



Strength-based faithfulness and the sibilant /s/ in Italian

Antonio Baroni

Università degli Studi di Padova
antonio.baroni@studenti.unipd.it

Abstract

Notwithstanding the primacy of the CV syllable, a number of languages allow for more complex types of syllables. In particular, word-initial consonant clusters are particularly challenging for any phonological theory. In this paper it is argued that obstruent clusters may be the result of casual speech processes where the most salient/frequent phonemes and features occurring in most pronunciation variants of a word are preserved. As a result, sibilants, being acoustically salient, tend to occur more often than other obstruents as the first member of word-initial obstruent clusters. A framework couched in Optimality Theory is presented, where a subfamily of faithfulness constraints refer to strength values stored in the underlying representation. The more salient and/or frequent a phoneme/feature is, the higher the strength value assigned to it. Finally, a number of languages are compared, arguing that their phonotactic differences may be due to the different ranking of markedness constraints and MAX-STRENGTHVALUE constraints.

Keywords: complexity; obstruent clusters; casual speech; phonotactics; acoustic salience.

1. Introduction¹

The current paper attempts to investigate the relationship between casual speech phenomena and the phonotactics of consonant clusters, focusing in particular on clusters that do not respect the Sonority Sequencing Generalization (henceforth SSG; Selkirk 1984; Clements 1990). SSG states that sonori-

¹ Conventions and abbreviations: C: Any consonant; O: Any obstruent; V: Any vowel; R: Any sonorant; #: Word boundary; P: Any labial stop; SSG: Sonority Sequencing Generalization; K: Any dorsal stop; OT: Optimality Theory; T: Any stop; MOA: Manner of articulation; T_{cor}: Any coronal stop; POA: Place of articulation; S: Any sibilant fricative; OCP: Obligatory Contour Principle; F: Any non-sibilant fricative.

ty should rise from the onset to the nucleus and fall from the nucleus to the coda. Well-formed consonant clusters therefore consist of an obstruent + sonorant sequence word-initially (#OR), and of a sonorant + obstruent sequence word-finally (RO#). The existence of obstruent clusters in a number of languages represents a challenge for most phonological theories. However, it is evident that some obstruent clusters are better formed than others, at least since they occur more frequently crosslinguistically. In the following section, it is argued that, regardless of the primacy of the CV syllable type, other types of syllables can arise through deletions in casual speech processes, as well as in morphological processes. Section 3 introduces the Optimality Theory framework and proposes a subfamily of faithfulness constraints, MAX[STRENGTHVALUE], whose role is to preserve the most frequent and the most salient phonemes and features occurring in all the pronunciation variants of a word. In Section 4 it is argued that the aforementioned subfamily of constraints is responsible for the preservation of acoustically prominent consonants, such as /s/, in pre-consonantal position, which explains the frequency of #SO clusters. Finally, in Section 5, it is suggested that taking into consideration casual speech phenomena, frequency and acoustic salience can help us understand crosslinguistic phonotactic preferences.

2. Syllable complexity and the primacy of CV

It is a well-known fact that languages differ in the degree of syllable complexity they allow. For instance, the only possible syllable type in Hawaiian is CV, whereas most Indo-European languages tolerate moderately complex consonant clusters word-initially and word-finally. Cases of extreme complexity are rare but do exist: Gilbertese displays very long tautosyllabic vowel sequences, as in the augmentative suffix *-kaaei* (Blevins 2004: 213), whilst Georgian allows consonant clusters of considerable length, as in *gvprckvnis* ‘he peels us’ (Comrie 1981: 200). Hawaiian, on the one hand, and Gilbertese and Georgian, on the other hand, can be seen as the two extremes of a continuum of syllable complexity, but with an important difference: the CV syllable type, the only possible sequence in Hawaiian, is the fundamental type of syllable. There is no language banning CV syllables (Maddieson 2011). This is particularly evident in the phonological adaptation of loanwords from a source language with a complex syllable structure in a target language allowing only CV syllables. For instance, the Dutch word *plan* ‘plan’ is

adapted in Sinhalese as *pālāna*, thus disrupting the consonant sequence (Borersma *et al.* 2000: 5). Besides the evidence from loanwords, there are other reasons to consider CV as the basic syllable type:

- It is the first type of syllable produced by infants (Levelt *et al.* 1999).
- It is perceptually optimal (Ohala and Kawasaki 1997). As a matter of fact, consonants (stops in particular) possess both internal and external cues for their recognition, and being followed by a vowel increases their perceptibility. Most of the times, listeners pay attention to the way that vocalic formants are affected by the preceding consonant in order to recognize its place of articulation.
- It is articulatorily simple. More specifically, MacNeilage and Davis (2000) argue that the CV syllable type derives from mandibular oscillation – the opening-closing movement of the jaw.

If CV is the preferred syllable type, why do languages tolerate CCV syllables so frequently? How do complex syllable types arise and how do they pass the test of time? I argue that, like many phonological phenomena, the emergence of complex syllable types results from casual speech processes, such as unstressed vowel reduction and deletion. Coarticulation and assimilation are pervasive in running speech, but in a number of languages they remain at the mere phonetic level, without affecting the phonological structure. Obviously, in other languages, labialization, palatalization, vowel harmony, etc. can instead be “promoted” and qualify as actual phonological processes (Blevins 2004: 126). Quite similarly, in $C_1V_1C_2V_2$ sequences, the unstressed vowel can be reduced to schwa or be deleted altogether, resulting in $C_1V_1C_2$ if the stress falls on V_1 and $C_1C_2V_2$ if the stress falls on V_2 . For example, Italian *specchio* ‘mirror’ /*ˈspɛkːjo*/ comes from Latin *speculu(m)* /*ˈspɛkulu(m)*/, which at some point lost the post-tonic /*u*/, resulting in */*ˈspɛklu(m)*/, and the cluster /*kl*/ evolved into /*kj*/ (Canello 1872: 209). The Spanish words *isla* ‘island’, *alma* ‘soul’, which contain consonant clusters, come from Latin words *insula* and *anima*, where the deletion of the unstressed vowel triggered a consonant sequence (and in the case of *anima* a dissimilatory process **n* > *l* [Penny 2004: 84–88]). English *family* /*ˈfæmɪli*/, *laboratory* /*ˈlæbrətəri*/ and *temperature* /*ˈtɛmpərətʃər*/ are typically pronounced with a word-internal cluster, as in [*ˈfæmli*, *ˈlæbrətəri*, *ˈtɛmprətʃər*].

Many words and phrases can have several pronunciation variants, one with a CVC sequence and one with a consonant cluster, e.g., English *suppose*

[sə'pəʊz] vs. ['spəʊz], French *regard* [ʁə'gɑʁ] vs. ['ʁgɑʁ] 'gaze', Italian *per esempio* [per'e'zempjo] vs. [pre'zempjo] 'for example'. Typically, the CVC-variant belongs to the *lento* style and the CC-variant to the *allegro* style. However, in the aforementioned cases, these predictable deletion processes have not been lexicalized, so it is highly unlikely that ['spəʊz], ['ʁgɑʁ] and [pre'zempjo] will appear in isolation in formal, *lento* style. Moreover, French phonotactics does not allow word-initial clusters that imply a sonority reversal, i.e., where a sonorant precedes a stop, and in Italian */pr/ is not a possible allomorph of *per* 'for'. Lexicalization of reduced forms does occur, however, and English forms such as *isn't*, *aren't*, *haven't*, *wouldn't* are evidently derived from casual speech pronunciations of unreduced *is not*, *are not*, *have not*, *would not*. Here, their lexicalization is also sanctioned by the existence of an official spelling. The emergence of both coda consonants and consonant clusters seems to be largely due to the deletion of unstressed vowels. For the remainder of this paper, the focus will be only on consonant clusters, as their analysis is still problematic in modern phonological theories. However, it is important to point out that, even if it is not the subject of the current work, also concatenative morphology can produce consonant sequences. Such clusters are investigated by *morphonotactics* (Dressler and Dziubalska-Kołaczyk 2006; Dressler et al. 2010; Dziubalska-Kołaczyk and Zielińska 2011; Dziubalska-Kołaczyk et al. 2013), i.e., a subfield of morphonology dealing with consonant clusters arisen because of morphological processes (typically concatenation, but also non-concatenative processes can produce the same effect, e.g., Polish *lew* 'lion' vs. *lwy* 'lions' (Dressler and Dziubalska-Kołaczyk 2006: 259)).

3. Optimality Theory

The model that will be proposed here is couched in the framework of Optimality Theory (Prince and Smolensky 1993/2004; McCarthy & Prince 1993, 1995, Kager 1999) and will attempt to account for the typological frequency of certain clusters as opposed to others, which occur rarely, if ever. According to Optimality Theory (henceforth OT), the grammar is articulated in three components: GEN (Generator), EVAL (Evaluator) and CON (Constraint set). Given an input, GEN generates all the possible outputs, which are subsequently fed to EVAL, where they are all evaluated simultaneously

according to a language-specific hierarchy of universal constraints. Phonological constraints are of two types: faithfulness and markedness. Faithfulness constraints preserve the input from undergoing changes. If all faithfulness constraints are undominated in the hierarchy, the input-output correspondence will be one-to-one. Markedness constraints militate against marked structures, e.g., since voiced obstruents are more marked than voiceless ones, there is a constraint *VOICEDOBSTRUENTS (Alderete 1997) assigning a violation for every voiced obstruent present in the output.

It is generally accepted by phonologists who work in OT that markedness constraints must be grounded in phonetics, i.e., they must be perceptually or articulatorily motivated (Steriade 2001; Flemming 2002; Hayes et al. 2004). For example, a constraint like NOCODA (Prince and Smolensky 1993) exists because the cues for the recognition of a consonant are considerably impoverished if it is not followed by a vowel, while a constraint like AGREE[FEATURE] (Baković 2005) expresses the general tendency towards coarticulation and assimilation. Grounding constraints in phonetics is a way for “constraining the constraints”, i.e., ensuring that no phonologist comes up with a CODA (“Syllables must have a coda”) or a NOONSET (“Syllables must not have onsets”) constraint. However, as many scholars have pointed out, grounded constraints are problematic with regards to the question of innateness: if constraints are learnable through experience with our acoustic and articulatory apparatuses, what is the use of replicating this knowledge in the mind (see Hale and Reiss 2000 for discussion)? Whether constraints are also innate or just learned is not relevant here, therefore I will not discuss this further. However, I will assume the existence of at least some “empty” constraints, i.e., constraints consisting of contentless templates, e.g., MAX[_____], ID[_____], NO[_____], following the proposal presented in Blaho (2008: 41). While it is quite apparent that markedness constraints are derived from (or at least, have some relation with) perceptual and articulatory ease, the origin of faithfulness constraints is less clear. As Collins points out, “faithfulness constraints are not conceptually difficult or puzzling in the same way as markedness constraints. [...] [T]hings don’t happen unless something causes them to happen. [...] There is no scientific problem to be answered regarding their origins because it is literally impossible for anything in the universe not to obey this principle” (Collins 2013: 43–44). While I mostly agree with Collins, I think it is nevertheless necessary to understand how faithfulness constraints work and are ranked. What is interesting about faithfulness constraints is not why they exist, but why they choose to pre-

serve what and why they are ranked the way they are. I argue that the content of faithfulness constraints and their position in the constraint hierarchy depends mainly on two factors: frequency in the acoustic input and salience. Let us take the example of the sibilant [s]. Experience with articulation teaches us that [s] is relatively hard to articulate, as opposed to stops and the glottal fricative [h]. Markedness would therefore militate against sibilants, as it is evident in the late production of this class of sounds in children (Sanders 1973). However, sibilants are among the most acoustically salient segments and they tend to be very frequent in speech (think, for example, of how often monophonemic morphemes consist of a sibilant in Indo-European languages). Consequently, faithfulness will be “informed” of the importance of [s] and the faithfulness constraint protecting it will be ranked high. This interaction between faithfulness and markedness is particularly evident in casual speech. Casual pronunciation in running speech is traditionally described as driven by articulatory ease, i.e., speakers aim to reduce effort while speaking when the situation is familiar/informal (Ernestus 2000). The most efficient way to reduce effort is to decrease the number of articulatory gestures necessary to produce speech. Obviously, not all gestures can be dispensed with, otherwise communication would fail. I call the acoustic image of a word which is preserved in casual speech the “invariant”. The invariant consists of phonemes and/or features that are essential for the recognition of a particular word. If the only drive in casual speech were to reduce articulatory effort, one should expect segments that are relatively costly from an articulatory point of view to be the first to go. However, this is not the case. A study on Italian casual speech (Baroni 2014, based on four dialogues extracted from the corpus of spontaneous Italian CLIPS (Savy and Cutugno 2009)) showed that sibilants were practically never reduced or deleted, regardless of their position, whereas all the other consonants and the vowels could potentially disappear.² More specifically, /t/ was the least likely to be preserved, especially after /s/, probably because of their spectral similarity (cf. Ernestus 2000: 113–114), followed by the unstressed vowels and the occlusive component of /n/. Conversely, nasalization and r-coloring tended to be preserved, to such an extent that the author concludes that the invariant of /sinistra/ ‘left’ is [s~sR], where [~] stands for nasality and [R] for r-coloring. It is hard to

² Consonant and vowel reduction and deletion appears to be a phonetic phenomenon in Italian, except for a few lexicalized cases (Bertinetto and Loporcaro 2005). See also Krämer (2009) for a standard OT account of Italian phonology.

see how pronouncing something like [s~sR] might be easier than pronouncing unreduced [sinistra], even though the articulatory gestures involved are fewer. In general, it is typical of casual speech to reduce and delete vowels to a much greater extent than consonants, even though vowels are arguably easier than consonants, articulation-wise (for languages other than Italian, see Shockey 2003 for English; Phrao 2009 for Danish; Ernestus 2000 for Dutch; Kohler 1990 for German). Harris (2006: 1491) points out that speech is somehow paradoxical, in that most of the sound energy is concentrated in vowels but most of the linguistically relevant information is borne by consonants. This inherent characteristic of speech is directly encoded in the Onset Prominence framework (Schwartz 2009, 2013) but is normally not explicitly formalized in other phonological theories. Going back to the example of *isn* *t* (see section 2), how can [ɪzn̩t] be said to be easier to pronounce than [ɪznɒt]? Markedness would suggest that the syllabic nucleus has to be preserved and that coda consonants must be dispensed with, but here the vowel is deleted and syllabicity is acquired by the nasal consonant, while the coda [z] is preserved. It is therefore apparent that the preservation of certain phonemes/features is listener-oriented. Of course, articulatory gestures are reduced, but only as long as acoustically salient components are preserved.

3.1 Faithfulness to the strong

A number of psycholinguistic studies (Cole et al. 1974; Goldinger 1998; McLennan et al. 2003; Pierrehumbert 2002; McLennan and Luce 2005; Goldinger 2007; Mattys and Liss 2008; Ernestus 2013) bring evidence that more than one pronunciation variant of the same word is stored in the mental lexicon and that reduction is word-specific and sensitive to both structural factors and frequency. As a result, frequent words are more likely to undergo reduction than rare words, short words are more likely than long words and function words more likely than content words (Bürki et al. 2010; Hinskens 2011). How to account for that in OT? First of all, following van Oostendorp's proposal (1997), I assume that faithfulness and markedness constraints are ranked differently depending on the speech style (e.g., casual, intermediate, careful, etc.). More specifically, faithfulness constraints maintain their ranking in relation with each other, and the same applies to markedness constraints, but the more casual the style, the higher markedness constraints

are ranked in relation to faithfulness constraints. This proposal is exemplified in (1).

- (1) F = faithfulness constraint, M = markedness constraint
 F1 > F2 > F3 > F4
 M1 > M2 > M3 > M4
 Careful style: F1 > F2 > F3 > F4 > M1 > M2 > M3 > M4
 Intermediate style: F1 > F2 > M1 > M2 > F3 > F4 > M3 > M4
 Casual style: F1 > M1 > M2 > M3 > M4 > F2 > F3 > F4.

Van Oostendorp's model is a possible solution to the problem of variation in OT, but does not tackle the issue of word-specific reduction. Since assuming a different constraint ranking for each word seems too costly and inelegant, I posit the existence of strength values associated to the different phonemes/features composing the underlying representation of words.³ These values are derived from the frequency and the salience of each phoneme/feature in the input (see also Baroni and Simonović 2014). Here by frequency I mean the persistence of a certain segment of feature in the pronunciation variant of a word, not absolute frequency. For instance, /ʃ/ in Italian is relatively rare, but in the pronunciation variants of the words in which it occurs is hardly deleted, as will be shown later. Let us consider again Italian /sinistra/ 'left'. When the child is exposed to pronunciation variants of /sinistra/, initially the strength of [s] might be, say, 1, and after 1000 tokens of the same word, all containing acoustic correlates of [s], its strength will be 1000, and so on. A similarly strong status will be assigned to nasality and rhoticity, whereas palatality will be slightly weaker and other elements even less strong. At some point, the child will construct a fully detailed representation of /sinistra/, including strength values assigned to the different phonological components of the word. The only way to account for different degrees of reduction depending on the frequency of the lexical item and on the relative salience of its phonemes is to posit the existence of a subfamily of faithfulness constraints that refer to strength values present in the underlying representation, rather than to a specific class of segments. Expressed differently, there are faithfulness constraints like MAX-C or MAX-V (McCarthy

³ A similar stance is taken by Schwartz (2009: 150): "In Prominence phonology, the representations of lexical items are assumed to be specified for the relative prominence of the segments they contain. [...] This relative prominence is meant to reflect which segments are the most susceptible to lenition and which must be preserved when the word is spoken".

2000) that refer to consonants and to vowels, respectively, as a whole class: “Do not delete consonants”, “Do not delete vowels”. The subfamily of faithfulness constraints that I am proposing here has a different structure, i.e., MAX[STRENGTHVALUE] (“Preserve segments/features with X strength value”). The strength values are stored in the underlying form together with the phonemes and features they are assigned to and obviously MAX[STRENGTHVALUE] constraints referring to greater strength values are ranked higher in the hierarchy. The (re)ranking of other faithfulness constraints and markedness constraints accounts for variation, as shown in (2), whereas MAX[STRENGTHVALUE] constraints ensure that the invariant is left untouched.

- (2) Careful style: Max[100] > Max[80] > Max[20] > M
 Intermediate style: Max[100] > Max[80] > M > Max[20]
 Casual style: Max[100] > M > Max[80] > Max[20].

MAX[100] indicates that the phonemes/features that appear in all the tokens of a word must be preserved. MAX[100], by definition, is undominated in the hierarchy. Some might argue that the reasoning risks being circular: the invariant cannot be deleted because it is the invariant. However, the question is less trivial than it might seem: the theory needs to be able to explain how the invariant is formed. As a matter of fact, the features and the segments that form part of the invariant are not selected arbitrarily, but they are typically the ones which bear the most salient cues for word recognition and that consequently appear in most/all pronunciation variants of that word. Imagine that the underlying representation of *sinistra*, /sinistra/, contains the following strength values: /s/, nasality, rhoticity → 100, palatality → 80, /t/ → 20. MAX[80] refers to palatality ([I] or [+high, -back]), not to the front vowel *per se*, so /i/ can be missing as long as palatality is borne by another segment. Instead, MAX[100] refers to all the features identifying stridents, it is therefore a shortcut standing for MAX[+CONSONANTAL, +STRIDENT, +CONTINUANT, +CORONAL, +ANTERIOR]. MAX[100] also protects nasality (or [+nasal]) and rhoticity (probably something like [low F3], cf. Blevins and Garrett 2004). Markedness constraints accounting for typical casual speech processes are AGREE-N (“Adjacent segments must agree in nasality”), AGREE-I (“Adjacent segments must agree in palatality”), and *V (“No vowels in the output”)⁴. *V

⁴ In order to account for the so-called “emergent stops” (Ohala 1997), another constraint should be postulated. In casual speech, a sequence of a nasal followed by a fricative or a liquid is often

is grounded in the tendency of speakers, in running speech, to delete vowels. I wish to remain theory-neutral with regards to features (binary or privative) or elements (such as the ones employed in Element Theory, cf. Kaye et al. 1985; Harris 1990; Backley 2011). Here I find more elegant to use AGREE-N and AGREE-I instead of, say, AGREE[+NASAL] or AGREE[+HIGH, -BACK], but the choice of acoustic primes is not relevant for the current analysis. The position in the ranking of AGREE-N, AGREE-I and *V, however, cannot characterize Italian grammar in the same way as, say, *COMPLEXCODA (Rose 2000) does. Italian native words cannot display complex codas (except in loanwords) and this fact is somehow wired in the core phonological grammar of Italian. Conversely, palatalization and nasalization are not pervasive phenomena: the former is limited to some specific cases (e.g., the plural of certain masculine nouns ending with /-ko, -go/, as *amico*, *amici* /a'mi:ko, a'mi:tʃi/ 'friend, friends' – but *antico*, *antichi* /an'ti:ko, an'ti:ki/ 'ancient'), the latter is a low-level phonetic phenomenon. It is therefore quite likely that they become relevant only when they are promoted in the hierarchy in less formal speech styles. In (3) three possible constraint rankings are proposed, each reflecting a different style.

- (3) Careful style: MAx[100] > Max[80] > Max[20] > Agree-N, Agree-I > *V.
 Intermediate style: Max[100] > Max[80] > Agree-N, Agree-I > Max[20] > *V.
 Casual style: Max[100] > Agree-N, Agree-I, *V > Max[80] > Max[20].

For each style, I selected a representative example from the CLIPS corpus: [sinistra, sɪ̃nsra, s~sr].

avoided by inserting a stop which is homorganic with the nasal, e.g., English *prince* pronounced [prints], Southern Italian *ansia* 'anxiety' pronounced [antsja]. Such pronunciations can be lexicalized, as in the English surnames *Thom[p]son*, *Sam[p]son* or in French *nombre* 'number' (<Latin **numerus*). Former OT analyses (Smolensky 1995; Alderete 1997) make use of conjoined constraints, such as *NC & *FRIC, which are violated when the candidate contains a sequence a nasal followed by a voiceless fricative. Warner (2002) proposes the articulatorily based constraint *NASREL, banning the release of the oral closure of a nasal. I refer to Warner's paper for a deeper analysis of the phenomenon.

Tableau 1. Careful style.

/sinistra/	MAX[100]	MAX[80]	MAX[20]	AGREE-N	AGREE-I	*V
☞ (a) sinistra				**	**	***
(b) sɪ̃nsra			*!	*	*	***
(c) s~sr		*!*	*	**		
(d) initra	*!*			**	**	***

In careful style, all the relevant faithfulness constraints are ranked higher than markedness constraints, so that (b) is ruled out because the second /i/ of the word is missing, (c) loses the competition because palatality (borne by the two /i/ vowels) has been deleted altogether and (d) will never emerge as a possible winner, under any ranking, since it violates MAX[100], which, by definition, is always undominated (i.e., the sibilants of *sinistra* can never be deleted).

Tableau 2. Intermediate style.

/sinistra/	MAX[100]	MAX[80]	AGREE-N	AGREE-I	MAX[20]	*V
(a) sinistra			*!*	**		***
☞ (b) sɪ̃nsra			*	*	*	***
(c) s~sr		*!*	**		*	
(d) initra	*!*		**	**		***

In Tableau 2, the two AGREE constraints dominate MAX[20], and therefore palatality can be spread from the vowel to the preceding nasal consonant, and at the same time, the nasal consonant is able to nasalize the preceding vowel, even though the deletion of /t/ violates MAX[20]. Note that MAX[80], which protects palatality, is not violated here, although the second /i/ vowel is missing, because its palatality is borne by another segment but is still present.

Tableau 3: Casual style.

/sinistra/	MAX[100]	AGREE-N	AGREE-I	*V	MAX[80]	MAX[20]
(a) sinistra		**	*!*	***		
(b) sɪ̃nsra		*	*	*!***		*
☞ (c) s~sr		**			**	*
(d) initra	*!*	**	**	***		

In casual style the winning candidate is (c), because *V now dominates MAX[80] and MAX[20], meaning that it is better to dispense with all the vowels rather than preserving /i/ and /t/. The result is [s~sr], something that would hardly form part of Italian lexicon but that surely surfaces in running speech.

The tendency of sibilants to be preserved even in heavily reduced speech does not show in *sinistra* exclusively. Baroni (2014) analyzed the behavior of Italian voiceless obstruents /s, ʃ, f, k, p, t/ in spontaneous speech. In the four dialogues of the CLIPS corpus under analysis, each occurrence of a voiceless obstruent was annotated and its realization (deleted, lenited, fully pronounced) was classified according to its position in the word (word-initial onset, word-internal onset, intervocalic, internal coda, part of a word-initial cluster, part of a word-internal cluster). It emerged that, in spoken Italian, the three fricatives are much more resistant and less context-dependent than the three stops. /t/ was the most likely to be deleted and also the most frequent, suggesting a correlation between frequency (predictability) and likeliness to undergo deletion. However, /s/ was the second most frequent obstruent in the sample and was fully pronounced 95% of the time and deleted less than 1% of the time. Crucially, /s/ is not only much more frequent than /ʃ, f/ in the dialogues under analysis (750 tokens of /s/ vs. 61 of /ʃ/ and 158 of /f/) but is also the only obstruent occurring in the coda (Italian phonology allows other obstruents to occupy the coda only when they are the first half of a geminate). In other words, frequency cannot be considered a good indicator for the deletion rate of a consonant, at least in Italian, since the acoustically salient /s/, albeit highly frequent, is rarely, if ever, deleted. Ernestus (2000: 110) and Shockey (2003: 14) suggest that also in Dutch and English, respectively, /t/ is particularly vulnerable and stridents particularly resistant, regardless of the

fact that they are both highly frequent. As for word frequency, assuming that the strength values associated to each segment/feature depend on the input received, the representation of highly frequent words will display a finer-grained distinction in strength between the units composing it. As a matter of fact, frequent words undergo a great deal of reduction and they are learned together with their intrinsic potential variation, i.e., with highly differentiated strength values.

4. Effects on phonotactics

One of the main assumptions of this paper is that consonant clusters arise, among other things, from deletion of vowels in casual speech. As the *sinistra* example showed, the complex cluster /str/ can be simplified in casual speech by deleting /t/, but the final outcome is far from being a sequence of CV syllables. However, it is probably not a coincidence that *sinistra* contains sibilants and sonorants and allows for such a great deal of variation, which arguably would not have taken place if the consonants of the word were only stops. As a matter of fact, the nature of the consonants of a word can be a good predictor of the degree of vocalic reduction that the word can undergo. Dalby (1986), in his study on casual American English, notices that the place and the manner of both the preceding and the following consonant are relevant to vowel reduction. Vowels are more likely to be deleted if preceded by a liquid (16%), a labial (obstruent or nasal, 11%), less likely if preceded by a coronal obstruent (6%) or a velar (4%). As for the following consonants, vowels are more likely to be deleted if followed by a coronal (obstruent or sonorant, 9%), less likely if followed by a velar (8%) or by a labial (6%). As for manner, vowel deletion occurs more often after continuants (fricatives, nasals and liquids) than after stops, and more often before stops than before continuants. The highest rate of deletion (22%) was found between a liquid and a coronal stop, such as in *ability*, followed by the liquid__velar context (16%), liquid__labial and labial__labial (9%). Not surprisingly, the most frequent consonant clusters in the world's languages normally contain a sonorant. It is well known that the preferred word-initial consonant cluster is #OR and the preferred word-final cluster is RO#. Accordingly, the rarest consonant clusters imply a reversal of sonority, such as word-initial #RO and word-final OR#. These are also the clusters more likely to be repaired by vocalic epenthesis (especially in the former case) or consonant deletion (especially in the latter

case). In the middle, there are plateau clusters, i.e., clusters of consonants of equal sonority, such as obstruent clusters and sonorant clusters. The idea is that, if phonotactic preferences are the result of frequent casual speech processes, since acoustic salience plays a role in the preservation of segments (sibilants are preserved regardless of their relative difficult articulation), plateau clusters containing highly salient segments will be more frequent in the world's languages. In particular, salience is likely to play a decisive role for consonants in pre-consonantal or word-final position, since they lack the support of a vowel. My predictions, based on acoustic salience and the dispreference of sequences of consonants with identical place or manner of articulation, are presented in (4).

- (4)
- (a) An obstruent cluster preferably contains a sibilant.
 - (b) Word-initially, #SO is the preferred order.
 - (c) Obstruent clusters in which one consonant is a stop and one is a fricative are more frequent than clusters where both consonants are either stops or fricatives.
 - (d) Consonants in an obstruent cluster prefer to differ in place of articulation.

In order to test my predictions, I gathered data from a sample of 39 languages, belonging to the following language families: Indo-European (Hodge 1946; Halle 1959; Steriade 1982; Nagamma Reddy 1987; Tankevičiūtė and Strimaitienė 1990; Hall 1992; Nestor 1993; De Schutter 1994; Kenstowicz 1994; Penzl 1995; Newmark 1998; Gussman 2007), Austronesian (Stern 1965; Wright 1996; Lock 2007), Mon Khmer (Henderson 1976; Nacaskul 1978), Sino-Tibetan (Jacques 2004), Siouan (Ullrich 2008), Benue-Congo (Ladefoged and Maddieson 1996), Caucasian (Butskhrikidze 2002), Semitic (Morelli 1999), Zapotec (Jaeger and Van Valin 1982; Marlett and Pickett 1987; Regnier 1993), Haida (Lawrence 1977), Yuman (Redden 1966), Plateau Penutian (Barker 1964), Chadic (Ladefoged 1964), Qiang (Hongkai 1986), Misanthla Totonac (MacKay 1994), Tsimshianic (Tarpent 1989), Otomi (Gibson 1956), Dravidian (Nagamma Reddy 1987), Chibchan (Oakes 2001), Caddoan (Rood 1975), Yuchi (Crawford 1973), Hokan (Marlett 1988). The data were extracted partly from Morelli (1999) and Parker (2012) and partly from Baroni (2014). In (5), for each possible word-initial obstruent cluster, I give the percentage of languages of the sample that allow for it.

- (5) SK (93.75%) > ST_{cor} (90.6%) > SP (87.5%) > PT_{cor} (59.4%) > PS (53.1%) > KS (50%) > T_{cor}K, KT_{cor} (34.4%) > PK (25%) > SF, KP (22%) > T_{cor}S, FP (18.75%) > TP (15.7%) > other clusters (less than 12.5%).

As expected, the three most frequent obstruent clusters begin with some kind of sibilant, while the fifth and the sixth most frequent clusters at least contain one. It is noteworthy that, in absence of a sibilant, another principle comes into play: labials prefer to occur first, whereas dorsals prefer not to. As a matter of fact, PT is much more frequent than TP, and PK occurs more often than KP. Even among the SO clusters, SK, where the second consonant is a dorsal, is more frequent than SP, where a labial occurs post-consonantly. This tendency is particularly interesting if compared to the acquisition of different places of articulation in infants. Fikkert and Levelt (2008) bring evidence that Dutch children pass through a phase during which in all the C₁VC₂ words they produce, C₁ is labial and C₂ is dorsal, so that a word like *kip* ‘chicken’ is pronounced as [pik] (cf. Ingram 1974 for similar findings in English). Another notable tendency is the avoidance of a sequence of consonants with the same place of articulation (traditionally formalized as OCP – Obligatory Contour Principle, Leben 1973) or a sequence of non-coronal consonants. As a matter of fact, labials and dorsals tend not to combine. I will make use of the feature [+grave] to unify labials and dorsals (Jakobson 1971[1938]: 274).⁵ All these preferences can be tentatively formalized using the OT constraints proposed in (6).

- (6) OCP[+GRAVE] Tautosyllabic [+grave] consonants are disallowed.
 OCP[+CONT] Tautosyllabic [+continuant] consonants are disallowed.
 OCP[–CONT] Tautosyllabic [–continuant] consonants are disallowed.
 OCP[PLACE] Tautosyllabic homorganic consonants are disallowed (Ussishkin 1999).
 *TO A stop cannot be followed by another obstruent in the same syllable (Morelli 1999: 48).
 [LABIAL] A labial consonant must appear syllable-initially (Fikkert and Levelt 2008).
 *[DORSAL] Dorsal consonants are banned from syllable-initial position (Fikkert and Levelt 2008).

⁵ Labials and dorsals are treated as a natural class also by some practitioners of Element Theory, such as Backley (2011). In his analysis, both labials and dorsals are characterized by the element [U], which plays the role of the head in the former and of the operator in the latter.

Tableau 4. Markedness of obstruent clusters.

	OCP [+grave]	OCP [+cont]	*TO	OCP [-cont]	[LABIAL]	*[DORSAL]	OCP [place]
(a) /sk/							
(b) /st/							*
(c) /sp/					*		
(d) /pt/			*	*			
(e) /ps/			*				
(f) /tk/			*	*			
(g) /kt/			*	*		*	
(h) /pk/	*		*	*			
(i) /sf/		*			*		
(j) /kp/	*		*	*	*	*	
(k) /ts/			*				*
(l) /xp/	*				*	*	
(m) /tp/			*	*	*		

Tableau 4 is a peculiar type of tableau displaying a constraint ranking that might explain the data in (5), without any particular claim on whether a language with such a ranking actually exists. Nonetheless, Tableau 4 captures the constraint interaction yielding the frequency of certain obstruent clusters and the relative rarity of others. It is shown that, among SO obstruents, /sk/ does not incur any violation, /st/ only violates the lowest-ranked constraint OCP[PLACE], whereas /sp/ violates [LABIAL]. It is also shown that SO clusters are better formed than both TT and FF clusters, since the latter two violate OCP[-CONT] and OCP[+CONT], respectively. Tableau 4 takes into consideration only one cluster of the FT type, i.e., /xp/. These clusters appear to be much rarer and that is, at least partially, probably due to the lower acoustic salience of /f, θ, x/ as opposed to sibilants. My guess is that non-sibilant fricatives are assigned lower strength values in the underlying representation and are therefore more likely to disappear over time. For instance, the results of

an experiment consisting in the identification of plateau clusters in noise revealed that /s/ is recognized correctly almost 100% of the time, whereas /f/ is very easily unheard or misheard, to a greater extent than stops (Baroni 2014). Other studies proved that dorsal fricatives, such as /x, ɣ/, have a similar degree of intensity as labial fricatives, and both are much weaker than sibilants (Gordon et al. 2002). According to the phonetic power scale proposed in Fletcher (1972: 82), /θ/ is the acoustically weakest of the English sounds, followed by /f/. The greater suitability of sibilants to occur in clusters can also be formalized positing the following constraint ranking: *FO, *TO > *SO, which basically states that it is worse for a non-sibilant fricative (F) and a stop (T) to precede another obstruent, than it is for a sibilant.

Tableau 5. /s/ as the best #O₁.

	*FO	*TO	*SO
a) /st/			*
b) /ft/	*!		
c) /pt/		*!	

As shown earlier, among languages that allow word-initial consonant clusters, all allow clusters of rising sonority (#OR). Among those that allow plateau clusters, all tolerate SO clusters. Some, but not all, tolerate other types of obstruent clusters but prefer those that begin with a labial or at least, where the second consonant is not labial. Very few languages allow clusters of reverse sonority (#RO). Whether a language allows consonant clusters or not and which type of clusters it allows depends on the interaction of MAX constraints and markedness constraints such as SSG and LABIALFIRST. For instance, imagine a series of consonant clusters that may have potentially arisen in casual speech: SO, OS, PT, TP, OR, RO. In a language like Czech (see 7a), faithfulness constraints dominate phonotactic constraints, and as a result, Czech allows for all these types of clusters, included those implying a sonority reversal. There is one notable exception: Czech has /#tsp/ (e.g., *cpát* ‘stuff’) but not /#tp/ or any other type of stop + labial cluster, indicating that [LABIAL must have been ranked higher than faithfulness but lower than MAX[+STRID] and LINEARITY (“The sequencing of output segments must be

the same as the input” (McCarthy and Prince 1995: 123)), since the affricate /ts/ can be described as a strident stop (La Charité 1993; Hall 2004). If a language has a phonotactics that typically follows the SSG but tolerates SO, OS and PT clusters, such as Italian, MAX[+STRID] and [LABIAL must be ranked higher than SSG, MAX and DEP⁶ (see 7b). English only tolerates OR and SO but, for example, its adaptation of Greek loanwords *psychology*, *pterodactyl* and *xylophone* (pronounced, respectively, [sar'kɒlədʒi], [ˌtɛɪə'dæktɪl] and ['zɑɪləfəʊn]) shows that *TO must be ranked higher than faithfulness constraints protecting stops (see 7c). Spanish only allows #OR, therefore SSG must be undominated. However, a ranking such as MAX[+STRID] > MAX > DEP triggers the insertion of an epenthetic [e] before SO clusters, such as in *estar* ‘to be’ from Latin *stare* (see 7d). Finally, there are languages like Finnish where not only are tautosyllabic SO clusters banned, but loanwords containing these sequences are simplified by omitting /s/, e.g., Swedish *spole* ‘spool’ is adapted in Finnish as *puola* (Luthy 1973: 6–7). In Finnish, therefore, SSG must dominate both MAX and MAX[+STRID] (see 7e). As Smith (2009: 155) points out, in Finnish (as well as in Hmong and Sranan), highly salient segments undergo deletion in loanwords adaptation, even though there is evidence that they are perceived. These deletion processes must therefore form part of the phonological grammar and are not based on (mis)perception⁷.

(7) Consonant clusters phonotactics

(a) Czech.

Input: SO, OS, PT, TP, OR, RO → MAX[+STRID], LINEARITY > [LABIAL > FAITH > SSG → SO, OS, PT, OR, RO but *TP.

(b) Italian.

SO, OS, PT, TP, OR, RO → MAX[+STRID], [LABIAL > SSG > FAITH → SO, OS, PT, OR, but *RO, *TP

⁶ DEP is the classic OT constraint against epenthesis (Prince and Smolensky 1993).

⁷ Similarly, the debuccalization and deletion of /s/ in several languages (such as Ancient Greek, Caribbean Spanish, Old French) is arguably due to articulatory difficulty rather than to perceptual factors.

(c) English.

SO, OS, PT, TP, OR, RO → MAX[+STRID], *TO > SSG > FAITH → SO, OR but *OS, *PT, *RO, *TP.

(d) Spanish.

SO, OS, PT, TP, OR, RO → MAX[+STRID] > SSG > FAITH → OR, but *SO, *OS, *PT, *TP, *RO, and SO → [e]SO.

(e) Finnish: SO, OS, PT, TP, OR, RO → SSG > MAX[+STRID], MAX > FAITH → OR, but *SO, *OS, *PT, *TP, *RO.

5. Conclusion

While it is uncontroversial that the basic syllable type is CV, at the low phonetic level there is great variation as for the type of syllables that are actually pronounced. Even languages with a simple syllabic structure admit complex consonant clusters in casual speech and these clusters do not always respect the SSG. As the /sinistra/ example has shown, casual speech phenomena are driven by both articulatory ease and the urge to preserve prominent phonemes and features, so that sibilants are more likely than other obstruents to be preserved in heavily reduced variants. The hypothesis proposed in this paper consists in the idea that frequent and/or salient segments and features form part of the invariant of words, i.e., the phonetic core of underlying representations that cannot be deleted, not even in heavily reduced speech. As the data in (5) suggest, when sonority is not at play, as in plateau clusters, the more salient a segment, the more likely it is preserved in pre-consonantal position, i.e., not directly adjacent to a vowel. As a result, #SO clusters are the preferred word-initial obstruent cluster type crosslinguistically. Although, for the moment, the proposals put forward here remain speculative and other experimental studies would be needed, investigating the relationship between casual speech processes and phonotactics appears to be promising and will hopefully expand our knowledge on the phonology of consonant clusters.

6. References

- Alderete, J. 1997. "Dissimilation as local conjunction". In: Kusumoto, K. (ed.), *Proceedings of North East Linguistic Society 27*. Amherst, MA: Graduate Linguistic Student Association. 17–31.
- Backley, P. 2011. *An introduction to Element Theory*. Edinburgh: Edinburgh University Press.
- Baković, E. 2005. "Antigeminatio, assimilation, and the determination of identity". *Phonology* 22(3). 279–315.
- Barker, M.A.R. 1964. *Klamath grammar*. Berkeley & Los Angeles, CA: University of California Press.
- Baroni, A. 2014. "On the importance of being noticed: the role of acoustic salience in phonotactics and casual speech". In: Calderone, B. and C. Celata (eds.), *Theory and evidence in the study of (mor)phonotactics*. (Special issue of *Language Sciences* 46(A). 18–36.)
- Baroni, A. 2014. *The invariant in phonology. The role of salience and predictability*. In preparation.
- Baroni, A. and M. Simonović. 2014. "A perception-based account of variation. Phonetics, phonology and the invariant". In: Cyran, E. and J. Szpyra-Kozłowska (eds.), *Crossing phonetics-phonology lines*. Cambridge: Cambridge Scholars Publishing. 115–140.
- Bertinetto, P.M. and M. Loporcaro. 2005. "The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome". *Journal of the International Phonetic Association* 35(2). 131–151.
- Blaho, S. 2008. *The syntax of phonology. A radically substance-free approach*. (PhD dissertation, University of Tromsø.)
- Blevins, J. and A. Garrett. 2004. "The evolution of metathesis". In: Hayes, B., R. Kirchner and D. Steriade (eds.), *Phonetically based phonology*. Cambridge, MA: Cambridge University Press. 117–156.
- Blevins, J. 2004. *Evolutionary phonology. The emergence of sound patterns*. New York: Cambridge University Press.
- Boersma, P., J. Dekkers and J. van der Weijer. 2000. "Optimality-theoretic phonology". In: Boersma, P., J. Dekkers and J. van der Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*. Oxford: Oxford University Press. 4–8.
- Bürki, A., M. Ernestus and U.H. Frauenfelder. 2010. "Is there only one 'fenêtre' in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words". *Journal of Memory and Language* 62. 421–437.
- Butskhrikidze, M. 2002. *The consonant phonotactics of Georgian*. Utrecht: LOT.
- Canello, U.A. 1872. "Il vocalismo tonico italiano". *Rivista di Filologia Romanza* 1. 209–225.
- Clements, G.N. 1990. "The role of the sonority cycle in core syllabification". In: Kingston, J. and M. Beckman (eds.), *Papers in laboratory phonology I*. Cambridge, MA: Cambridge University Press. 283–333.

- Cole, R.A., M. Coltheart and F. Allard. 1974. "Memory of a speaker's voice: Reaction time to same-or-different-voiced letters". *The Quarterly Journal of Experimental Psychology* 26. 1–7.
- Collins, J. 2013. Modal-dependence and naturalness in phonology: Confronting the ontogenetic question. (MA thesis, University of Tromsø.)
- Comrie, B. 1981. *The languages of the Soviet Union*. Cambridge, MA: Cambridge University Press.
- Crawford, J.M. 1973. "Yuchi phonology". *International Journal of American Linguistics* 39(3). 173–179.
- Dalby, J. 1986. *Phonetic structure of fast speech in American English*. Bloomington, IN: Indiana University Linguistics Club.
- Dressler, W.U. and K. Dziubalska-Kołaczyk. 2006. "Proposing morphonotactics". *Italian Journal of Linguistics/Rivista di Linguistica* 18(2). 249–266.
- Dressler, W.U., K. Dziubalska-Kołaczyk and L. Pestal. 2010. "Change and variation in morphonotactics". *Folia Linguistica Historica* 31(1). 51–67.
- Dziubalska-Kołaczyk, K., P. Zydorowicz and M. Jankowski. 2013. "English morphonotactics: A corpus study". *The Phonetician* 107/108. 53–67.
- Dziubalska-Kołaczyk, K. and D. Zielińska. 2011. "Universal phonotactic and morphonotactic preferences in second language acquisition". In: Dziubalska-Kołaczyk, K., M. Wrembel and M. Kul (eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010*. Frankfurt Am Main: Peter Lang. 53–63.
- Ernestus, M. 2013. "Acoustic reduction and the roles of abstractions and exemplars in speech processing". *Lingua* 142. 27–41.
- Ernestus, M. 2000. *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology-phonetics interface*. Utrecht: LOT
- Fikkert, P. and C. Levelt. 2008. "How does Place fall into place? The lexicon and emergent constraints in children's developing phonological grammar". In: Avery, P., B. Elan Dresher and K. Rice (eds.), *Contrast in phonology. Theory, perception and acquisition*. Berlin: Mouton de Gruyter. 231–268.
- Flemming, E. 2002. *Auditory representations in phonology*. New York: Routledge.
- Fletcher, H. 1972. *Speech and hearing in communication*. Huntington, NY: Robert E. Krieger.
- Gibson, L.F. 1956. "Pame (Otomi) phonemics and morphophonemics". *International Journal of American Linguistics* 22. 242–265.
- Goldinger, S.D. 1998. "Echoes of echoes? An episodic theory of lexical access". *Psychological Review* 105. 251–279.
- Goldinger, S.D. 2007. "A complementary-systems approach to abstract and episodic speech perception". *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken. 49–54.
- Gordon, M., P. Barthmaier and K. Sands 2002. "A cross-linguistic acoustic study of fricatives". *Journal of the International Phonetic Association* 32. 141–174.
- Gussman, E. 2007. *The phonology of Polish*. Oxford: Oxford University Press.
- Hale, M. and C. Reiss. 2000. "Phonology as cognition". In: Burton-Roberts, N., P. Carr and G. Docherty (eds.), *Phonological knowledge. Conceptual and empirical issues*. Oxford: Oxford University Press. 161–184.

- Hall, T.A. 1992. "Syllable structure and syllable-related processes in German". *Linguistische Arbeiten* 276.
- Hall, T.A. 2004. "Assibilation in Modern German". *Lingua* 114. 1035–1062.
- Halle, M. 1959. *The sound pattern of Russian*. The Hague: Mouton.
- Harris, J. 1990. "Segmental complexity and phonological government". *Phonology* 7. 255–300.
- Harris, J. 2006. "The phonology of being understood: Further arguments against sonority". *Lingua* 116. 1483–1494.
- Hayes, B., R. Kirchner and D. Steriade (eds.). 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.
- Henderson, E.J.A. 1976. "Khasi initial clusters". In: Jenner, P.N., L.C. Thompson and S. Starosta (eds.), *Austroasiatic studies*. Honolulu: The University Press of Hawaii. 523–538.
- Hinskens, F. 2011. "Lexicon, phonology and phonetics. Or: Rule-based and usage-based approaches to phonological variation". In: Siemund, P. (ed.), *Linguistic universals and language variation*. Berlin: Mouton de Gruyter. 416–456.
- Hodge, C.T. 1946. "Phonemes of Serbo-Croatian". *Language* 22. 112–120.
- Hoffman, C. 1963. *A grammar of the Margi language*. London: Oxford University Press for International African Institute.
- Hongkai, S. 1986. *Notes on Tibeto-Burman consonant clusters. Linguistics of the Tibeto-Burman area*. Beijing: Chinese Academy of Seoul Sciences.
- Ingram, D. 1974. "Fronting in child phonology". *Journal of Child Language* 1. 233–241.
- Jacques, G. 2004. "The laterals in Tibetan". Talk presented at the 10th Himalayan Languages Symposium, Thimphu, Bhutan, 1–3 December.
- Jaeger, J.J. and R. van Valin Jr. 1982. "Initial consonant clusters in Yateé Zapotec". *International Journal of American Linguistics* 48(2). 125–138.
- Jakobson, R. 1971 [1938]. "Observations sur le classement phonologique de consonnes". In: Rudy, S. (ed.), *Selected writings 1: Phonological studies*. The Hague: Mouton. 272–279.
- Kager, R. 1999. *Optimality Theory*. Cambridge, MA: Cambridge University Press.
- Kaye, J.D. 1990. "Government in phonology: The case of Moroccan Arabic". *The Linguistic Review* 6. 131–159.
- Kaye, J., J. Lowenstamm and J.-R. Vergnaud. 1985. "The internal structure of phonological representations: a theory of Charm and Government". *Phonology Yearbook* 2. 305–328.
- Kenstowicz, M. 1994. *Phonology in Generative Grammar*. Cambridge, MA & Oxford: Blackwell.
- Kohler, K.J. 1990. "Segmental reduction in connected speech in German: Phonological facts and phonetic explanations". In: Hardcastle, W.J. and A. Marchal (eds.), *Speech production and speech modelling*. Dordrecht: Kluwer Academic Publishers. 69–92.
- Krämer, M. 2009. *The phonology of Italian*. Oxford: Oxford University Press.
- La Charité, D. 1993. The internal structure of affricates. (PhD dissertation, University of Ottawa.)

- Ladefoged, P. and I. Maddieson 1996. *The sounds of the world's languages*. Cambridge: Blackwell.
- Ladefoged, P. 1964. *A phonetic study of West African Languages*. London: Cambridge University Press.
- Lawrence, E. 1977. *Haida dictionary*. Fairbanks: Alaska Native Language Center.
- Leben, W.R. 1973. Suprasegmental phonology. (PhD dissertation, MIT.)
- Levelt, C., N. Schiller and W. Levelt. 1999. "The acquisition of syllable types". *Language Acquisition* 8(3). 237–264.
- Lock, A. 2007. *Phonology essentials – Abau language*. Ukarumpa, Papua New Guinea: SIL International.
- Luthy, M.J. 1973. *Phonological and lexical aspects of colloquial Finnish*. (Uralic and Altaic series 119.) Bloomington, IN: Indiana University.
- MacKay, C.J. 1994. "A sketch of Misantra Totonac phonology". *International Journal of American Linguistics* 60(4). 369–419.
- MacNeilage, P.F. and B.L. Davis. 2000. "Origin of the internal structure of word forms". *Science* 288. 527–531.
- Maddieson, I. 2011. "Consonant-vowel ratio". In: Dryer, M.S. and M. Haspelmath (eds.), *The world atlas of language structure online*. Munich: Max Planck Digital Library, chapter 3. Available online at <<http://wals.info/chapter/3>>.
- Marlett, S.A. and V.B. Pickett 1987. "The syllable structure and aspect morphology of Isthmus Zapotec". *International Journal of American Linguistics*. 53(4). 398–422.
- Marlett, S.A. 1988. "The syllable structure of Seri". *International Journal of American Linguistics*. 54(3). 245–278.
- Mattys, S.L. and J.M. Liss. 2008. "On building models of spoken word recognition: when there is as much to learn from natural 'oddities' as artificial normality". *Perception and Psychophysics* 70. 1235–1242.
- McCarthy, J. and A. Prince 1993. *Prosodic morphology I: Constraint interaction and satisfaction*. (Technical Report #3, Rutgers University Center for Cognitive Science.)
- McCarthy, J. and A. Prince 1995. "Faithfulness and reduplicative identity". *University of Massachusetts Occasional Papers in Linguistics* 18. 249–384.
- McCarthy, J. 2000. "The prosody of phrase in Rotuman". *Natural Language and Linguistic Theory* 18. 147–197.
- McLennan, C.T. and P.A. Luce. 2005. "Examining the time course of indexical specificity effects in spoken word recognition". *Journal of Experimental Psychology: Learning, Memory and Cognition* 31. 306–321.
- McLennan, C.T., P.A. Luce and J. Charles-Luce. 2003. "Representation of lexical form". *Journal of Experimental Psychology: Learning, Memory and Cognition* 29. 539–553.
- Morelli, F. 1999. The phonotactics and phonology of obstruent clusters in Optimality Theory. (PhD dissertation, University of Maryland, College Park, MD.)
- Nacaskul, K. 1978. "The syllabic and morphological structure of Cambodian words". *Mon-Khmer Studies* 7. 183–200.
- Nagamma R.K. 1987. "Constraints on consonant sequences across some Indian Languages: A typological view". *Osmania Papers in Linguistics* 13. 39–57.

- Nespor, M. 1993. *Fonologia*. Bologna: Il Mulino.
- Newmark, L. 1998. *Albanian-English dictionary*. Oxford: Oxford University Press.
- Oakes, P.J. 2001. "A description of Teribe phonology". *SIL Electronic Working Papers*.
- Ohala, J.J. 1997. "Emergent stops". *Proceedings of the 4th Seoul Conference on Linguistics (SICOL)*. 84–91.
- Ohala, J.J. and H. Kawasaki-Fukumori. 1997. "Alternatives to the sonority hierarchy for explaining segmental sequential constraints". In: Eliasson, S. and E.H. Jahr (eds.), *Language and its ecology: Essays in memory of Einar Haugen*. (*Trends in Linguistics. Studies & Monographs* 100.) Berlin: Mouton de Gruyter. 343–365.
- Oostendorp, M. van. 1997. "Style levels in conflict resolution". In: Hiskens, F. and R. van Hout (eds.), *Variation, change and phonological theory*. John Benjamins: Amsterdam. 207–230.
- Parker, S. 2012. Onset consonant clusters database. (Unpublished manuscript. Dallas, TX: Graduate Institute of Applied Linguistics and SIL International.)
- Penny, R.J. 2004. *A history of the Spanish language*. Cambridge, UK: Cambridge University Press.
- Penzl, H. 1955. *A grammar of Pashto*. Washington, DC: American Council of Learned Societies.
- Pharao, N. 2009. Consonant reduction in Copenhagen Danish: A study of linguistic and extra-linguistic factors in phonetic variation and change. (PhD dissertation, University of Copenhagen.)
- Pierrehumbert, J.B. 2002. "Word-specific phonetics". In: Gussenhoven, C. and N. Warner (eds.), *Laboratory phonology VII*. Berlin: Mouton de Gruyter. 101–140.
- Prince, A. and P. Smolensky. 1993/2004. *Optimality Theory: Constraints interaction in generative grammar*. Malden, MA: Blackwell. (Originally distributed as Rutgers University Center of Cognitive Science Technical Report #2.)
- Redden, J.E. 1965. Walapai phonology and morphology. (PhD dissertation, Indiana University.)
- Regnier, S. 1993. "Quiégolani Zapotec phonology". *Work Papers of the Summer Institute of Linguistics* 37, University of North Dakota.
- Rood, D.S. 1975. "The implications of Wichita phonology". *Language* 51(2). 315–337.
- Rose, S. 2000. "Rethinking geminates, long-distance geminates and the OCP". *Linguistic Inquiry* 31(1). 85–122.
- Sanders, E.K. 1973. "When are speech sounds learned?" *Journal of Speech and Hearing Disorders* 37. 55–63.
- Savy, R. and F. Cutugno. 2009. "CLIPS. Diatopic, diamesic and diaphasic variations in spoken Italian". In: Mahlberg, M., V. González-Díaz and C. Smith (eds.), *Proceedings of the 5th Corpus Linguistics Conference*, article #213.
- Schutter, G. de 1994. "Dutch". In: Koenig, E. and J. van der Auwera (eds.), *The Germanic languages*. London & New York, NY: Routledge. 439–477.
- Schwartz, G. 2009. *Phonology for the listener and language learner*. Poznań: Wydawnictwo Naukowe UAM.
- Schwartz, G. 2013. "A representational parameter for onsetless syllables". *Journal of Linguistics* 49(3). 613–646.

- Selkirk, E.O. 1984. "On the major class features and syllable theory". In: Aronoff, M. and R. Oehrle (eds), *Language sound structure*. Cambridge, MA: MIT Press. 107–136.
- Shockey, L. 2003. *Sound patterns of spoken English*. Oxford: Blackwell.
- Smith, J.L. 2009. "Source similarity in loanword adaptation: Correspondence Theory and the posited source-language representation". In: Parker, S. (ed.), *Phonological argumentation: Essays on evidence and motivation*. London: Equinox. 155–177.
- Smolensky, P. 1995. "On the internal structure of the constraint component Con of UG". Handout, UCLA, 7 April. ROA #86-0000.
- Steriade, D. 1982. Greek prosodies and the nature of syllabification. (PhD dissertation, MIT.)
- Steriade, D. 2001/2008. "The phonology of perceptibility effects: The P-map and its consequences for constraint organization". In: Hanson, K. and S. Inkelas (eds.), *The nature of the word*. Cambridge, MA: MIT Press. 151–179. (Originally circulated in 2001.)
- Sterner, J.K. 1975. "Sobei phonology". *Oceanic Linguistics* 14. 146–167.
- Swanton, J.R. 1910. *Haida, an Illustrative Sketch*. Bureau of American Phonology, Bulletin 40, pt I. 205–282.
- Tankevičiūtė, M. and M. Strimaitienė. 1990. "Initial consonant clusters in Prussian". *Baltistica* 26(2). 105–110.
- Tarpen, M.L. 1989. A grammar of the Nisgha language. (PhD dissertation, University of Victoria.)
- Ullrich, J. 2008. *New Lakota dictionary & incorporating the Dakota dialects of Santee-Sisseton and Yankton-Yanktonai*. Lakota Language Consortium.
- Ussishkin, A. 1999. "Head dominance in Modern Hebrew prosodic morphology". In: Ussishkin, A., D. Herrick, K. Kurisu and N. Sanders (eds.), *Phonology at Santa Cruz* 6. Santa Cruz, CA: University of Santa Cruz. 71–82.
- Warner, N. 2002. "The phonology of epenthetic stops: implications for the phonetics-phonology interface in Optimality Theory". *Linguistics* 40. 1–27.
- Wright, R. 1996. Consonant clusters and cue preservation in Tsou. (PhD dissertation, UCLA.)

Address for correspondence

Antonio Baroni
DISLL
Università degli Studi di Padova
via Beato Pellegrino 1
35100 Padua
Italy
antonio.baroni@studenti.unipd.it