and that the input only consists in the phonological form. In standard orthography, on the other hand, the input must also contain some information about the orthographic form, otherwise no ranking could explain the coexistence, in the same system, of <igh> and <iCe> to note /aɪ/. It appears likely that, at least for proficient readers and spellers, orthographic representations are present underlyingly, paired with the phonological ones (Katz & Frost 2001). This is also consistent with the fact that Eye Dialect is a stylistic device aiming to convey non-cultivated speech, ignoring the norm (especially the orthographic norm) on purpose.

Tableau 9: Standard English <tonight>

| /tʊnaɪt/ <tonight> | INPUT-TO-OUTPUT FAITHFULNESS | <igh> ↔ /aɪ/ | <VCe> ↔ TENSE V | DEP-L |
|---|---|---|---|---|
| ☞a) /tʊnaɪt/ <tonight> | | | * | ** |
| b) /tʊnaɪt/ <tonite> | *! | * | | * |
| c) /tʊnaɪt/ <tonit> | *! | * | * | |

In Tableau 9 an INPUT-TO-OUTPUT FAITHFULNESS constraint is ranked high in the hierarchy and bans any candidate that does not contain <gh>. Even if the orthographic form were ignored and the input only consisted in the phonological form, (b) and (c) would still be ruled out since <igh> ↔

/aɪ/ is ranked higher than <VCe> ↔ TENSE V.

Tableau 10: Eye Dialect <tonite>

| /tʊnaɪt/ | INPUT-TO-OUTPUT FAITHFULNESS | <VCe> ↔ TENSE V | <igh> ↔ /aɪ/ | DEP-L |
|---|---|---|---|---|
| a) /tʊnaɪt/ <tonight> | | *! | | ** |
| ☞b)/tʊnaɪt/ <tonite> | | | * | * |
| c) /tʊnaɪt/ <tonit> | | *! | * | |

Tableau 10 shows the Eye Dialect ranking. The input only contains the phonological form and the candidates are possible phonological form – orthographic form pairs. Since there is no orthographic information in the input, faithfulness constraints are not violated by any of the candidates. The more general constraint <VCe> ↔ TENSE V dominates the more specific one, <igh> ↔ /aɪ/, and both are ranked higher than DEP-L. Candidate (b) is the winner, since the tense vowel is rendered by <i> + <Ce>. Candidate (a) is dismissed because it violates the highest ranked constraint and so is (c), although it is the only candidate without any silent letter.

## 7. CONCLUSION

The aim of this paper was to show that Optimality Theory is suitable to describe orthographic phenomena, given that writing systems and orthographies are normally based on a series of conflicting principle that can be easily translated into violable constraints. In particular, I have tried to focus on the grounding of the constraints that might play a role in orthographic systems. I proposed that some constraints might be grounded in cognition (e.g., SIMPLICITY constraints), some in perceptual salience (e.g., VISIBILITY and *VISIBILITY families) and others might be simply based on system-internal factors (e.g., bidirectional grapheme ↔ phoneme constraints). Even languages employing the same writing system may rank constraints differently, representing the same phonological features with different solutions. Both German and Finnish represent vowel length, but in German not all vowel graphemes can be doubled, so other solutions have to be preferred. Two languages that possess phonological long consonants can employ different writing systems and constraint rankings, for example, Italian represents long consonants by doubling the consonant grapheme whereas Punjabi, that employs an abugida, places a diacritic sign on the grapheme standing for the syllable that precedes the long consonant. When it comes to opaque orthographies, such as English and French, spontaneous phenomena like Spelling Pronunciation and Eye Dialect promote 1:1 relationships between sound and letters and/or extend the more general patterns at the expense of the more specific ones. Even if one of the most striking facts about writing systems is their external dissimilarity, with this paper I hope to have shown that at least some constraints seem to be active both in related and unrelated orthographies. Beyond the visual diversity, writing systems and orthographies rely on a finite set of constraints that are ranked differently depending on the orthography considered. Put differently, the analysis of orthographic phenomena in OT allows us to account for differences among writing systems through constraint ranking but at the same time shows us what is universal about writing.

## BIBLIOGRAPHY

ARTYMOVYĆ, Agenor (1932) « Fremdwort und Schrift », *Charisteria Guilelmio Mathesio quinquagenario a discipulis et Circuli Linguistici Pragensis sodalibus oblata*, Prague, Pražský Linguistický Kroužek.

BARONI, Antonio (2011) « Alphabetic vs. non-alphabetic writing: Linguistic fit and natural tendencies », *Italian Journal of Linguistics/Rivista di Linguistica* 23.2, 127-159. BARONI, Antonio (2013). « Eye Dialect and Casual Speech Spelling: Orthographic Variation in OT », *Writing Systems Research* 5.1, 24-53. BARONI, Marco (2001). « How Do Languages Get Crazy Constraints? Phonetically-Based Phonology and the Evolution of the Galeata Romagnolo Vowel System », *UCLA Working Papers in Linguistics* 7, 152-178.

BLOOMFIELD, Leonard (1933) *Language,* New York, Henry Holt.

BOWDRE, Paul Hull jr. (1964) *A Study of Eye Dialect*, thèse de doctorat, Université de Floride.

CHOMSKY, Noam et Halle, Morris (1968). *The Sound Pattern of English*, New York, Evaston & Londres, Harper & Row.

COULMAS, Florian (1989) *The writing systems of the world*, Oxford, Blackwell.

COULMAS, Florian (2009). « Evaluating merit – the evolution of writing reconsidered », *Writing Systems Research* 1.1, 5-17.

DRESSLER, Wolfgang U., MAYERTHALER, Willi, WURZEL, Wolfgang U., et PANAGL, Oswald (1987) *Leitmotifs in Natural Morphology*, Amsterdam, John Benjamins.

FISCHER-JØRGENSEN, Eli (1949) « Remarques sur les principes de l'analyse phonémique », *Travaux du Cercle Linguistique de Copenhague* 5, 213-234.

GRAY, William Scott (1956) *The Teaching of Reading and Writing*, Paris, UNESCO.

GRIMSHAW, Jane (1997) « Projections, heads and optimality », *Linguistic Inquiry* 28, 373-422.

HEINZ, Matthias (2012) « Complex syllable types in Romance languages: *central* or *peripheral*? », handout, *International Conference on Romance Languages – in memoriam Alf Lombard, 7-8 Novembre 2012*, Université de Lund, Suède.

HJELMSLEV, Louis (1943) *Omkring sprogteoriens grundlæggelse.* Copenhague, Akademisk forlag.

HOŘEJŠÍ, Vladimir (1971) « Formes parlés, formes écrites et systèmes orthographiques des langues », *Folia linguistica* 5, 185-193.

JAKOBSON, Roman (1963) *Essais de linguistique générale,* Paris, Minuit.

JUSTESON, John S. (1976) « Universals of Language and Universals of Writing », JUILLAND, Alphonse, DEVINE, A. M. et STEPHENS, Laurence D. (éd.), *Linguistic studies offered to Joseph Greenberg* 1, Saratoga, Anma Libri, 57-94.

KATZ, Leonard et FROST, Stephen J. (2001) « Phonology constrains the internal orthographic representation », *Reading & Writing: An Interdisciplinary Journal* 14, 297-332.

KENSTOWICZ, Michael (1994) *Phonology in Generative Grammar.* Oxford, Blackwell.

LEGENDRE, Géraldine (2000) « Morphological and prosodic alignment of Bulgarian clitics », DEKKERS, Joost, VAN DER LEUW, Frank et VAN DER WEIJER, Jeroen (éd.), *Optimality Theory : phonology, syntax and acquistion,* Oxford, Oxford University Press, 423-462.

LEGENDRE, Géraldine, GRIMSHAW, Jane et VIKNER, Sten (2001) *Optimality-theoretic syntax,* Cambridge, MIT Press.

LEHTONEN, Annukka (2006) « Sources of Information Children Use in Learning to Spell: the Case of Finnish Geminates », MALATESHA, Joshi et AARON, P. G. (éd.), *Handbook of Orthography and Literacy*, Oxford et New York, Routledge.

LEVITT, Jesse (1978) « The Influence of Orthography on Phonology: A Comparative Study (English, French, Spanish, Italian, German) », *Linguistics* 208, 43-67.

MCCARTHY, John et PRINCE, Alan (1993) *Prosodic Morphology I: Constraint Interaction and Satisfaction*, Rutgers University.

NEUMAN, Yishai (2009) *L'influence de l'écriture sur la langue*, thèse de doctorat, Université de Paris III, Sorbonne Nouvelle.

PONTECORVO, Clotilde (1994) « Iconicity in Children's First Written Texts », Simone, Raffaele (éd.), *Iconicity in Language*, Amsterdam et Philadelphia, John Benjamins, 277-307.

PRIMUS, Beatrice (2004) « A Featural Analysis of the Modern Roman Alphabet », *Written Language & Literacy* 7.2, 235-274.

PRINCE, Alan, et SMOLENSKY, Paul (1993) *Optimality Theory: constraints interaction in generative grammar,* Rutgers University.

SAMPSON, Geoffrey (1985) *Writing systems*, London, Hutchinson.

SAUSSURE, Ferdinand de (1916) *Cours de linguistique générale,* Paris, Payot.

SCHEER, Tobias (2004*) A Lateral Theory of Phonology. Vol. 1: What is CVCV and why should it be?*, Berlin, Mouton de Gruyter.

SEBBA, Mark (2006) « Ideology and alphabets in the former USSR », *Language Problems and Language Planning* 30.2, 99-125.

SMALLEY, William A. (1964) *Orthography Studies*, London/Amsterdam, United Bible Society/North-Holland Publishing Company.

SONG, Hye Jeong et WIESE, Richard (2010) « Resistance to complexity interacting with visual shape – German and Korean orthography », *Writing Systems Research* 2.2, 87-103.

SWART, Henriette de (2006) « Marking and Interpretation of Negation: A bidirectional-OT approach », ZANUTTINI, Raffaella et CAMPOS, Hector (éd.), *Negation, Tense and Clausal Architecture: Cross-linguistic Investigations*, Georgetown, Georgetown University Press, 199-218.

VACHEK, Josef (1939) « Zum Problem der geschriebenen Sprache », *Travaux du Cercle Linguistique de Prague.*

WEINGARTEN, Rüdiger (2011) « Comparative graphematics », *Written Language & Literacy* 14.1, 12-38.

WIESE, Richard (2004) « How to optimize orthography », *Written Language & Literacy* 7.2, 305-331.