Acta Linguistica Hungarica, Vol. 57 (2–3), pp. 189–209 (2010)
DOI: 10.1556/ALing.57.2010.2–3.2

ANTICIPATORY COARTICULATION IN HUNGARIAN V*n*C SEQUENCES

MÁRIA GÓSY[a] – ANDRÁS BEKE[b] – ROBERT M. VAGO[c]

[a]Research Institute for Linguistics, Hungarian Academy of Sciences
Benczúr u. 33., H–1068 Budapest, Hungary
gosy@nytud.hu

[b]Department of Phonetics, Eötvös Loránd University
Múzeum krt. 4/A, H–1088 Budapest, Hungary
beke_andras@freemail.hu

[c]Department of Linguistics and Communication Disorders
Kissena Hall 349, Queen College of CUNY
65-30 Kissena Boulevard, Flushing NY 11367-1597, USA
robert.vago@qc.cuny.edu

Abstract: The duration of the vowel and the nasal was analyzed in the casual pronunciation of Hungarian words containing the sequence V*n*.C, where '.' is a syllable boundary and C is a stop, affricate, fricative, or approximant. It was found that due to anticipatory coarticulation the duration of *n* is significantly shorter before fricatives and approximants than before stops and affricates.

A teaching algorithm was used to distinguish between stops/affricates and fricatives/approximants in V*n*C sequences. We used an approach to the classification of C by means of the support vector machine (SVM) and the properties of Radial basis function (RBF) kernel (using MATLAB, version 7.0). The results show close to 95% correct responses for the stop/affricate vs. fricative/approximant distinction of C, as opposed to about 60% correct responses for the classification of the voicing feature of C.

Keywords: vowel duration, nasal duration, anticipatory coarticulation, coarticulatory effects, speech classification

1. Introduction

Speech is produced as a sequence of sounds as the vocal tract moves from one articulatory configuration to the next. Coarticulation refers to the transition from one gesture to another, whereby the different articulation gestures combine with different timing patterns (Hardcastle–Hewlett 1999; Boyce et al. 1990; Chafcouloff–Marchal 1999). Research has shown the existence of different types of nasalization effects in various sequences across languages. There are only four languages out of 317 analyzed that have no nasal segments of any kind, and only six have no nasal stops in their phonemic inventory (Maddieson 1984). Phonetic studies have concentrated on the production, perception, and acoustic consequences of nasal sounds, or on the nasal process and nasalization effects in various contexts (e.g., Ohala 1984; Ladefoged 1993; Laver 1994; Farnetani–Recasens 1999; Beddor 2007; Chen et al. 2007; Delvaux et al. 2007). Research resulted in important findings concerning the acoustic patterns of nasals and nasalized sounds (Malécot 1960; Ohala–Busà 1995; Pickett 1999; Busà 2007, among others). For example, in vowels followed by a nasal and then an oral consonant, coarticulatory nasalization was shown to be more extensive when the oral consonant was voiceless or a fricative, and in nasal contexts lower and longer vowels tended to be more heavily nasalized than higher and shorter vowels, respectively (Bell-Berti 1993; Whalen–Beddor 1989).

The Hungarian speech sound inventory contains a number of nasal consonants that correspond to three nasal phonemes: /m n ɲ/; see Vago (1980), Kenesei et al. (1998), Siptár–Törkenczy (2000), and Gósy (2004), among others. The dentialveolar [n] is the most frequent nasal consonant in spontaneous speech (Gósy 2004). There are no nasal vowel phonemes; however, vowels tend to be naso-orally articulated either preceding or following nasal consonants. In Hungarian, on the basis of acoustic data (Horváth 2005; 2008; Beke–Horváth 2009), progressive nasalization has a stronger effect on vowels than regressive nasalization does.

In this study, anticipatory coarticulation was analyzed in the following Hungarian sound sequences:

$$\begin{array}{ccccc} \text{V} & + & [\text{n}] & + & \text{C} \\ \left\{\begin{array}{l}[\text{ɔ ɛ o ø}]\\ [\text{aː eː iː}]\end{array}\right\} & & & & \left\{\begin{array}{l}[\text{t d ts ʧ}]\\ [\text{s z ʃ ʒ}]\\ [\text{l j}]\end{array}\right\}\end{array}$$

The vowels in the analyzed contexts were followed by the dentialveolar nasal [n] and a postnasal consonant, which was a noncontinuant obstruent (stop or affricate), a continuant obstruent (fricative), or an approximant (lateral or central). (Stops and affricates will be referred to as 'noncontinuant consonants', as opposed to both fricatives and approximants.) The voiced counterparts of the affricates were not used, due to their low frequency. In Hungarian casual speech, the dentialveolar nasal may not be fully pronounced before fricatives and approximants; closure for the nasal may be incomplete, resulting in a shorter nasal sound and a longer vowel before the nasal, as well as in a vowel that is (to a certain extent) nasalized (see Siptár–Törkenczy 2000; Olaszy 2007; Gósy–Vago 2009).

The present research has a twofold goal: (1) to show the acoustic-phonetic consequences of anticipatory coarticulation in Hungarian V*n*C sequences as regards the duration of V and *n*, and (2) to develop a model using a teaching algorithm to distinguish the various V*n*C sequences.

2. Subjects, material, method

Ten native female speakers of Hungarian with no known speech or hearing defects read randomized isolated words in a sound-proof chamber (ages ranged from 20 to 28). The word list consisted of 65 Hungarian words and 5 word combinations (altogether 70 items) that contained the dentialveolar nasal [n] followed by one of the following classes of consonants: voiceless and voiced fricatives ([s z ʃ ʒ]), voiceless and voiced stops ([t d]), voiceless affricates ([ts ʧ]), central and lateral approximants ([j l]). Preceding the nasal, seven vowels were selected that can readily co-occur with all intended clusters in Hungarian words: three long and four short vowels ([aː eː iː ɔ o ø ɛ]). All these speech sounds are phonemes in Hungarian. The nasal and the following consonant were all heterosyllabic, as in *ka**n.d**úr* 'tomcat', *vo**n.z**ás* 'attraction'.[1] All words and word combinations were stressed on the initial syllable, according to the general pattern of the language.

The words were recorded in a soundproof chamber and digitalized using 44.1 kHz sampling rate with 16-bit resolution. Acoustic-phonetic analysis was carried out by Praat software, version 5.4 (Boersma–Weenink 2005). The nasal, the preceding vowel, and the following consonant were identified in each word for each speaker. The words were manually annotated using the Praat software by the authors. The duration of the nasal consonant and of the preceding vowel was measured. The duration of the vowels was measured as the interval from the second formant onset to the onset of the nasal formants. Nasal duration was measured to the onset of a following obstruent consonant or to the onset of the second

[1] Further examples of the list: *kancsal* 'cross-eyed', *hentes* 'butcher', *színdús* 'colorful', *döntés* 'decision', *kénsav* 'vitriol', *láncok* 'chains', *vénlány* 'spinster', *tanszék* 'department', *fenség* 'majesty', *pánsíp* 'panpipe', *bontás* 'demolition', *önjáró* 'self-propelling', *kínzók* 'tormentors', *önző* 'selfish', *lenzsák* 'sack of flax', *bonszáj* 'bonsai', *vén zár* 'old lock', *igen jó* 'very good'.

formant in the case of a following approximant; see Figure 1. Other cues such as formant discontinuity or differences in intensity were also used as segmentation cues.

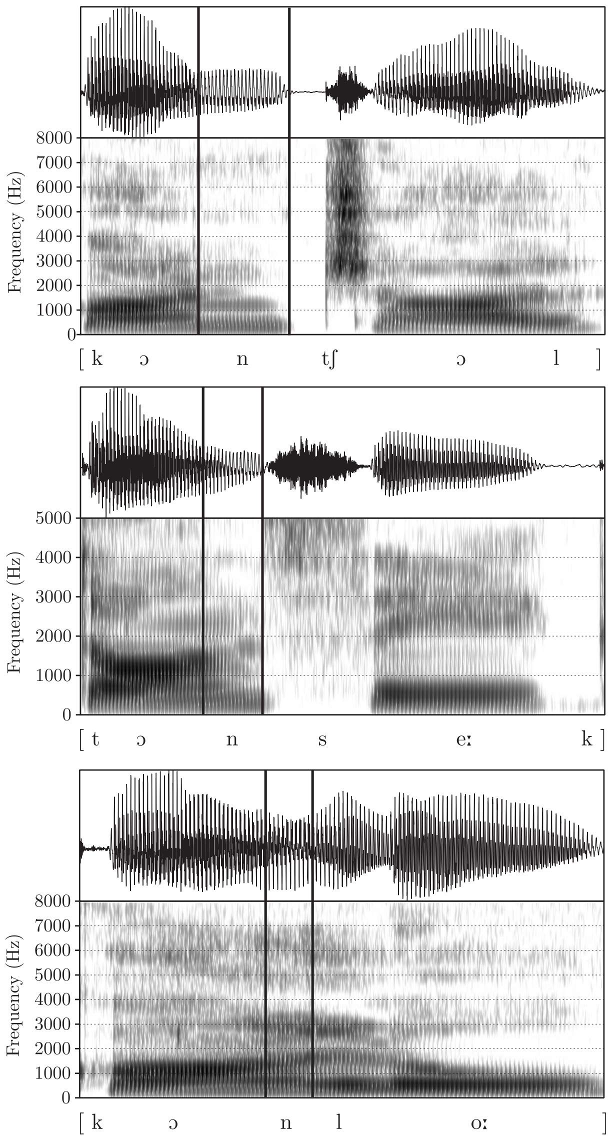

Fig. 1

The acoustic structure of the words *kancsal* ‘cross-eyed’, *tanszék* ‘department’, *kanló* ‘stallion’

The corresponding spectrographic and waveform displays were consulted; auditory perception was also considered. The MATLAB implementation was used for further refinements. Statistical analysis of the data was carried out using ANOVA, UNIANOVA, and MANOVA (using SPSS 14.0). In all cases, the confidence level was set at the conventional 95%.

3. Results

An acoustic-phonetic analysis of the durations confirmed our expectation that anticipatory coarticulation occurs in V*n*C sequences. Specifically, fricatives and approximants affect the duration of the nasal, which appears shorter (as compared to a nasal that appears preceding noncontinuant obstruents). In turn, the temporally modified, i.e., shortened, nasal has an effect on the preceding vowel, which appears longer. The extent of vowel lengthening is larger when the nasal is followed by fricatives than before approximants. We turn to the details in the ensuing subsections.

3.1. Coarticulation effects on nasal duration

In V*n*C sequences, the duration of the dentialveolar nasal varies considerably, depending on the manner of articulation of the following consonant. The data reveal that the nasal is shorter before fricatives (mean value: 59.6 ms) than before noncontinuants (mean value: 107.8 ms). The effect of approximants on the duration of the nasal is similar to that of fricatives (mean value: 53.25 ms); the nasal is shorter before [j] (mean value: 48 ms) than before [l] (mean value: 58.4 ms). See Table 1 and Figure 2.

Table 1

V*n*C sequences: The duration of *n* and the manner of articulation of C

	Duration of *n* (ms)	
C types	Mean	Std. dev.
Noncontinuant	108	30
Fricative	60	24
Approximant	53	21

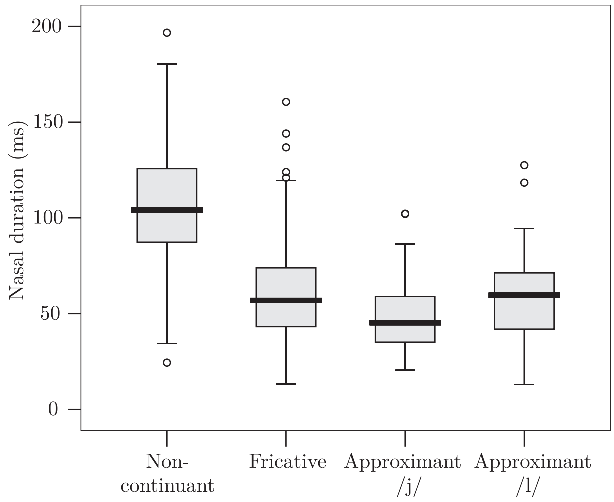

Fig. 2

V*n*C sequences: The duration of *n* and the manner of articulation of C (medians & ranges)

The difference in nasal duration before fricatives vs. noncontinuant obstruents is significant ($F(3, 629) = 193.2$; $p = 0.001$). Post hoc Tukey tests confirm, however, that there is no significant difference in nasal duration before fricatives vs. approximants. The nasal is shorter before the central approximant than before the lateral approximant; however, there is no significant difference in nasal duration between them, either.

Analysis was carried out as to whether the place of articulation of the consonant following the nasal affects the duration of the nasal. There is no significant difference among the consonants in this respect; see Table 2.

Is the duration of the nasal affected by the phonemic length of the preceding vowel? In our data, we found no significant differences in this regard that might be explained by the fact that there were no objective durational differences between the phonemically short and long vowels (see section **3.3**).

In some languages, nasals are shorter before voiceless obstruents than before voiced ones (Beddor 2007). Analysis was made to find out whether this was the case in Hungarian. The data reveal that the nasal is indeed shorter before a voiceless obstruent. The mean duration of the nasal before voiceless obstruents is 82.6 ms (SD = 40.3 ms), while before voiced obstruents it is 85.75 ms (SD = 33.7 ms). Although the difference seems to be small, it turns out to be significant ($F(1, 628) = 71.3$, $p = 0.001$).

Table 2

V*n*C sequences: The duration of *n* and the place of articulation of C

	Duration of *n* (ms) Place of articulation of C					
	Dentialveolar		Palatoalveolar		Palatal	
C types	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Voiced non-continuant	138 /__[d]	28				
Voiceless non-continuant	97 /__[ts]	21	101 /__[ʧ]	21		
	109 /__[t]	24				
Voiced fricative	72 /__[z]	18	60 /__[ʒ]	20		
Voiceless fricative	49 /__[s]	19	49 /__[ʃ]	19		
Approximant	59 /__[l]	22			48 /__[j]	17

The duration of nasals was analyzed by means of a two-factor UNI-ANOVA, based on the duration of nasals as the dependent variable, and on the voicing and the type of the consonant (noncontinuant, fricative, or approximant) following the nasal as the two fixed factors. The univariate analysis confirmed a significant difference for both factors, indicating their decisive effects on nasal duration (factor of voicing: $F(1, 629) = 41.75$, $p = 0.001$; factor of consonant type: $F(1, 629) = 620.22$, $p = 0.001$). The interaction of the two factors also proves to be significant: $F(1, 629) = 12.67$, $p = 0.001$. Refined statistical analyses showed that consonant type has greater impact on nasal duration than voicing (consonant type: Partial Eta Squared = 0.498; voicing: Partial Eta Squared = 0.063; interaction: Partial Eta Squared = 0.020). Analysis revealed that the range of the nasal durations explains 50.8% of all possible factors. This percentage comes from the two fixed factors; however, the consonant type defines the nasal duration to a greater extent, close to 50% (R Squared = 0.508). In sum, both fixed factors and their co-effect define the nasal duration.

The durational changes in the nasal consonant are explained in terms of anticipatory coarticulatory effects. The nasal preserves the closure in its articulation if it shares the same articulation gesture with the following consonant. This is the case when [n] is followed by a dentialveolar stop or affricate, as in [nd] and [nts] sequences. All consonant types involved here, the nasal as well as the stop or the affricate, have a closure in their articulation gesture. These originally distinct closures—the one that belongs to the nasal and the one that belongs to the following consonant—merge into one gestural movement, since normally the closure of the nasal is not released before the homorganic closure of the following stop or affricate.[2]

The motor command of the nasal closure is modified in the case of a following fricative consonant. Except in very careful pronunciation, the articulation of the nasal in this context is characterized by an early onset of the velar gesture and by the lack of articulatory closure. The blade of the tongue approaches the dentialveolar area, but never reaches the teeth or alveolar ridge. There should be a two-way effect on the *n*C consonant cluster: an anticipatory and a progressive one. On the one hand, the fricative consonant influences the nasal closure anticipatorily, while on the other hand the fricative itself is nasalized as a consequence of the progressive articulatory effect of the nasal. During this coarticulation the nasal consonant loses its closure; hence, the air is allowed to flow through both the oral and the nasal cavities. The consequence of all these effects and the modified articulation gestures in the speech chain is a shortened nasal consonant.

Since the fricatives and the approximant [j] share the same tongue gesture (or very similar tongue gestures) in their articulation, the central approximant is expected to show a similar coarticulatory effect on the preceding nasal. The experimental data confirm this expectation. The approximant [l] does have a closure, as well as two lateral apertures, simultaneously. In this case, the question arises as to which of the two features of [l] has a stronger effect on nasal duration: closure or aperture? The two apertures allow this consonant to be uttered continuously; in this sense, [l] is a continuant. The closure, however, makes this consonant similar to a noncontinuant consonant. In point of fact, our durational data of the nasal show that the lateral approximant behaves like a continuant consonant; see Figure 2. Thus, aperture is a more important feature than

[2] Since the place of articulation of [n] is in the middle of the place-of-articulation range, it is relatively easy to adjust its closure slightly to match that of a following consonant. As a consequence, the dentialveolar nasal easily assimilates in place of articulation (Vago 1980; Kenesei et al. 1998; Siptár–Törkenczy 2000, among others). Indeed, there are no word-internal instances of [n] not being homorganic to a following consonant (Siptár–Törkenczy *op.cit.*, 92, fn. 17).

closure in the articulation gesture of [l]. In sum, both approximants are like other continuant consonants in terms of their durational effects on a preceding nasal.

3.2. Classification of C in V*n*C sequences based on the duration of *n*

A model was established using a teaching algorithm to distinguish among the various types of V*n*C sequences. We used an approach to the classification of the consonant type that follows the nasal by means of the support vector machine (SVM) and the properties of Radial basis function (RBF) kernel, based on the duration of the vowel and the nasal (MATLAB version 7.0 was used).

A classification algorithm, the ROC (Receiver Operating Characteristic) curve analysis was used that shows that the duration of the nasal predicts whether the following consonant is a noncontinuant or a fricative. These curves show the percentage of hits (correct detection of a positive feature value) vs. the percentage of false alarms (incorrect detection of a positive feature value) for each possible threshold value. There are two points in each ROC curve—highest hit rate and lowest false alarm rate—that provide a threshold for the best classification. The points on the ROC curve occurring closest to top left corner (in a Euclidean sense [0;1]) were considered to be the optimal threshold. The result is a 'fair test' (meaning a good result) where the cutpoint turns out to be at 75.5 ms, and the result of the separation is AUC = 74.082%; see Figure 3 (overleaf).

We also assumed that the voicing of the following consonant could be determined on the basis of the duration of the nasal. The result is a 'poor test' (meaning that the result is poor) where the cutpoint turns out to be at 77.0 ms. The separation result is low: the AUC is 60.434%. This model shows that, based on the duration of the nasal, a much more reliable evaluation can be made as to whether the following consonant is a fricative or a noncontinuant consonant than whether it is voiced or voiceless.

3.3. Coarticulation effects on vowel duration

The length of the vowel in a V*n*C sequence is also affected by spreading coarticulation from the postnasal consonant: the vowel is shorter when

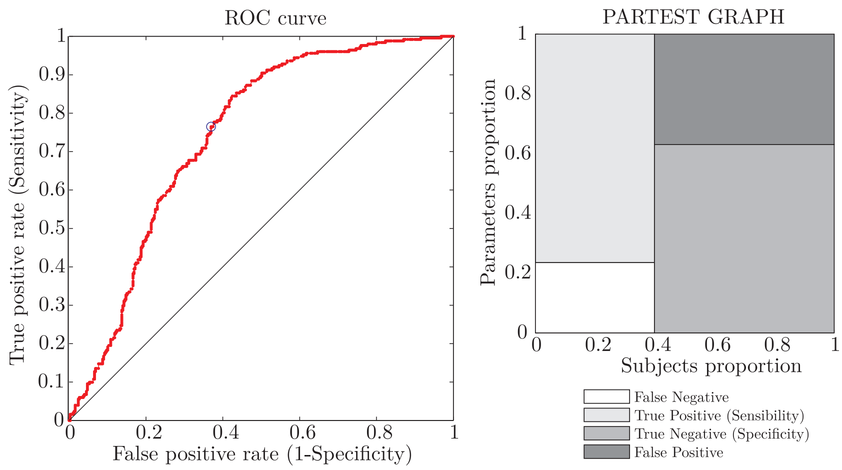

Fig. 3

The results of the ROC separation test based on the duration of the nasal for the judgment of the following consonant as a fricative or a noncontinuant consonant

the nasal is followed by a noncontinuant obstruent, and longer when the nasal is followed by a fricative (cf. Kavitskaya 2002). The mean duration of vowels in the case of postnasal fricatives is 185 ms, as opposed to 155.5 ms in the case of postnasal noncontinuants. The difference in spreading coarticulation effects on vowel duration is significant (one-way ANOVA: $F(3, 629) = 49.1$, $p = 0.001$), cf. Table 3. The effect of approximants on vowel duration is similar to the effect of fricatives. The mean vowel duration when the nasal is followed by palatal [j] is 173.31 ms, while it is 176.0 ms when the nasal is followed by the lateral [l] (see Figure 4). The Tukey post hoc tests among consonant groups confirmed that the duration of vowels in the case of postnasal approximants is significantly different from the case of postnasal noncontinuant obstruents.

Table 3

V*n*C sequences: The duration of V and the C types

	Duration of V (ms)	
C types	Mean	Std. dev.
Noncontinuant	156	29
Fricative	185	35
Approximant	175	34

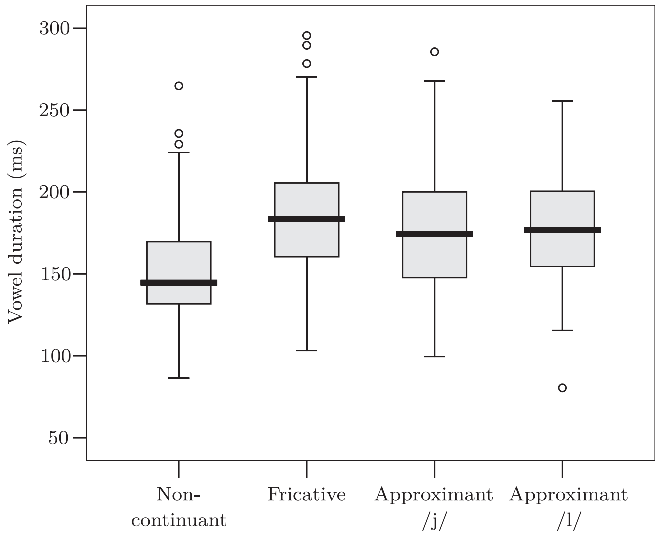

Fig. 4

V*n*C sequences: The duration of V and the C types (medians & ranges)

Vowel duration does not differ significantly depending on whether the nasal is followed by fricatives or by approximants. This means that the effect of these consonants on vowel duration is alike. Nor is there a significant difference between the two types of approximants as far as their effect on vowel duration is concerned.

Inherent articulation gestures also contribute to differences in vowel duration. Table 4 presents data on the duration of the various vowel types in V*n*C sequences.

Table 4

V*n*C sequences: The duration of the V types

	Duration of V (ms)|Standard deviation V types						
C types	[aː]	[eː]	[ø]	[ɔ]	[ɛ]	[iː]	[o]
Noncontinuant	181|31	150|26	156|25	140|17	142|25	143|28	144|30
Fricative	213|45	185|33	167|37	186|24	183|26	183|34	176|31
[j]	206|38	179|29	188|30	160|25	176|26	180|44	146|28
[l]	207|25	168|43	183|27	184|23	160|33	154|27	170|26

The mean duration of the class of back vowels is 172.27 ms (SD = 37.23 ms), while the mean duration of the class of front vowels is 165.4 ms (SD = 34.3 ms). The durational difference between the class of front vowels and the class of back vowels is significant ($F(1, 629) = 5.502$, $p = 0.019$); see also Figure 5.

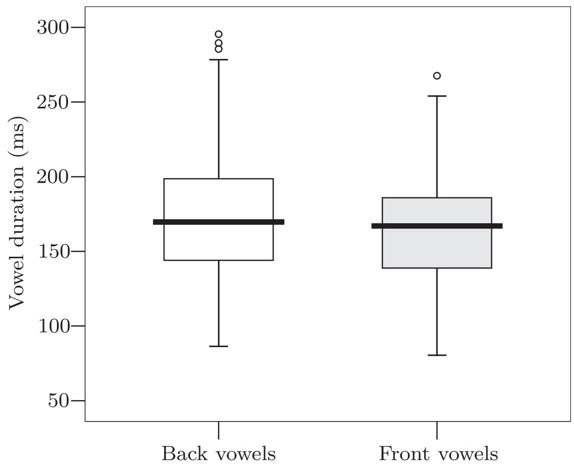

Fig. 5

V*n*C sequences: The duration of the front and back vowels (medians & ranges)

The coarticulation effect on vowel duration was also analyzed taking the phonological length of the vowels in question into account. The orthographic representations of the short and long vowels (*e* vs. *é, a* vs. *á*) were assumed to cue the subjects in pronouncing the isolated words. The mean duration of the long vowels is 177.18 ms (SD = 39.9 ms), while the mean duration of the short vowels is 163.4 ms (SD = 31.8 ms). Although we had expected to find objective durational differences in phonemically long and short vowel durations depending on the type of the following consonant, the data did not confirm this expectation. It is only the vowel [aː] that has also objectively long durations (mean value: 199.11 ms) as opposed to the durations of the phonologically short vowels (whose mean duration is 164.87 ms for [ɔ], 162.78 ms for [ɛ], 159.29 ms for [o], and 165.63 ms for [ø]). The speakers did not pronounce other phonemically long vowels as long as they did in the case of [aː] (cf. the mean duration values 167.19 ms for [eː], 162.43 ms for [iː]). Therefore, [aː] alone is responsible for the differences between the phonologically short and long vowels.

The duration of the vowels was further analyzed in relation to the voicing feature of the postnasal consonant. The data reveal that vowel duration is longer when a voiced consonant follows the nasal. In this case, the mean duration is 180.9 ms (SD = 31.8 ms), while the mean duration is 157.7 ms (SD = 36.5 ms) when a voiceless consonant follows the nasal; see Table 5. The difference is significant (one-way ANOVA: $F(1,628) = 71.8$, $p = 0.001$).

Table 5

V*n*C sequences: Details of V durations and the voicing of C

	Mean durations of V (ms)|Standard deviation V types						
C types	[aː]	[eː]	[ø]	[ɔ]	[ɛ]	[iː]	[o]
voiced	214|38	185|35	182|32	176|29	172|30	175|38	162|31
voiceless	183|35	152|24	151|25	153|28	155|31	153|33	156|34

We have analyzed the effects of four fixed factors on vowel duration (using UNIANOVA): the voicing and type of the postnasal consonant and the front/back dimension and the phonological length of the vowel. The results show that all the fixed factors have an effect on vowel duration (factor of phonological length of the vowel: $F(1,629) = 36.69$, $p = 0.001$; backness of the vowel: $F(1,629) = 7.109$, $p = 0.008$; voicing of the postnasal consonant: $F(1,629) = 16.71$, $p = 0.001$; postnasal consonant type: $F(1,629) = 78.88$, $p = 0.001$). The model explains 0.312% of all frames (R Squared = 0.329). The interaction among the factors is not significant, except for one case: the horizontal movement of the tongue and the phonological length of the vowels ($F(1,629) = 22.97$, $p = 0.001$). The F-ratio calculation shows that consonant type has the greatest effect on the classification of the postnasal consonant. This is confirmed by the Partial Eta Squared (PRE) ratio (PRE = 0.114).

3.4. Classification of the postnasal consonant based on vowel duration

The ROC separation curve shows that the duration of the vowel predicts whether the following consonant is a noncontinuant or a fricative. The result of the ROC separation curve test is ‘fair’, the cutpoint turns out to be at 168.67 ms, and the result of the separation is AUC = 71.134%; cf. Figure 6.

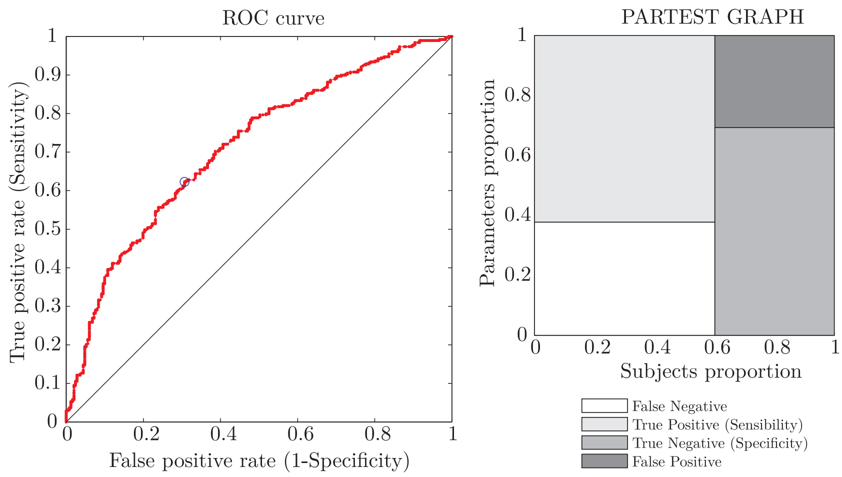

Fig. 6

The results of the ROC separation test based on vowel duration for the classification of the postnasal fricative or noncontinuant consonant

In case the voicing character of the postnasal consonants was to be recognized on the basis of the vowels' duration, the ROC separation test shows a 'poor' result. The cutpoint is at 171.5 ms; the separation result is AUV= 68.776%. This means that the voicing character of the postnasal consonant cannot be predicted by means of the duration of the vowels that precede the nasal.

3.5. Classification of the postnasal consonant based on the ratio of vowel and nasal duration

We supposed that the duration of the vowel and the nasal together would be close to a constant value and the ratios would show differences depending on the type of the postnasal consonant. The assumption has partly been confirmed. The vowel + nasal durations are significantly different (one-way ANOVA: ($F(2, 628) = 28.9992$, $p = 0.001$), and do not show a constant-like value (Figure 7). This means that the changing durations of the vowels and the *n* did not equalize each other.

The ratio of vowel and nasal duration refers to the double temporal modifications of the V*n* sequence depending on the postnasal consonant (this kind of ratio in percentage terms was obtained by means of dividing the vowel duration by the nasal duration and multiplying it by 100). The V*n* duration mean ratio is 33% before the fricatives (SD = 15.2%)

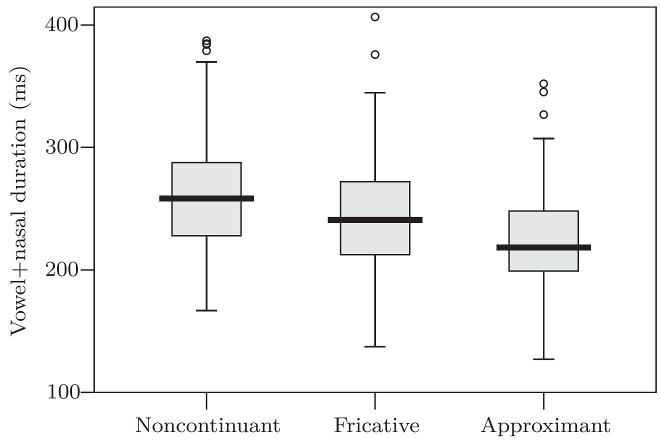

Fig. 7

The vowel + nasal durations depending on the postnasal consonant (medians & ranges)

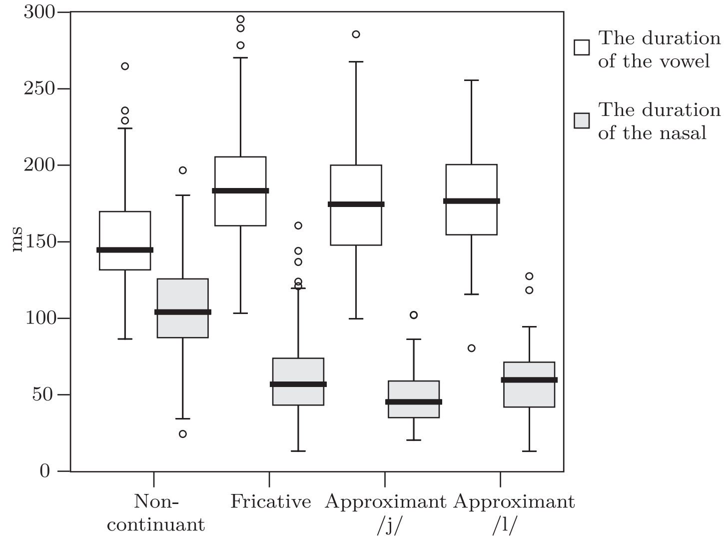

Fig. 8

VnC sequences: Durational data of n and V as a function of C (medians & ranges)

while the mean ratio is 73.2% (SD = 22.2 ms) before the noncontinuants. The mean ratio of the duration of the Vn sequence is 27.2% (SD = 10.9%) before [j] while it is 34.7% (SD = 15.7%) before [l]. These differences turned out to be significant (one-way ANOVA $F(3, 629) = 257.2$, $p = 0.001$). The Tukey post hoc tests confirmed again that there are no significant differences either between the fricatives and the approximants or between the two approximants. The ratios of the vowel/nasal durations provide a basis for the statistical evaluation of the postnasal consonant type for classification; see Figure 8.

The vowel + nasal durations were also analyzed according to the voicing character of the postnasal obstruents (cf. Figure 9). Statistical analysis confirms that the difference is significant ($F(1, 502) = 48.6$, $p = 0.001$). The V*n* sequence is shorter before voiceless obstruents than before voiced ones.

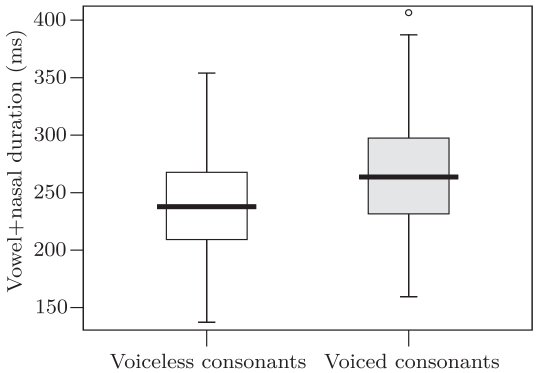

Fig. 9

The vowel + nasal durations depending on the voicing character of the postnasal obstruents

The mean ratio of vowel/nasal durations is significantly higher when the postnasal consonant is voiced than when it is voiceless ($F(1, 629) = 37.3$, $p = 0.001$). The mean ratio in the former case is 55.23% (SD = 26.32%), in the latter case 42.4% (SD = 26.2%); see Figure 10.

The ROC separation curve, however, resulted in 'poor tests' in both cases using the mean ratio data. The mean ratio of vowel/nasal durations shows that the cutpoint is at 43.87 ms; the separation result is AUC = 65.188%. This means that neither the type nor the voicing character of the postnasal consonant can be evaluated based on these ratios.

4. Automatic classification of consonant types

The second goal of this study was to verify that vowel and nasal durations can be used for the automatic classification of the consonant types following the nasal in the V*n*C sequences analyzed here. The SVM classifier

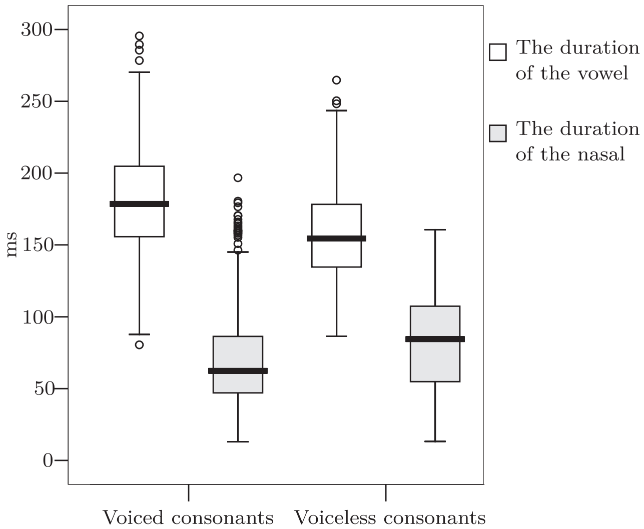

Fig. 10

V*n*C sequences: Durational data of *n* and V and the voicing of C (medians & ranges)

algorithm was selected for the reason that this is a widely and successfully used algorithm in speech classification systems. SVMs have a relatively good generalization capability, with less requirement of the amount of training data, and they have been particularly well developed for binary classification tasks. Furthermore, they are scalable for high dimensional data without a corresponding increase in the number of training samples. The experiments were carried out using the SVMlight toolkit (cf. Joachims 1999).

The training data were selected randomly, based on 2/3 of the frames in each of two consonant types: fricatives and noncontinuants. The testing data differed from those of the training data in that the remaining 1/3 of the frames were used as well. The training of the SVM classifiers was done using (i) two parameters (vowel duration and nasal duration) and (ii) three parameters (vowel duration, nasal duration, and ratio of the two durations). The SVM classifiers were used for training with Radial Basis Function (RBF) kernels. The accuracy of an SVM model is largely dependent on the selection of the model parameters. In this study two methods were used—a 3-multiple cross-validation and a grid search—for finding optimal parameter values (gamma and kernel), Figure 11. Huang et al. (2004) proposed to use the C and γ values in the following ranges: C: $\{-5^{\wedge 2}; -3^{\wedge 2}; \ldots; 13^{\wedge 2}; 15^{\wedge 2}\}$ and γ: $\{-15^{\wedge 2}; -13^{\wedge 2}; \ldots; 1^{\wedge 2}; 3^{\wedge 2}\}$. The statistical results show that automatic classification results in 94.39% correct classification for all data of the stricture feature of the postnasal consonant. 59.28% correct classification is reached for all data of the voicing feature of the same consonant (as it can be detected on the darkest space of the figure, close to 1 = 100%).

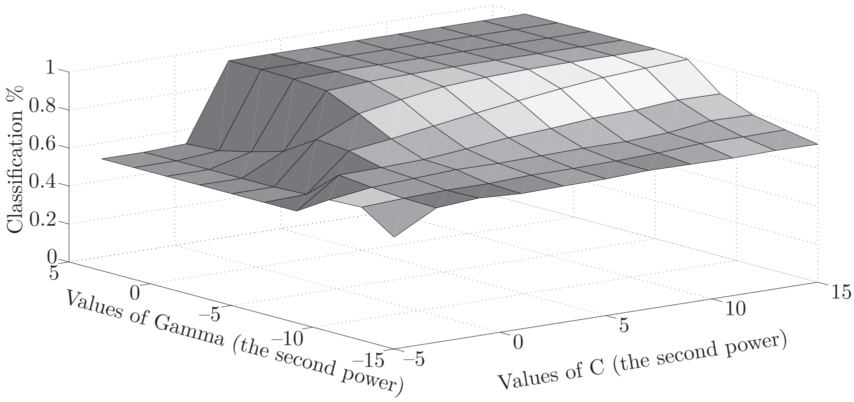

Fig. 11

V*n*C sequences: classification ratio of the fricatives and noncontinuants by means of 3-cross-validation

The fewer parameters used, the better the classification results for the stricture of the postnasal consonant. There is no such difference in the classification results for the voicing of the postnasal consonant (Table 6).

Table 6

V*n*C sequences: The classification matrices for the classification of the C types using two or three parameters

	Classification matrix (%)			
	Using two parameters		Using three parameters	
C type	Fricative	Noncontinuant	Fricative	Noncontinuant
Fricative	93.75	6.25	61.33	38.66
Noncontinuant	5.15	94.84	23.62	76.37
Voicing character	Voiced	Voiceless	Voiced	Voiceless
Voiced	67.85	32.14	70.0	30.0
Voiceless	49.28	50.71	50.71	49.28

5. Conclusions

In this study on anticipatory coarticulation within Hungarian V*n*C sequences, two main issues were addressed. The first one concerned the durational effects of C on both the preceding nasal and vowel; the second one concerned the possibility of the automatic detection of the C type based on the temporal data. It appears that in the V*n*C sequences analyzed, anticipatory coarticulation emanating from a postnasal fricative affects the duration of both the nasal and the vowel. The coarticulatory effect is also confirmed for approximants; the effect is similar to that found with fricatives. The articulatory structures are characterized by longer gesture in the case of vowels and shorter gesture in the case of the nasal if a fricative or an approximant follows. Conversely, vowels have shorter gestures and the nasal has a longer gesture if a noncontinuant obstruent follows. Our data do not provide sufficient information concerning whether the timing modification of the vowel is the result of the modified temporal gesture of the nasal itself or of the C type showing the spreading coarticulation effect. Right-to-left coarticulation effects and the observed spread of coarticulatory effects over several segments are difficult to explain (cf. Tatham–Morton 2006). We assume that the V*n*C sequences show a contextually determined coarticulatory adjustment that appears across syllable boundaries and might be a specific property of the motor control system.

Both the vowel and the nasal in our V*n*C sequences have significantly longer duration before voiced obstruents than before voiceless obstruents. This might be attributed to the anticipation of glottal opening before voiceless consonants (for similar results, see Farnetani–Recasens 1993).

The present results bear also on the artificial classification of speech. The model developed confirms that the duration of the vowel and the nasal is of crucial importance for the classification of the following consonant as to whether it is a fricative, approximant or noncontinuant.

References

Beddor, Patrice S. 2007. Nasals and nasalization: The relation between segmental and coarticulatory timing. In: Trouvain – Barry (2007, 249–55).

Beke, András – Viktória Horváth 2009. A nazális koartikuláció variabilitása a spontán beszédben [On the variability of nasal coarticulation in spontaneous speech]. In: Mária Gósy (ed.): Beszédkutatás 2009 [Speech research 2009], 28–45. MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium, Budapest.

Bell-Berti, Fredericka 1993. Understanding velic motor control: Studies of segmental context. In: Marie K. Huffman – Rena A. Krakow (eds): Nasals, nasalization, and the velum, 63–85. Academic Press, New York.

Boersma, Paul – David Weenink 2005. Praat: Doing phonetics by computer. (Version 5.4) [Computer program]. Retrieved from http://www.praat.org/.

Boyce, E., Susan – Rena A. Krakow – Fredericka Bell-Berti – Carol E. Gelfer 1990. Converging sources of evidence for dissecting articulatory movements into gestures. In: Journal of Phonetics 18 : 173–88.

Busà, Grazia M. 2007. Coarticulatory nasalization and phonological development: Data from Italian and English nasal–fricative sequences. In: Marie J. Solé – Patrice S. Beddor – Manjari Ohala (eds): Experimental approaches to phonology, 311–36. Oxford University Press, Oxford.

Chafcouloff, Michael – Alain Marchal 1999. Velopharyngeal coarticulation. In: Hardcastle – Hewlett (1999, 69–79).

Chen, Nancy F. – Janet L. Slifka – Kenneth L. Steven 2007. Vowel nasalization in American English: Acoustic variability due to phonetic context. In: Trouvain – Barry (2007, 905–8).

Delvaux, Veronique – Angelique Amelot – Solange Rossato 2007. Nasal processes at the interface between phonetics and phonology. In: Trouvain – Barry (2007, 245–8).

Farnetani, Edda – Daniel Recasens 1993. Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. In: Language and Speech 36 : 279–303.

Farnetani, Edda – Daniel Recasens 1999. Coarticulation models in recent speech production theories. In: Hardcastle – Hewlett (1999, 31–65).

Gósy, Mária 2004. Fonetika, a beszéd tudománya [Phonetics, the science of speech]. Osiris Kiadó, Budapest.

Gósy, Mária – Robert M. Vago 2009. A case of grounded compensatory lengthening. Poster presented at the Ninth International Conference on the Structure of Hungarian, Debrecen, Hungary.

Hardcastle, William – Nigel Hewlett (eds) 1999. Coarticulation. Cambridge University Press, Cambridge.

Horváth, Viktória 2005. A magánhangzók nazalizációjáról [On the nasalization of vowels]. In: Mária Gósy (ed.): Beszédkutatás 2005 [Speech research 2005], 51–62. MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium, Budapest.

Horváth, Viktória 2008. Nazális hatás a magánhangzóképzés függvényében [An analysis of the nasal effect depending on vowel type]. In: Mária Gósy (ed.): Beszédkutatás 2008 [Speech research 2008], 22–31. MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium, Budapest.

Huang, Zan H. – Hsinchrum Chen – Chia-Jung Hsu – Wun-Hwa Chen – Soushan Wu 2004. Credit rating analysis with support vector machine and neural networks: A market comparative study. In: Decision Support Systems 37 : 543–58.

Joachims, Thorsten 1999. Making large-scale SVM learning practical. In: Bernhard Schölkopf – Chris Burges – Alexander J. Smola (eds): Advances in kernel methods – support vector learning, 73–91. MIT Press, Cambridge MA.

Kavitskaya, Darya 2002. Compensatory lengthening: Phonetics, phonology, diachrony. Routledge, New York.

Kenesei, István – Robert M. Vago – Anna Fenyvesi 1998. Hungarian. Routledge, London & New York.

Ladefoged, Peter 1993. A course in phonetics. Harcourt, Brace and Jovanovich, Fort Worth, Texas.

Laver, John 1994. Principles of phonetics. Cambridge University Press, Cambridge.

Maddieson, Ian 1984. Patterns of sounds. Cambridge University Press, Cambridge.

Malécot, André 1960. Vowel nasality as a distinctive feature in American English. In: Language 26 : 222–9.

Ohala, John J. 1984. Phonological evidence for top-down processing in speech perception. In: Joseph Perkell – Dennis H. Klatt (eds): Invariance and variability in speech processes, 386–401. Lawrence Erlbaum, Hillsdale NJ.

Ohala, John J. – Grazia M. Busà 1995. Nasal loss before voiceless fricatives: A perceptually-based sound change. In: Rivista di Linguistica 7 : 125–44.

Olaszy, Gábor 2007. Mássalhangzó-kapcsolódások a magyar beszédben [Consonant clusters in Hungarian]. Tinta Könyvkiadó, Budapest.

Pickett, James M. 1999. The acoustics of speech communication: Fundamentals, speech perception theories, and technology. Ally and Bacon, Boston & New York.

Siptár, Péter – Miklós Törkenczy 2000. The phonology of Hungarian. Oxford University Press, Oxford.

Tatham, Mark – Katherine Morton 2006. Speech production and perception. Palgrave Macmillan, New York.

Trouvain, Jürgen – William J. Barry (eds) 2007. Proceedings of the 16th International Congress of the Phonetic Sciences. Saarbrücken University, Saarbrücken.

Vago, Robert M. 1980. The sound pattern of Hungarian. Georgetown University Press, Washington DC.

Whalen, Douglas H. – Patrice S. Beddor 1989. Connections between nasality and vowel duration and height: Elicitation of the Eastern Algonquian intrusive nasal. In: Language 65 : 457–86.