The laryngeal properties of Slovak three-consonant clusters

THE LARYNGEAL PROPERTIES OF SLOVAK THREE-CONSONANT CLUSTERS

Zsuzsanna Bárkányi – Zoltán G. Kiss

Introduction

In this paper,[1] we study the phonetic properties of three-consonant clusters (CC#C) in Slovak. More precisely, we will investigate the laryngeal properties of the velar–alveolar stop clusters /kt/ and /gd/, and the alveolar fricative–stop clusters /st/ and /zd/ in word-final position when followed by a voiced or a voiceless obstruent, or a sonorant consonant. This topic is of interest for two reasons: (i) there are not many studies dealing with the laryngeal characteristics of three-consonant clusters, and (ii) the study of consonant clusters can shed further light on the issue whether or not regressive voicing assimilation (RVA) in general, including pre-sonorant voicing, is a neutralizing process in Slovak.

We will seek to answer the following research questions: (i) Is voicing assimilation in Slovak (with obstruent and sonorant consonants as triggers) neutralizing or incomplete? (ii) Does pre-sonorant voicing in Slovak differ from pre-obstruent voicing? (iii) Is the word-final devoicing of obstruent clusters a completely neutralizing process in Slovak, or is it an example of incomplete laryngeal neutralization?

Background

Neutralization

Understood in its most well-known sense, phonological neutralization refers to the case when two or more contrastive sound segments suspend their contrast under specific conditions, whereby only a limited set of the contrastive segments can occur in a particular position. Examples include vowel reduction in English (where only certain vowels may appear in an unstressed syllable) and laryngeal neutralization (see Silverman 2012 for a detailed discussion of neutralization, as well as Jansen 2004). Neutralization processes are considered to be *complete* when there is no difference between the underlyingly contrasting members either in production or in perception for any of the possible phonetic correlates of a given contrast in a given context. That is, forms (e.g., voiced and voiceless obstruents) that are distinguishable in certain contexts (e.g., in intervocalic position) are phonetically completely

[1]We would like to thank the two anonymous reviewers for their valuable suggestions.

indistinguishable in the neutralizing contexts (e.g., word-finally or in pre-obstruent position). Thus, in such positions a devoiced segment cannot be distinguished from an underlyingly voiceless segment either in its phonetic properties or in its phonological behaviour in any way.

Neutralization interpreted this way, that is, true phonetic neutralization, rarely results in homophones though. Firstly, it has been observed (e.g. Charles-Luce 1993; Kaplan 2011) that voicing alternation is more likely to be nearly neutralized – as opposed to completely neutralized – in contexts that would otherwise be semantically ambiguous. That is, phonological patterns are sensitive not only to contrasts among segments, but also to contrasts among individual lexical items.

Secondly, even if a neutralization process does derive homophony, it will rarely be the case that there is semantic ambiguity because languages resort to other strategies – especially in diachronic terms – to avoid homophony. Silverman (2012) discusses Korean, a language that has numerous neutralizing alternations but where the amount of homophony resulting from these alternations is surprisingly low. Korean counterbalanced the attrition of root-final consonantal values by resorting to root compounding. We can still assume though that processes such as voicing assimilation can be completely neutralizing phonetically.

Theoretically, there are at least three facets of phonetic neutralization (which themselves could be classified even further, see Dinnsen 1985, for instance). First, in the case of *complete neutralization,* there is no difference between the underlyingly contrasting members either in production or in perception for any of the possible phonetic correlates of the given contrast. This is the traditional generative assumption of neutralization, which – with the advent of more and more experimental work – turns out to be rarely the case.

Another possibility is that there is some systematic acoustic and/or articulatory difference between the segments in question, but this difference is *not perceived,* or at least speakers are not aware of the contrast. Allophonic differences and the first stages of sound change typically belong to this group (Dinnsen 1985). A subclass of this group is when there is a slight articulatory difference which does not manifest itself acoustically. Beňuš and Gafos (2007), using a combination of magnetometry and ultrasound, found that Hungarian transparent vowels that trigger back harmony (*híd* 'bridge') showed a more retracted tongue body posture than phonemically identical vowels that trigger front harmony (*víz* 'water') even in isolation (that is, not in a suffixed form, which would be a simple coarticulatory phenomenon). Note that no acoustic or perception study so far has shown any differences between them.

A further possibility is that the members of a "neutralized" contrast are not identical after all. Some production-acoustic features might remain that are consistently and significantly different in the contrasting sounds, and which are perceived by speakers. Processes belonging to this group can be quite varied again: contrast preservation despite the loss of a primary acoustic cue might be fairly robust in some cases, while very weak in others (see Steriade's p-map theory, Steriade 2008).

The notions *neutralization, categoricality* and *graduality* are closely connected. A variation is generally thought to be *categorical* if it can be described with the categorical values of phonological features, i.e., when an alternation occurs between two discrete categories (e.g., voiced and voiceless) with no intermediate values. An alternation is thought to be *gradient* if the acoustic characteristics of the variants reflect values in between these categories (e.g., partly voiced), even if these in-between categories are systematic. Partial neutralization is gradient according to this view since some acoustic characteristics might signal more voicing (e.g. vowel length), whiles others might signal less voicing (e.g. phonation itself), for instance.

The phonological context regarding voicing neutralization studied in this paper is the word-final position. It has been reported for German (Port et al. 1981; O'Dell–Port 1983; Charles-Luce 1985), Catalan (Dinnsen–Charles-Luce 1984; Charles-Luce 1993), and Polish (Slowiaczek–Dinnsen 1985; Slowiazcek–Szymanska 1989) that word-final laryngeal neutralization leaves some residual cues to the phonological voicing of obstruents. However, Fourakis and Iverson (1984) and Kahlen-Halstenbach (1990) found that word-final devocing is phonetically complete in German. Jassem and Richter (1989) report the same for Polish. Experimental evidence concerning voicing assimilation is varied. There is experimental work demonstrating that regressive voicing assimilation is non-neutralizing, and therefore it is a low-level, phonetic process (e.g., Charles-Luce 1993 on Catalan and Burton–Robblee 1997 on Russian). In contrast, Hallé and Adda-Decker (2011) found that whenever it occurs, voicing assimilation is categorical in French. Strycharczuk and Simon (2013) claim the same about West-Flemish. The issue of complete vs. incomplete laryngeal neutralization is far from being settled either empirically or theoretically. And there are very few experimentally-based studies that deal with pre-sonorant voicing.

Pre-sonorant voicing

Pre-sonorant voicing is a type of regressive voicing assimilation whereby a word-final voiceless/devoiced obstruent is assimilated in voicing to a following sonorant consonant or vowel in the next word. This process has raised recurrent interest among phonologists mostly due to the fact that the apparent trigger of voicing assimilation is a segment which is not contrastively specified for voicing. Phonetically considered, sonorants may be suitable triggers of regressive voicing assimilation as they are phonetically voiced and rather resistant to devoicing. Yet, typologically, pre-sonorant voicing is much less frequent than pre-obstruent voicing. In phonetically-based models this is explained by the passive or modal phonation of sonorants as opposed to the active voicing of voiced obstruents (see especially Jansen 2004 and the references therein).

There are some interesting restrictions that seem to apply to pre-sonorant voicing, which do not apply to "regular", pre-obstruent voicing assimilation: it typically occurs in languages which display final devoicing (this, however, does not mean that in all languages with word-final devoicing we will find pre-sonorant voicing as well). Pre-sonorant voicing is also generally restricted to the word-final (or syllable-final) position. Slovak is a language displaying both pre-obstruent and pre-sonorant voicing assimilation, as reported by Pauliny (1979) and Rubach (1994). In Slovak, a word-final obstruent is realized voiced if it is followed by a voiced obstruent (1c), or by a sonorant consonant or a vowel in the next word (1a). The latter process also applies to clusters, but is not operative within the word (1b).

(1) Voicing assimilation and pre-sonorant voicing in Slovak (Pauliny 1979: 152–153)

a. pre-sonorant voicing across word-boundary
vták letí [fta:g lɛci:] 'bird is flying'
chlap ani nejedol [xlab aɲi ɲɛjɛdol] 'man didn't even eat'
jesť a piť [jezɟ a pic] 'eat and drink'

b. no pre-sonorant voicing within the word
tma [tma] 'darkness'
kladivo [klaɟiʋo] 'hammer'
astma [astma] 'asthma'
chlap-mi [xlapmi] 'man-INSTR'

c. regressive VA among obstruents
kto [kto] 'who', *gde* [gɟɛ] 'where'
chlap dochodí [xlab doxoɟi:] 'man comes'
hrad pri [ɦrat pri] 'castle next to'

d. final devoicing
plod [plot] 'fruit'
plot [plot] 'fence'

As far as the trigger of pre-sonorant voicing is concerned, significant variation is observed among languages. In some languages – like Slovak, shown in (1), Kraków Polish (Rubach 1996) or West-Flemish (Strycharczuk–Simon 2013) – sonorant consonants and vowels pattern together and induce voicing assimilation. West-Flemish differs from the other Southern Dutch dialects in that in those dialects, as reported by de Schutter and Taeldeman (1986), only vowels voice the final fricative of the preceding word, while in West-Flemish, fricatives are voiced before sonorant consonants as well across word-boundaries: *zes jaar* [zɛz jaːr] 'six years'. Similarly to the Southern Dutch dialects, /s/-voicing in Ecuadorian Spanish is also induced only by vowels. Standard Peninsular Spanish is exactly the other way round: /s/ is voiced when followed by a voiced obstruent or a sonorant consonant. The process is not limited to word-final position, syllable-final /s/ also undergoes voicing (Hualde 2005).

It has been reported in a number of studies that pre-sonorant voicing targets only subclasses of obstruents: in Dutch only fricatives undergo voicing assimilation induced by sonorants (Simon 2010), in Spanish only /s/. An illustrative example is provided by Jiménez–Lloret (2008), who report a dialect continuum in Catalan: in Central Valencian there is no voicing of word-final consonants before vowels, Alguerés and the Valencian dialect of la Costera have sibilant voicing, in the Valencian dialect of Palmera – apart from word-final sibilants – alveolar affricates also become voiced in pre-vocalic position, Central Catalan has variable /f/ voicing as well, while in Alicantino all word-final obstruents undergo voicing when followed by a vowel.

Three-consonant clusters

As we have mentioned in the introduction, there are not many studies dealing with the laryngeal properties of three-consonant clusters. Here we briefly cite a few studies that discuss the focus of our investigation, namely voicing assimilation in three-member clusters.

Central Catalan shows an intriguing asymmetrical system: pre-vocalic voicing affects word- and prefix-final sibilants and stop + sibilant clusters to the exclusion of singleton stops (Bonet–Lloret 1998; Wheeler 2005; Strycharczuk 2012). As far as sonorant consonants are concerned, they affect all obstruents equally. This type of "undergoer asymmetry" is problematic for any phonetically and/or functionally-based explanation. There are several competing hypotheses as to why fricative voicing may be preferred over stop voicing before sonorants but none of them can straightforwardly account for the question why stop + fricative clusters undergo voicing while singleton stops do not. Note that any output-oriented rule or constraint-based formal analysis can easily account for this pattern. (A vowel will voice the word-final sibilant, which then will voice the preceding stop.) However, the general issue of pre-sonorant voicing remains a problem for these models, too.

Strycharczuk (2012) analyzes sibilant voicing using a diachronic phonetic-functional model. According to her, the process originated as intervocalic sibilant voicing rather than being a pre-vocalic voicing process. She claims that the pattern started off as intervocalic voicing that targeted delaryngealized sibilants. These neutralized final obstruents are less likely to resist voicing spill over from the neighbouring sounds, as no active devoicing gesture is executed to counteract voicing. Passive voicing may also be less perceivable in stops that in sibilants. The next diachronic step was that listeners re-interpreted intervocalic sibilant voicing as pre-vocalic. The final stage in the development of the Catalan pre-vocalic voicing involved rule telescoping (Hyman 1975), when a voiced pre-vocalic sibilant becomes an input to VA, which operates independently in the language. In the case of Catalan not only the undergoer asymmetry is puzzling, but the trigger asymmetry as well: vowels only voice sibilants and sibilant-final clusters (like [ps], [ks]) while sonorant consonants cause gradient voicing in all obstruents with significant inter- and intra-speaker variation. If we assume that the right-hand environment does have an effect on the duration of passive voicing, but what is essential is the presence or absence of a voicing target (i.e., a delaryngealized final obstruent), we do not expect any differences between sonorant consonants and vowels as triggers of pre-sonorant voicing. Strycharczuk (2012) hypothesizes that pre-vocalic voicing is the older pattern of the two in Catalan, which is supported by the high amount of variation observed in the pre-sonorant consonant process as well as the assumption that vowels are more conductive of passive voicing as they are more open.

Recasens and Mira (2013) also examine Catalan from an articulatory perspective, but they focus on $C_1C_2\#C_3$ sequences where C_2 is always an obstruent, while C_1 and C_3 may be an obstruent or a sonorant. The goal of their study is to investigate the extent to which word-final obstruents assimilate in voicing to the following word initial voiced consonant. The authors work within the Degree of Articulatory Constraint (DAC) model of coarticulation, which is based on the principle that the extent to which consonants resist the coarticulatory effects of other phonetic segments *(coarticulation resistance)* and exert coarticulatory effects on these adjacent segments *(coarticulation aggressiveness)* ought to increase with the involvement of a given articulator in their production. Thus, for example, since the tongue dorsum is more actively involved in the production of palatal consonants than in the case of labials and alveolars, the former consonants ought to be more resistant to tongue dorsum coarticulation effects from the adjacent vowels than the latter, while at the same time exerting more prominent coarticulatory effects on the vowels in question. Similarly, consonants which – because of their production requirements – are more prone to exhibit overall voicing are the ones that ought to be the most resistant to changes in voicing degree induced by the adjacent consonants and should also be the most aggressive as triggers of voicing. Thus, for example, sonorants (nasals, laterals) are expected to exert more voicing coarticulation on preceding obstruents (stops, fricatives) than obstruents since they exhibit more voicing and are less prone to devoice across contextual conditions.

In an earlier study (Recasens–Mira 2012), the authors found, contrary to the initial expectation, that syllable-final fricatives and stops showed much less voicing than expected before nasals and laterals (above 80% voicing in C_2, less than 45% voicing in C_1), and voicing differences as a function of place of articulation did not extend into C_1. Note that C_2 in this case is the target consonant not the trigger. According to the authors, the presence of little voicing during obstruents followed by nasals and laterals appears to be due to the need to preserve the pressure difference across the oral constriction for intense turbulence and thus the integrity of the frication noise for fricatives, and to allow for a sufficient intraoral pressure build-up for the generation of a salient burst for stops, which could be impaired if regressive voicing occurred simultaneously with anticipatory nasalization for nasals and with anticipatory tongue front raising for laterals.

Data for three-consonant clusters reported by Recasens–Mira (2013) show lower percentages of vocal fold vibration in all three consonants as a general rule. Thus, voicing percentages across speakers and contextual conditions for syllable final obstruents subjected to voicing assimilation amounted to 5–45% in CCC sequences and to 30–45% in CC sequences in the case of fricatives, and to 5–55% in three-consonant clusters and to 55–60% in two-consonant clusters in the case of stops. These percentages confirm the hypothesis stemming from DAC that the degree of voicing should decrease with the number of consonants in the cluster and thus with an increase in the aerodynamic and articulatory demands involved. Consonant voicing percentages in three-consonant clusters differ considerably as a function of manner and place of articulation. Voicing coarticulation effects from specific consonants on others yielded little support for the Catalan regressive voicing rule, as the contribution of C_3 to voicing in the preceding syllable/word-final consonants was relatively small and did not always agree with the initial prediction that regressive voicing should increase with voicing degree in the triggering consonant. In particular, there was little voicing during obstruents when followed by a nasal or a lateral, which contradicts DAC, as the authors speculate, perhaps in order to allow for sufficient intra-oral pressure build-up for the generation of turbulent airflow and a burst which could be impaired by anticipatory nasalization for nasals and an earlier apical constriction for laterals.

Duration data reveal that the effect in question may be accompanied by C_2 shortening mostly when C_3 is a nasal. The patterns of voicing interaction between C_1 and C_2 lend some support to the hypothesis that voicing effects should be stronger if involving consonants located within the same syllable and word than across a syllable and word boundary. C_3 stop burst duration was also greater for clusters with a voiceless C_3 than for those with a voiced C_3 in stop + /s/ + stop clusters. Duration effects associated with the C_3 voicing distinction could not be traced during C_1 or the vowel preceding the cluster. These segment duration and intensity data suggest that speakers of languages where voiced stops exhibit voicing lead may use not only vocal fold vibration but other phonetic characteristics that depend more closely on air pressure and airflow for cueing the voicing contrast in clusters – as supported by a number of studies from different languages. The vocal fold vibration and segmental duration and intensity data just summarized indicate that, contrary to current descriptive and phonological accounts, voicing assimilation in Catalan three-consonant clusters with a voiced C_3 cannot be modelled as a purely regressive process (Wheeler 2005). C_3-dependent regressive voicing effects occur less than predicted by the phonological rule: obstruents are mostly voiceless when occurring in C_1. C_3 position effects extend to some extent into C_2 but barely into C_1. Vocal fold vibration data provide some support for voicing dependency between C_1 and C_2 and thus consonants placed in the same syllable final position. It thus appears that voicing assimilation may be conditioned by syllable and word affiliation as well. Moreover, considerable voicing effects between the two syllable final consonants occur at the progressive but not at the regressive level.

Markó et al. (2010) investigated CC and CCC clusters within the word and across the word-boundary in spontaneous and read speech in Hungarian. Here we only mention clusters that were not interrupted by pause of any length. Measurements were carried out manually by the authors, and realizations were classified into three groups: a consonant was considered as voiced if it contained a quasi-periodic signal in at least 80% of its duration. A consonant was considered as voiceless if it contained quasi-periodic signal in at most 20% of its duration. Between these values the consonant was designated to be partially voiced. It is somewhat difficult to evaluate the results of this study since both the manner and the place of articulation of the members of CCC clusters were quite varied in the spontaneous corpus, as well as the number of occurrences. The demonstrative *azt* 'that-ACC', for instance, was highly overrepresented, and there were very few tokens with three obstruents. There was much inter- and intra-speaker variation as well. Nonetheless, the authors conclude that for partially voiced realizations, a large difference is found between the voicing and devoicing types of assimilation. They assume that this difference is due to articulatory concomitances like the interaction of voicing assimilation and the physical constraint of devoicing, their reasoning is similar to Recasens–Mira (2013). They claim that the variability of the data confirms that Hungarian voicing assimilation is a gradient and sometimes only partly regressive process. They also observe that the process most of the time seems to operate obligatorily and that speech style can override it.

Singleton consonants in Slovak

In the remainder of this section, we will briefly summarize our earlier findings on the laryngeal properties of word-final alveolar obstruents in Slovak (Bárkányi–G. Kiss 2012, 2013). In Slovak, word-final single /t/, /d/, /s/ and /z/ were realized completely voiceless before a silent pause (with over 90% of unvoiced frames for all target consonants under scrutiny). There was no statistically significant difference between the voiced and the voiceless obstruents: /t/ vs. /d/: $b = 0.444$, $t(15) = 0.25$, $p = 0.806$; /s/ vs. /z/: $b = 2.692$, $t(15) = 1.54$, $p = 0.143$ (Figure 1). (The methodology of the experiment on singleton consonants summarised here is identical to that presented in the section *Experiment* below.)

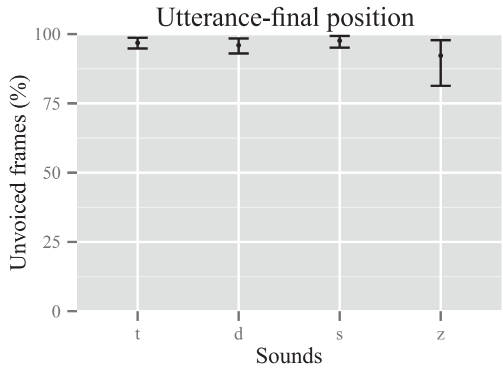

Figure 1.
Means of the ratio of the unvoiced part to total consonant length in Slovak utterance-final /t d s z/ (error bars indicate 95% confidence intervals)

We found a statistically significant difference in the case of /s/ vs. /z/ for two of the acoustic correlates of the voicing contrast. Their duration was not significantly different; however, the duration of the preceding vowel turned out to be significantly different [$b = -6.619$, $t(15) = -2.95$, $p = 0.0099$, effect size: $r = 0.61$], and consequently the vowel-to-consonant (V : C) duration ratio was also significantly larger for /z/ than for /s/ [$b = -0.051$, $t(15) = -2.44$, $p = 0.028$, effect size: $r = 0.53$]. Table 1 sums up the phonetic variables measured in utterance-final position and whether the members of each obstruent pair differed in a statistically significant way for them.

Table 1: Acoustic correlates of obstruent voicing in utterance-final position
* stands for a statistically significant difference ($p < 0.05$)

Acoustic correlates	**/t/–/d/**	**/s/–/z/**
Unvoiced frames		
Voicing duration		
Consonant duration		
Preceding vowel duration		*
V : C duration ration		*

Note that contrast-preservation between the voiced–voiceless pairs in word-medial intervocalic position is robust in the language. Word-final utterance-medial consonants, i.e., obstruents followed by another consonant in the next word showed sporadic significant differences between the acoustic properties of the voiced–voiceless members of the alveolar obstruent pairs, but mostly point to the direction of voicing neutralization. Before /p/, the stops /t/ and /d/ differed with respect to the vowel duration variable only (not even in V : C ratio); before /b/ we found no difference in the case of stops, but /s/ and /z/ differed in V : C duration ratio.

In pre-sonorant position the fricatives were found to be statistically different in the amount of voicing, in consonant duration and vowel duration as well, although not in their ratio. Note that both /s/ and /z/ were realized with a fair amount of voicing (mean percentage of unvoiced frames for /s/: 20.95%, SD = 33.64%, for /z/: 39.32%, SD = 39.49%; mean voicing duration for /s/: 44 ms, SD = 18 ms, for /z/: 39 ms, SD = 22 ms); therefore, we might suspect that both fricatives are perceived as voiced by speakers, but this must be backed up by a follow-up perception experiment. Furthermore, both the vowel and the fricative itself were longer in the case of the underlyingly voiced fricative /z/, which partly contradicts universal trends. Figure 2 sums up the mean percentages of voicing in the final alveolar obstruents in Slovak in three assimilation environments.

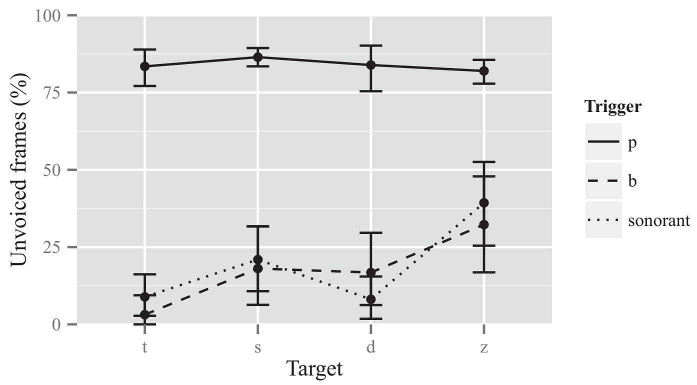

Figure 2.
Interaction graphs showing the mean ratio of the unvoiced part to total consonant length in word-final utterance-medial /t d s z/ followed by the voiceless obstruent /p/, the voiced obstruent /b/ and the sonorant consonants /m l/ in Slovak (error bars indicate 95% confidence intervals)

As mentioned above, the voicing distinction in Slovak single consonants seems to be almost fully neutralized in utterance-final position (final devoicing), as well as before voiced and voiceless consonants. Note that sonorants in Slovak have as much "voicing power" as voiced obstruents. Word-final obstruents do not differ in their voicing (unvoiced frames, voicing duration) whether they are followed by /b/ or sonorants, but they are significantly less voiced when followed by /p/ [/t/ followed by /b/ vs. the sonorants: $b = -2.90$, $t(10) = -1.13$, $p = 0.284$; /t/ followed by /b/ and the sonorants vs. /p/: $b = 25.80$, $t(10) = 16.0$, $p < 0.001$, effect size: $r = 0.98$; /s/ followed by /b/ vs. the sonorants: $b = -1.61$, $t(10) = -0.430$, $p = 0.675$; /s/ followed by /b/ and the sonorants vs. /p/: $b = 22.25$, $t(10) = 9.63$, $p < 0.001$, effect size: $r = 0.95$]. This indicates that sonorants in Slovak do not form an intermediate category as triggers of voicing assimilation. Pre-sonorant voicing in Slovak clearly and categorically patterns with pre-voiced obstruent voicing.

The present study aims to further investigate whether voicing assimilation in Slovak is really taken to the "end", that is to say, whether it really is a stabilized categorical process, or it is more of a low level coarticulatory/phonetic phenomenon. We assume that if the process is categorical (including final devoicing), it should be neutralizing. If, on the other hand, voicing assimilation in Slovak is coarticulatory, the absolute duration of the voiced part across single consonants and consonant clusters should be fairly constant.

Experiment

Material

Words ending in /kt/–/gd/ and /st/–/zd/ were tested in the following three positions: (i) absolute word-final (utterance-final) position; (ii) word-final sentence-medial position, where the target obstruents were followed by one of the following triggers: (a) voiced obstruent /b/, (b) voiceless obstruent /p/, (c) sonorant consonant (/l/ or /m/), and (iii) sentence-medial intervocalic position.

The target obstruents were always preceded by the vowel /a/ or /o/; in intervocalic position, the vowel following the target consonants was /a/. The test words were *kontrakt* 'contract', *smaragd* 'emerald', *chvost* 'tail' and *drozd* 'blackbird'. The use of minimal pairs was avoided on purpose because in our experience, despite the use of a fair number of distractors, subjects tend to overemphasize the differences in their pronunciation. Stimuli were embedded in carrier sentences: e.g., *Kontrakt bez pečate je neplatný* 'The contract without a stamp is invalid'. The carrier sentences were 10–13 syllables long, neutral sentences, the target and the trigger occured in the same intonational phrase; word-stress in Slovak falls on the first syllable. We did not find a significant difference in the behaviour of /l/ and /m/ with regard to their voicing capabilities, therefore we decided to collapse the data from pre-/m/ and pre-/l/ positions together into a common 'pre-sonorant consonant' context.

Methods

Six native speakers of Slovak participated in the experiment aged 20–52, none of them reported any speaking, hearing or reading disorder. They were all naive as to the aims of the experiment and participated as a courtesy to the authors. Subjects read the test sentences and fillers from a monitor screen in a randomized order, which was generated by SpeechRecorder.[2] Each test sentence was read five times, but the first reading was considered as the familiarization phase, and was not taken into consideration. We investigated six contexts for four words by six subjects with four repetitions, which resulted in altogether 576 test items. Recordings were made in a sound-attenuated room with a Sony ECM-MS907 microphone connected to a laptop through an M-Audio MobilePre USB preamplifier external sound card. The material was recorded at a 44,100 Hz sampling rate, and was resampled at 22,050 Hz for the various acoustic measurements.

Measurements

The acoustic analysis was carried out in Praat (version 5.3.12, Boersma–Weenink 2012), for the statistical analysis we used R (version 2.15.0). The spectrograms were segmented manually by the authors and the following measurements were carried out on the basis of the inserted boundaries:

(2) Correlates of laryngeal contrast measured in the experiment

a. *Phonation*-related correlates of laryngeal contrast
   i. the absolute length of the voiced interval
   ii. ratio of the unvoiced part compared to the total length of the consonant (“% of unvoiced frames”)

b. *Duration*-related correlates of laryngeal contrast
   i. duration of the preceding vowel
   ii. duration of the target consonant
   iii. vowel-to-consonant duration ratio

Voicing was measured manually, based on the visual inspection of the spectrograms and oscillograms. In the case of stops, voicing was measured during the closure phase, i.e., up to the burst, but the release phase was not included (similarly to the methodology applied by Strycharczuk 2012, for instance). In the case of fricatives, voicing was measured during the whole duration of the frication noise. We measured two parameters: the absolute length of the voiced interval within the target consonants in seconds (referred to as “voicing duration” in the figures below) and the ratio of the unvoiced part compared to the total length of the consonant (referred to as “unvoiced frames” in the figures below).

As for the duration-related parameters, in the case of fricatives, the interval of frication noise was measured. In the case of stops, closure duration and release burst duration were measured. Since absolute segment durations are highly variable due to different speaking rates, the vowel-to-consonant dura-

[2] http://www.bas.uni-muenchen.de/Bas/software/speechrecorder/

tion ratio was also measured. It has been observed (e.g., Port–Dalby 1982) that the ratio between vowel duration to stop closure or fricative constriction remains relatively constant in words with the same underlying voicing feature: the vowel-to-consonant duration ratio is generally larger for voiced obstruents than for voiceless obstruents. Many perception-driven accounts derive the inverse patterning of voiced–voiceless obstruent length and preceding vowel duration as a form of mutual auditory enhancement for the voicing contrast. The idea is that increased vowel duration makes the duration of a following obstruent appear shorter, and conversely that a decrease in vowel duration increases the perceived duration of a following obstruent, and that vowel duration and obstruent duration are therefore integrated into a single percept (Port–Dalby 1982; Massaro–Cohen 1983; Kluender et al. 1988; Port–Leary 2005). This hypothesis has been largely supported by experimental evidence. Thus, listeners pay attention especially to the relative duration of a vowel and the constriction duration of a following obstruent (Javkin 1976; Parker et al. 1986; Kingston–Diehl 1994), which may serve to preserve the voicing contrast in phonetically unfavourable positions, known as "pre-fortis clipping" in the English phonetics/phonological literature (e.g., Wells 1982). For an extensive overview on the choice of these (and other) acoustic correlates of voicing contrast and voicing assimilation, see Jansen (2004), Bárkányi–Kiss (2007) and Strycharczuk (2012), and the references therein. In the present paper we will not discuss other cues that are also cited in the literature as correlates of laryngeal contrast, such as intensity of the frication and the burst, or the f_0 and F_1 of the surrounding vowels.

Statistical analysis

The acoustic correlates of voicing were analyzed with linear mixed-effects models, using the *nlme* package in R (Pinheiro et al. 2013; for the method see Field et al. 2012). For each phonetic outcome variable a mixed-effects model was fitted with random intercepts for subjects. The fixed parts of the model were: target clusters and trigger sounds (when a following sound followed the target cluster). The contrast coding (using planned orthogonal contrasts) distinguished between 1. target class: stop + stop (/kt, gd/) vs. fricative + stop (/st, zd/); 2. /kt/ vs. /gd/, and 3. /st/ vs. /zd/. The effect size measure used in the paper is Pearson's correlation coefficient *r*.

Miscellaneous issues

We discarded from the analysis those cases where a silent pause of any length appeared between the word-final cluster and the following consonant in the sentence-medial word-final position. This amounted to 6.4% of all the tokens; such pauses were observed in the speech of two subjects. Four instances had to be discarded due to a technical error, thus 548 items could be analysed eventually.

For the current experiment we aimed to investigate the same type of final consonants (alveolar stops and fricatives) as in our earlier study on singletons (see the previous section and Bárkányi–G. Kiss 2012; 2013) so that the two sets of results could be consistently compared across the two studies. Furthermore, our aim was to elicit a fairly natural speech tempo. These factors partly contributed to C_2 being deleted in over half of the cases. In stop + stop clusters before C_3, C_2 was deleted in 66.07% of the cases. It was in only 16.98% that both C_1 and C_2 were kept and fully released. The percentage of C_2 deletion in clusters composed of a fricative and a stop was 66.41%, but C_2 was kept and fully released in 30.79% of the cases.

Results and discussion

Utterance-final position

In this context both consonants (C_1 and C_2) were preserved. (This was not the case in word-final utterance-medial context as mentioned above in the previous section). The duration of the clusters was on average 145–158% of that of singleton consonants. In this position we did not find statistically significant differences between the underlyingly voiced vs. voiceless clusters for any acoustic cues. [Stops, unvoiced frames: $b = 1.012$, $t(15) = 1.304$, $p = 0.211$; fricatives, unvoiced frames: $b = -0.335$, $t(15) = -0.436$, $p = 0.669$; stops, vowel duration: $b = -0.002$, $t(15) = -1.48$, $p = 0.159$; fricatives, vowel duration: $b = 0.0004$, $t(15) = 0.273$, $p = 0.788$; stops, consonant duration: $b = 2.325$, $t(15) = 0.69$, $p = 0.5$; fricatives, consonant duration: $b = 0.04$, $t(15) = 0.012$, $p = 0.99$]. As Figure 3 exhibits, all the clusters were over 95% devoiced in utterance-final position.

Intervocalic position

The intervocalic word-medial position was included in the experiment for the following reason. We assumed that this context was a phonetically "favourable" position, where contrast-preservation should be relatively robust.[3] Our results backed up this expectation. Figure 4 shows the differences in the voicing ratio of the clusters [/kt/ vs. /gd/: $b = 47.64$, $t(15) = 26.40$, $p < 0.001$, $r = 0.99$; /st/ vs. /zd/: $b = 38.99$, $t(15) = 21.61$, $p < 0.001$, $r = 0.98$].

In this position, similarly to the absolute final context, both members of the cluster were maintained, we found no deletions here. We observed that in 20% of the cases, /g/ was not realized as a stop but rather as an approximant without a closure phase and a noticeable release burst. Vowel duration again showed a similar pattern to singletons. It was before /kt/ and /gd/ that vowel length did not significantly differ [/kt/ vs. /gd/: $b = -2.65$, $t(15) = -0.89$, $p =$

[3] On the phonetically favourable nature of the prevocalic, intervocalic context for contrast preservation, see, among others, Steriade (1997); Hayes (1999); Hayes–Steriade (2004); Wright (2001, 2004).

0.3849; /st/ vs. /zd/: $b = -8.77$, $t(15) = -2.97$, $p = 0.0096$, $r = 0.61$), but overall if we consider the vowel–consonant duration ratio, both the stop and the fricative cluster pairs showed shortening effects; see Figure 5 [/kt/ vs. /gd/: $b = -0.082$, $t(15) = -3.21$, $p = 0.0058$, $r = 0.64$; /st/ vs. /zd/: $b = -0.19$, $t(15) = -7.33$, $p < 0.001$, $r = 0.88$].

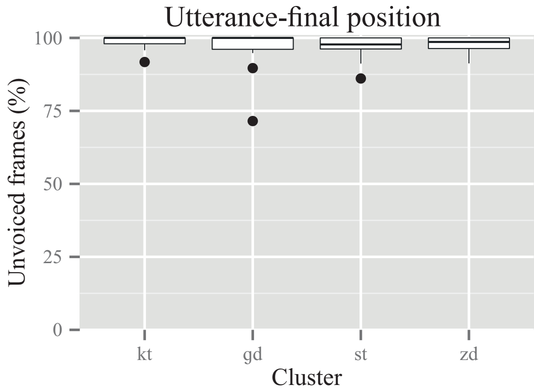

Figure 3.
Boxplots showing the ratio of the unvoiced part to total consonant length in Slovak utterance-final /kt gd st zd/ clusters

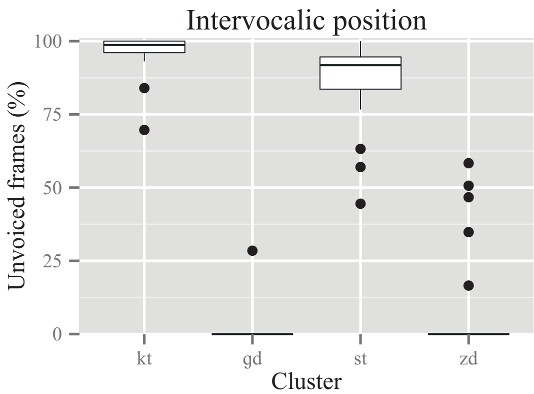

Figure 4.
Boxplots showing the ratio of the unvoiced part to total consonant length in Slovak word-medial intervocalic /kt gd st zd/ clusters

We conclude that while there seems to be complete voicing neutralization in utterance-final position, intervocalic obstruent clusters are fully contrastive in the language. Let us now turn to the focus cases of this paper, namely, the voicing properties of CC#C clusters.

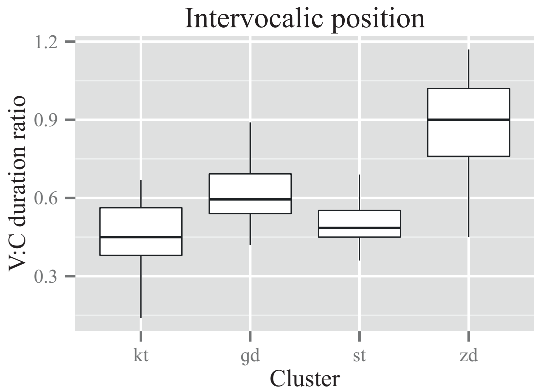

Figure 5.
Boxplots showing the V:C ratio in Slovak word-medial intervocalic /kt gd st zd/ clusters

CC#C clusters

Note that this is the context where in many cases we observed C_2 deletion. First, we will present the results for those cases where C_2 was preserved so that we can observe the voicing properties of consonant clusters in pre-consonant position. And then, we will proceed to compare voicing in C_1 in those cases where C_2 was preserved with those realizations where it was deleted.

We begin with those cases where we expected devoicing, that is, obstruent clusters followed by /p/ in the next word. Our expectations were borne out: all the clusters were realized with around 90% unvoiced frames, and we found no statistically significant differences between the underlyingly voiced vs. voiceless clusters [stop + stop clusters $b = -10.715$, $t(9) = -1.615$, $p = 0.14$; fricative + stop clusters $b = -0.755$, $t(9) = -0.138$, $p = 0.893$]. We did not find significant differences for any of the duration-related correlates either. The situation is very similar for the voicing correlates in the pre-/b/ and pre-sonorant contexts: comparing the voicing ratio of stop + stop clusters before /b/: $b = 19.128$, $t(4) = 1.859$, $p = 0.136$; and the fricative + stop clusters before /b/: $b = -12.78$, $t(4) = -1.414$, $p = 0.23$. Again, we did not find statistically significant differences for any of the duration-related parameters either. As for the pre-sonorant position, stop + stop clusters seem to be neutralized for voicing: $b = -1.163$, $t(11) = -0.239$, $p = 0.814$.

In the case of fricative clusters, we did find a significant difference; however, it occurred unexpectedly: it was /zd/, and not /st/, that displayed less voicing; actually, these clusters were much less voiced than expected (mean unvoiced frames: 50%) compared to /st/ [average unvoiced frames: 24.19%): $b = -11.282$, $t(11) = -2.4$, $p = 0.0348$, $r = 0.96$]. We think that these unexpected results were a consequence of the small number of tokens we could measure (due to the deletion of C_2), and therefore, it is difficult to draw meaningful conclusions for these findings. We sporadically obtained statistically significant results for other variables, too, for similar reasons (low number of tokens). While we did not find significant differences either in vowel length or in consonant length, their ratio turned out to differ significantly for stop + stop clusters in pre-sonorant position, although the effect size of this significance was very low: $b = 0.123$, $t(11) = 2.436$, $p = 0.033$, $r = 0.038$. We note again that for the underlyingly voiced cluster /gd/ we found a lower value (duration ratio = 0.807) than for /kt/ (duration ratio = 1.00), which contradicts the usually observed tendencies for vowel-to-consonant duration ratio (for comparison, see Figure 5, which illustrates the findings of this duration variable in intervocalic position).

A much more interesting question is whether we can find differences in the voicing "aggressiveness" of /b/ vs. sonorant consonants, that is, whether the clusters under scrutiny are more voiced before an actively voiced obstruent than before a modally voiced sonorant. Another question related to the categoricality of voicing assimilation and within that pre-sonorant voicing in Slovak concerns whether the voicing properties of singleton consonants differ from those of consonant clusters. If we find important differences, we might assume that voicing assimilation is coarticulatory after all. If, however, the voicing properties of C vs. CC targets are similar, it points to the direction of a categorical process. Note that in this latter case there still might be a phonetic, aerodynamic difficulty in implementing voicing for a longer time, which can give rise to some differences. Let us compare the results on the voicing behaviour of singleton consonants (our earlier study cited above) and consonant cluster targets (present experiment).

We divided our data according to the following parameters: (i) stops vs. fricatives; (ii) singletons vs. clusters; (iii) clusters with deletion vs. no deletion. In this way we obtained the following "target class" groups: single stops (labelled as "singST" in the graphs below), single fricatives ("singFR"), stop + stop clusters with no deletion ("NoDelST"), stop + stop clusters where C_2 is deleted and therefore they are realized as single stops ("DelST"), fricative + stop clusters with no deletion ("NoDelFR"), and fricative + stop clusters where C_2 is deleted and therefore they are realized as single fricatives ("DelFR"). Figure 6 summarizes the amount of voicing in the six target classes in four different contexts (trigger classes): before sonorant consonants, before /b/, before /p/ and "nothing", which stands for the absolute final position where there is no triggering segment. Note that in this final context, both members of the clusters were systematically articulated and therefore the groups "DelST" and "DelFR" are not applicable here.

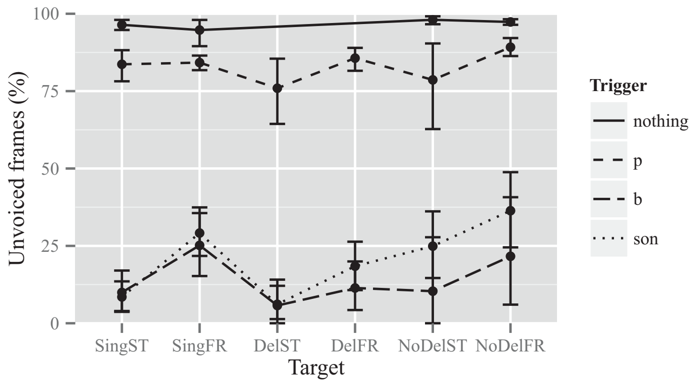

Figure 6.
Interaction graphs showing the mean proportion of voicing in word-final utterance-medial /t d s z/ and /kt gd st zd/ followed by the voiceless obstruent /p/, the voiced obstruent /b/ and the sonorant consonants /m l/, as well as in utterance-final position in Slovak
(error bars indicate 95% confidence intervals)

As far as the voicing aggressiveness of /b/ vs. sonorant consonants is concerned, if we add all six target groups up, we do find a statistically significant difference with a medium effect size [$b = -3.21$, $t(26) = -2.231$, $p = 0.0345$, $r = 0.53$), despite the fact that both trigger full voicing (mean 19% of unvoiced frames in the case of sonorants and 14% for /b/). However, there are no interaction effects (as can be seen on Figure 6), that is, /b/ vs. sonorants do not cause differences in voicing to the six classes that are examined here. Figure 7 illustrates how small the difference between the two groups is.

This suggests that RVA, which seems to be categorical in Slovak, is extended to the pre-sonorant position as well. The statistically significant difference observed is not due to less amount of voicing in obstruents before sonorant consonants as opposed to /b/, but rather to the fact that there are more instances when RVA fails to apply in pre-sonorant position (Figure 8). In these cases word-final devoicing occurs, which – when averaged across the board – gives a result between voiceless and voiced obstruents. This is in accordance with Strycharczuk (2012), who claims that pre-sonorant voicing is categorical but optional.

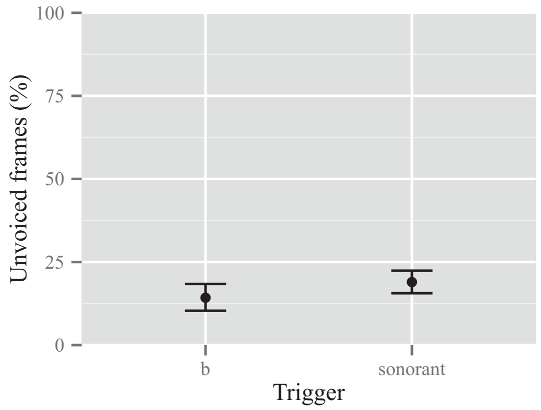

Figure 7.
Mean values for the voiceless–voiced ratio before /b/ and before sonorant consonants (error bars indicate 95% confidence intervals)

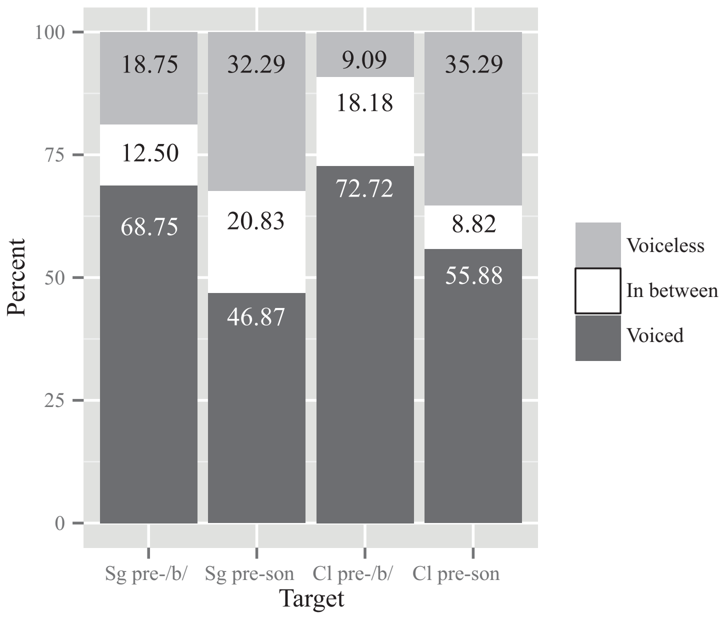

Figure 8.
Percentages of voiceless, partially voiced and voiced realizations of singleton obstruents and obstruent clusters in pre-/b/ and pre-sonorant position

We divided our data into three categories: (i) fully voiced realization, (ii) partially voiced realization and (iii) voiceless realization. The criteria for the classification were the following arbitrarily determined values. We classified those instances as "fully voiced" which contained up to 29% of unvoiced frames, that is they were at least 70% voiced. "Voiceless" occurrences were those which contained at least 71% of unvoiced frames, the rest were labelled as "in between", that is, partially voiced. Figure 8 clearly demonstrates that it is not so much the partially voiced realizations that are considerably more numerous in the case of pre-sonorant obstruents but the instances of voiceless realizations. Therefore, in the remainder of this section, we will treat the pre-sonorant and pre-/b/ contexts as one group.

In Figure 9 we compare the amount of voicing between single consonants – this group comprises singleton obstruents and those single obstruents that remain from clusters with C_2 deletion – and CC clusters in pre-/b/ and pre-sonorant position. The two groups do not differ with regard to the voiced–voiceless portion during the obstruent(s): $b = -1.919$, $t(24) = -1.522$, $p = 0.141$.

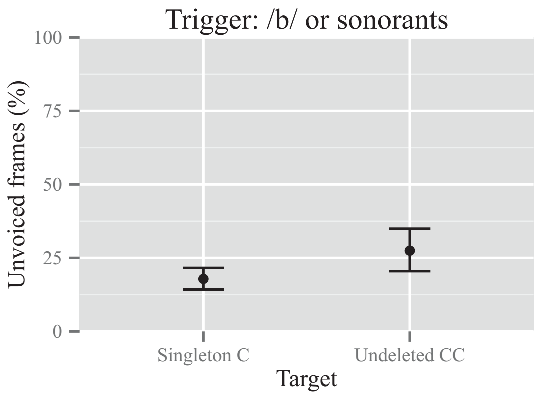

Figure 9.
Mean values for the ratio of the unvoiced part to total consonant length of single obstruents vs. obstruent clusters
(error bars indicate 95% confidence intervals)

This result is indicative of a non-coarticulatory process since it is not the absolute voicing duration that is similar (see Figure 10), but the *proportion* of voicing. If the absolute voicing duration is the same or similar across different consonant lengths, we might suspect that voicing is due to articulatory inertia, so it is coarticulatory. If it is the proportion of voicing that is similar in single consonants and consonant clusters, it should probably be attributed to a pre-planned articulatory gesture. As Figure 10 shows, the duration of the voiced portion of the obstruent cluster is significantly longer than the voiced portion of the single consonant, although the magnitude of the effect is very small [$b = -0.004$, $t(22) = -4.638$, $p = 0.0001$, $r = 0.001$].

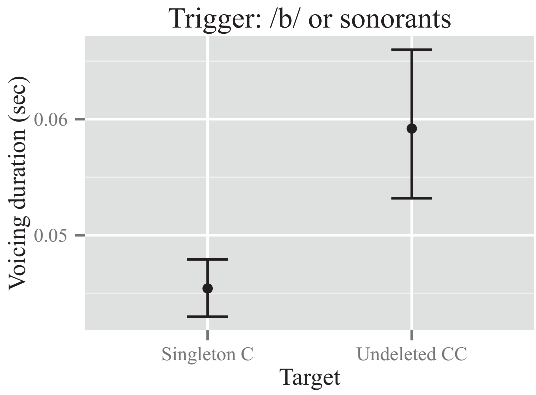

Figure 10.
Mean values for the duration of the voiced part in single obstruents vs. obstruent clusters in voicing context
(error bars indicate 95% confidence intervals)

If we tease apart the single consonant class and compare true singletons with single consonants that result from cluster simplification in pre-consonant environment, we find that the two groups are not different with regard to their duration [$b = 2.99$, $t(10) = 1.352$, $p = 0.2062$], but they significantly differ in the voiced portion (again, however, the effect size is very small): $b = -0.003$, $t(24) = -3.879$, $p = 0.0007$, $r = 0.001$ (see Figure 11).

Interestingly enough, if we compare the voiced interval of reduced clusters with those of fully realized clusters we do not find a statistically significant difference [$b = -0.001$, $t(24) = -1.497$, $p = 0.147$], while they do differ with regard to their voicing ratio [$b = -1.958$, $t(24) = -2.98$, $p = 0.0065$, $r = 0.37$]. These results indicate that in the case of reduced stops there is a planned but unrealized cluster the voicing of which is implemented by speakers despite the deletion of C_2. This finding suggests that RVA in Slovak is not coarticulatory, although, we must warn the reader that this experiment should be replicated with a larger set of data, which also includes non-alveolar stops so that cluster simplification may be avoided. In Figure 12 we sum up the effects of voicing on obstruents in Slovak.

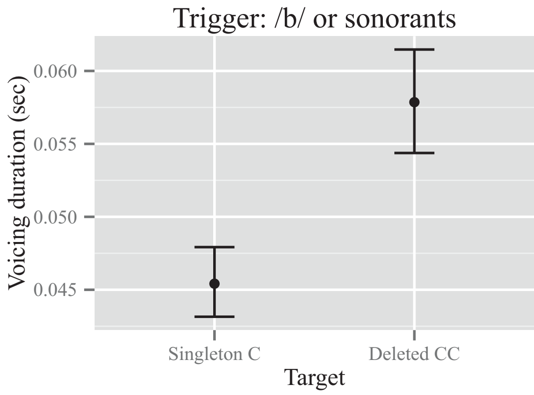

Figure 11.
Mean values for the duration of the voiced part in intended single obstruents vs. reduced obstruent clusters in voicing context
(error bars indicate 95% confidence intervals)

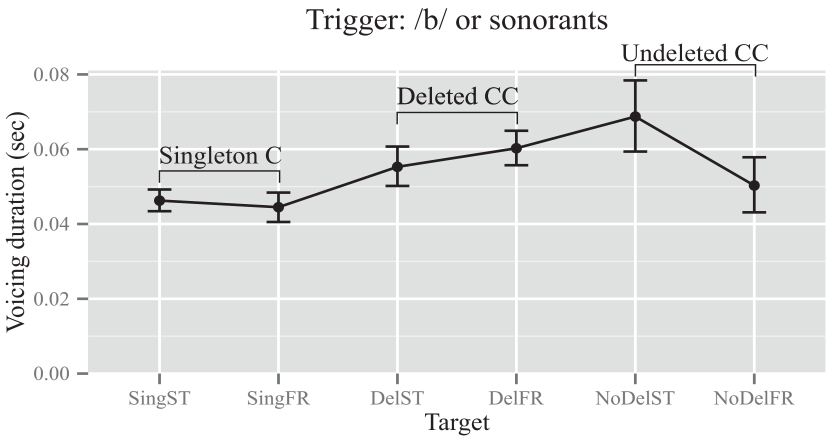

Figure 12.
Mean values for the duration of the voiced part in intended single stops and fricatives, reduced stop and fricative clusters and fully realized stop + stop and fricative + stop clusters in voicing context
(error bars indicate 95% confidence intervals)

We can see in Figure 12 that single consonants are different from simplified and undeleted clusters, while the latter two are not statistically significant with regard to their voicing duration only their voicing ratio. Stops tend to be more voiced than fricatives. This tendency is more robust in the case of longer intervals, that is, in undeleted clusters. It seems that fricatives reach a "voicing ceiling" earlier than stops, which is not unexpected due to the aerodynamic difficulty of initiating and maintaining voicing in fricatives as opposed to stops (e.g., Ohala 1983; Stevens 1998). As this paper has also shown, vowels are consistently longer before fricatives than before stops, which appears to be the case before single fricatives as well as fricative-initial clusters. We leave the investigation of this issue for future research.

A last piece of evidence we cite here supporting our claim that voicing assimilation in Slovak is not coarticulatory but rather categorical (and at times, optional), comes from Beňuš–Trnka (2014), who demonstrate that conversational fillers starting with a voiced schwa-like vowel like *umm, ur,* etc. function as prosodic breaks and as such induce word-final devoicing. However, in a non-negligible number of cases they do trigger voicing assimilation, indicating that speakers display a bimodal behaviour with a choice between two categorical options: they either produce word-final devoicing or they implement pre-sonorant voicing.

Conclusions

This paper has presented a modest contribution to the study of voicing assimilation in Slovak, a language for which instrumental/experimental phonetic and phonological research is lacking. More specifically, we have investigated the voicing properties of three-consonant clusters (CC#C), and how voicing assimilation affects them, an area of Slovak phonetics and phonology which has not received enough attention either. This paper is also a contribution to the study of pre-sonorant voicing, a topic of growing interest both empirically and theoretically. Our experiment has shown that word-final obstruent clusters in Slovak (just like singletons) are realized completely voiceless. This finding indicates that Slovak obstruents are categorically targetless for voicing in this position, which is claimed to be a pre-requisite of pre-sonorant voicing. It has been also shown that sonorant consonants and voiced stops do not differ in their voicing "capabilities" in this language, thus sonorants do not form an in-between category between voiceless obstruents and voiced obstruents. A novel finding of the paper is that pre-sonorant voicing assimilation in Slovak appears to be optional but categorical, rather than obligatory or gradual. We have not found any evidence for voicing assimilation, including pre-sonorant voicing, being a coarticulatory process for any of the speakers of our experiment. An unexpected result of the paper that is in need of further clarification is that vowels before fricatives are realized consistently longer than vowels before stops, irrespective of the prosodic position. It would also be interesting to study on a larger set of data whether C_2 deletion varies systematically with voicing.

References

Bárkányi, Zsuzsanna – Kiss, Zoltán 2007. *A phonetically-based approach to the phonology of [v]: A case study from Hungarian and Slovak*. Paper presented at the 4th Old World Conference in Phonology. Rhodes, Greece. 18–21 January.

Bárkányi, Zsuzsanna – G. Kiss, Zoltán 2012. *On the border of phonetics and phonology: Sonorant voicing in Hungarian and Slovak*. Paper presented at the 20th Manchester Phonology Meeting, Manchester. May 24–26th.

Bárkányi, Zsuzsanna – G. Kiss, Zoltán 2013. *Phonetics or phonology? Why do sonorants not voice in Hungarian?* Paper presented at the 11th International Conference on the Structure of Hungarian, Piliscsaba. August 29–31.

Beňuš, Štefan – Gafos, Admantios 2007. Articulatory characteristics of Hungarian 'transparent' vowels. *Journal of Phonetics* 35. 271–300.

Beňuš, Štefan – Trnka, Marián 2014. *Prosody, voice assimilation and conversational fillers*. Manuscript. Institute of Informatics, Slovak Academy of Sciences.

Boersma, Paul – Weenink, David 2012. Praat: Doing phonetics by computer. (Version 5.3.12) [Computer program]. Retrieved from http://www.praat.org/

Bonet, Eulalia – Lloret, María-Rosa 1998. *Fonologia catalana*. Ariel, Barcelona.

Burton, Martha – Robblee, Karen 1997. A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25. 97–114.

Charles-Luce, Jan 1985. Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics* 13. 309–324.

Charles-Luce, Jan 1993. The effects of semantic context on voicing neutralization. *Phonetica* 50. 28–43.

De Schutter, Georges – Taeldeman, Johan 1986. Assimilatie van Stem in de Zuidelijke Nederlandse Dialekten. In Devos, Maud – Johan Taeldeman (eds.): *Vruchten van z'n akker: opstellen van (oud-) medewerkers en oud-studenten voor Prof. V.F. Vanacker.* Seminarie voor Nederlandse Taalkunde, Ghent. 91–133.

Dinnsen, Daniel A. 1985. A re-examination of phonological neutralization. *Journal of Linguistics* 21. 265–279.

Dinnsen, Daniel A. – Jan Charles-Luce 1984. Phonological neutralization, phonetic implementation, and individual differences. *Journal of Phonetics* 12. 49–60.

Field, Andy – Miles, Jeremy – Field, Zoe 2012. *Discovering statistics using R.* Sage, London.

Fourakis, Marios – Iverson, Gregory 1984. On the 'incomplete neutralization' of German final obstruents. *Phonetica* 41. 140–149.

Hallé, Pierre A. – Adda-Decker, Martine 2011. Voice assimilation in French obstruents: Categorical or gradient? In Goldsmith, John – Hume, Elizabeth – Wetzels, Leo (eds.): *Tone and features: phonetic and phonological perspectives.* De Gruyter, Berlin. 149–175.

Hayes, Bruce 1999. Phonetically driven phonology: The role of Optimality Theory and inductive grounding. In Darnell, Michael – Moravcsik, Edith A. – Noonan, Michael – Newmeyer, Frederick J. – Wheatly, Kathleen (eds.): *Functionalism and formalism in linguistics.* Vol. 1. John Benjamins, Amsterdam–Philadelphia. 243–285.

Hayes, Bruce – Steriade, Donca 2004. The phonetic bases of phonological markedness. In Hayes, Bruce – Kirchner, Robert M. – Steriade, Donca (eds.): *Phonetically based phonology.* Cambridge University Press, Cambridge. 1–33.

Hualde, José-Ignacio 2005. *The sounds of Spanish*. Cambridge University Press, Cambridge.

Hyman, Larry M. 1975. *Phonology: Theory and analysis*. Holt, Rinehart and Winston, New York.

Jansen, Wouter 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. Doctoral dissertation, Rijksuniversiteit Groningen.

Jassem, Wiktor – Richter, Lutoslawa 1989. Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17. 317–325.

Javkin, Hector R. 1976. The perceptual basis of vowel duration differences associated with the voiced/voiceless distinction. *Report of the Phonology Laboratory, UC Berkeley* 1. 78–92.

Jiménez, Jesús – Lloret, María-Rosa 2008. Asimetrías perceptivas y similitud articulatoria en la asimilación de sonoridad del catalán. *Cuadernos de Lingüística del I.U.I. Ortega y Gasset* 15. 71–90.

Kahlen-Halstenbach, Birthe 1990. Zur psychologischen Realität der Auslautverhartung im Deutschen. *Zeitschrift für phonetische Sprachwissenschaft und Kommunikationsforschung* 43. 645–655.

Kaplan, Abby 2011. How much homophony is normal? *Journal of Linguistics* 48/2. 1–41.

Kingston, John – Diehl, Randy L. 1994. Phonetic knowledge. *Language* 70. 419–454.

Kluender, Keith R. – Diehl, Randy L. – Wright, Beverly A. 1988. Vowel length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16. 153–169.

Markó, Alexandra – Gráczi, Tekla Etelka – Bóna, Judit 2010. The realisation of voicing assimilation rules in Hungarian spontaneous and read speech: Case studies. *Acta Linguistica Hungarica* 57. 210–238.

Massaro, Dominic W. – Cohen, Michael M. 1983. Consonant/vowel ratio: An improbable cue in speech perception. *Perception and Psychophysics* 33. 502–505.

O'Dell, Michael – Port, Robert F. 1983. Discrimination of word final voicing in German. *Journal of the Acoustical Society of America* 73. Supplement 1.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage, Peter (ed.): *The production of speech.* Springer, New York. 189–216.

Parker, Ellen M. – Diehl, Randy L. – Kluender, Keith R. 1986. Trading relations in speech and non-speech. *Perception and Psychophysics* 39. 129–142.

Pauliny, Eugén 1979. *Slovenská fonológia*. [Slovak phonology]. Slovenské Pedagogické Nakladateľstvo, Bratislava.

Pinheiro, Jose – Bates, Douglas – DebRoy, Saikat – Sarkar, Deepayan 2013. The nlme Package. Linear and Nonlinear Mixed Effects Models. *R package version 3*. 1–108.

Port, Robert F. – Dalby, Jonathan. 1982. C/V ratio as a cue for voicing in English. *Perception and Psychophysics* 2. 141–152.

Port, Robert F. – Leary, Adam P. 2005. Against formal phonology. *Language* 81. 927–964.

Port, Robert F. – Mitleb, Fares – O'Dell, Michael 1981. Neutralization of obstruent voicing in German is incomplete. *The Journal of the Acoustical Society of America* 70. Supplement 1.

R Core Team 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from http://R-project.org.

Recasens, Daniel – Mira, Meritxell 2012. Voicing assimilation in Catalan two-consonant clusters. *Journal of Phonetics* 40. 639–654.

Recasens, Daniel – Mira, Meritxell 2013. Voicing assimilation in Catalan three-consonant clusters. *Journal of Phonetics* 41. 264–280.

Rubach, Jerzy 1994. *The lexical phonology of Slovak*. Clarendon Press, Oxford.

Rubach, Jerzy. 1996. Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27. 69–110.

Silverman, Daniel. 2012. *Neutralization*. Cambridge University Press, Cambridge.

Simon, Ellen 2010. Phonological transfer of voicing and devoicing rules. Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences* 32. 63–86.

Slowiaczek, Louisa M. – Dinnsen, Daniel A. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13. 325–341.

Slowiaczek, Louisa M. – Szymanska, Helena J. 1989. Perception of word-final devoicing in Polish. *Journal of Phonetics* 17. 205–212.

Steriade, Donca. 1997. *Phonetics in phonology: The case of laryngeal neutralization*. Manuscript. University of California Los Angeles.

Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In Inkelas, Sharon – Hanson, Kristin (eds.): *The nature of the word*. MIT Press, Cambridge, MA. 151–180.

Stevens, Kenneth N. 1998. *Acoustic phonetics*. MIT Press, Cambridge, MA.

Strycharczuk, Patrycja. 2012. *Phonetics–phonology interactions in pre-sonorant voicing*. Doctoral dissertation, University of Manchester.

Strycharczuk, Patrycja – Simon, Ellen 2013. Obstruent voice before sonorants. The case of West-Flemish. *Natural Language and Linguistic Theory* 31/2. 563–588.

Wells, John C. 1982. *Accents of English 1–3*. Cambridge University Press, Cambridge.

Wheeler, Max W. 2005. *The phonology of Catalan*. Oxford University Press, Oxford.

Wright, Richard 2001. Perceptual cues in contrast maintenance. In Hume, Elizabeth – Johnson, Keith (eds.): *The role of speech perception in phonology*. Academic Press, San Diego. 251–277.

Wright, Richard 2004. A review of perceptual cues and cue robustness. In Hayes, Bruce – Kirchner, Robert M. – Steriade, Donca (eds.): *Phonetically based phonology.* Cambridge University Press, Cambridge. 34–57.

The present research was supported by the Hungarian National Research Fund (OTKA K104897).