Durational aspects of fast speech in German

Reinhold Greisbach[1,2], Kornélia Juhász[5,4,3], Zsuzsa Szánthó[3,4,5], Szabina Zsoldos[3], Andrea Deme[3,4,5]

[1]IfL-Phonetics, University of Cologne, Germany
[2]Work Area Psycholinguistic Appr. in Rehabilitation Sciences, University of Cologne, Germany
[3]ELTE Eötvös Loránd University, Hungary
[4]MTA–HUN-REN NYTK Lendület "Momentum" Neurophonetics Research Group, Hungary
[5]HUN-REN Hungarian Research Centre for Linguistics, Hungary

reinhold.greisbach@uni-koeln.de, juhasz.kornelia@nytud.hun-ren.hu, zsuzsa.szantho@gmail.com, szamboc@gmail.com, deme.andrea@btk.elte.hu

Abstract

It is debatable whether the phonetic manifestation of the central phonological contrast in the German vowel system is a quantitative (vowel length) or qualitative contrast. An acceleration of articulation has an effect on the duration of the vowels and thus on the quantitative contrast. Acceleration can also lead to greater centralization and thus influence the qualitative contrast. The investigation of fast speech is intended to provide information on which parameter is more resistant and can therefore considered to be the distinctive feature. In the experiment, the corner vowels of 19 native German speakers were examined in normal and fast speech using minimal pairs and analyzed with regard to their duration and quality parameters (formants). A new algorithm for normalizing the formant frequencies is presented. The results go both ways. The qualitative contrast is more robust to speeding up speech. However, the quantitative contrast seems to be more important as it extends from the vowels to all sounds of the analyzed minimal pairs. The two strategies are not correlated, which indicates that the speakers use different individual strategies to increase their speaking rate.

Index Terms: speech rate, vowel length contrast, fast speech, vowel quantity, vowel triangle normalization

1. Introduction

Phonetic and phonological descriptions of the sound inventory and sound structure of a language are based on precise pronunciation by speakers of that language or (mechanically speaking) articulation at a slow rate of speech. Words that are transcribed in pronunciation dictionaries or in books on foreign language learning are in most cases only the slowly and clearly spoken version of a word.

But this does not reflect the situation in real speech. According to Lindbloom's H&H theory [15] speakers adapt to the communication situation by constantly changing their speech production. The effects of these adaptations may be ordered along the time axes, on a scale of more or less information transfer.

To transfer more information in the same portion of time the speakers may employ different strategies. They may speed up articulation, which physically means that the articulators (lips, tongue tip, tongue body, velum) must move faster to produce more sounds in a shorter period of time. Another strategy would be omitting some articulatory movements which would lead, as a consequence, to missing sounds in the flow of speech ("elision"). The third strategy is a reduction of magnitude of the articulatory movements which may lead to a less clear pronunciation. If the intended articulatory goal is not achieved, the sound quality produced will change and with it the perceived sound quality [14]. Examples of this are "spirantization" for plosives or "centralization" for vowels.

However, the H&H theory says nothing about what the reductions in articulatory movements look like in specific cases and what acoustic consequences they have. In our study, we investigate the effects of increased speech rate on the central distinctive feature of the German vowel system, which could be called vowel length: In what way is the contrast between long and short vowels maintained when the speaking rate increases?

2. German vowel system

In German, almost all of the vowels in accented syllables can be combined into pairs. These pairs differ in duration and, in most cases, also by an additional phonetic feature, so that we have the pairs /ɪ/-/iː/, /ʏ/-/yː/, /ʊ/-/uː/, /ɔ/-/oː/, /œ/-/øː/, /ɛ/-/eː/ and /a/-/aː/.

In German phonology, in addition to length, 3 other parameters are discussed that may be considered the primary feature of the contrast from which the others are predictable. This way, there are 4 possible features discussed (e.g. [8]):

- Quantity (based on the auditorily perceptible duration: long vs. short)
- Quality (based on the auditorily perceptible sound quality: peripheral vs. centralised, schwa-like)
- Tenseness (based on the proprioceptively perceptible articulation effort, perceived by the speaker as muscle tension when speaking: tense vs. lax)
- Syllable cut (based on the auditorily perceptible end of a vowel, as smoothly fading out (soft cut, loose connection) vs. abrupt ending (sharp cut, fixed connection).

It is important to emphasize that all these features are assumed to capture one single phonologically relevant contrast between the two vowel groups.

10.21437/ISAPh.2024-6

In German phonology, one of these parameters is, therefore, usually considered as most important and it forms the so-called distinctive feature. All of the other differences that can be found between the pairs are predictable, hence redundant, and of secondary importance.

Phonetic analyses on this topic have so far primarily been carried out to determine which phonetic feature speaks in favor of the choice of one of these 4 parameters as a distinctive feature of the vowel length difference but these attempts have ultimately lead to inconclusive results.

In our study, we want to find out which of these parameters (to be more precise: quantity or quality) is more resistant to an increase of articulation speed, which would point to this parameter being the distinctive feature of the contrast. Due to the great number of vowel pairs we have in German, we focus only on the corner vowel pairs (Figure 1).

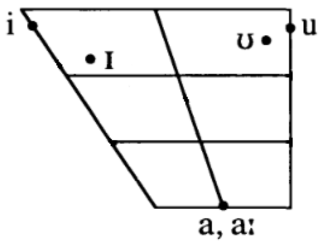

Figure 1: *German vowel pairs according to [5]. (Only the corner vowels are displayed.)*

Figure 1 shows the German corner vowel pairs according to the description of the German phonetic system based on phonetic transcription by [5]. The vowel pairs differ in sound quantity (only noted for [a:] in Figure 1) and sound quality. According to [5], the greatest difference in quality is found in the /i/-vowels, while the /a/-vowels do not differ in quality.

3. Experimental Design

In our study we focus on the corner vowels of the vowel triangle, which are embedded into [C_1V(:)C_2ə]-shaped real words of German with 2 syllables. In these words, the target vowels are placed in the first syllable, which is accented. The vowel of the second syllable is a schwa, thus forming the typical structure of a 2 syllabic native word of German.

We used minimal pairs to reduce variational effects of the surrounding consonants on the opposition, a fact which was already observed by [16] in 1960. We choose the pairs *Schiefe-Schiffe* [ʃi:fə] – [ʃɪfə] (‘skew’ – ‘ships’), *biete-bitte* [bi:tə] – [bɪtə] (‘offer’ – ‘please’) for the /i/-vowels, *Maße-Masse* [ma:sə] – [masə] (‘measures’ – ‘mass’) , *wate-Watte* [va:tə] – [vatə] (‘wade’ – ‘cotton wool’) for the /a/-vowels and *Buße-Busse* [bu:sə] – [bʊsə] (‘penance’ – ‘buses’) , *Muße-Musse* [mu:sə] – [mʊsə] (‘leasure’ – ‘must’ (pl.)) for the /u/-vowels. In the last case, we had the rather unusual plural form of the word *Muss* [mu:s] (‘must’), like in a phrase *Es ist ein Muss, dass …* (‘It is a must, that …’).

There were 2 minimal pairs for every vowel, which totalled in 12 words. The consonants were not balanced but chosen due to existing minimal pairs in German. The minimal pairs were embedded into a carrier phrase *Mehr* <target word> *bitte!* (‘More <target word> please!’). The phrases were repeated 12 times in random order by each speaker, resulting in a total of 144 phrases per speaker.

19 native German speakers (10 females and 9 males aged 22 to 63) from the western parts of Germany were recorded, with no dialectal or regional characteristics. For the first set of recordings, the speakers were asked to speak at their normal rate. For the second set, they were asked to speak as quickly as possible. The experimenter conducted a short training session in which he encouraged the speakers to speak as quickly as possible.

The acoustic analysis was performed manually using PRAAT [17]. Only the 6 fastest articulations of every word were evaluated for the ‘fast speech rate’ in each speaker.

For the operationalization of the velocity of speech, typically longer stretches of speech are measured and an average is taken. Terminologically a difference is made between speech including pauses (speech rate) and speech excluding pauses (articulation rate), see e.g. [9]. In our experimental set up using short phrases, the articulation rate seems to be an appropriate measure to quantify speed. To be comparable with the data in literature, we divided the number of syllables in the phrase (5) by the total duration of the phrase to get the syllable rate per second.

We measured and analyzed vowel duration to quantify length, and the formant frequencies F_1 and F_2 to capture vowel quality.

In order to be able to compare the contrasts between long and short vowels in both the temporal and spectral domain, we decided to normalize all measured values by setting the raw values of the long vowels at normal speech rate to 100% for each vowel. In this way the contrast between short and long and between normal and fast vowels are operationalized as ratios between the measured values in all cases.

4. Results

4.1. Syllable rates

There were almost no elisions, assimilations or lenitions observed in fast speech. That is, the durational reductions must have been primarily caused by a faster movement of the articulators. On average, the articulation speed increased from 5.69 syllables/second to 7.67 syllables/second which corresponds to an acceleration to 73.8 % of the time relative to the normal articulation speed.

Thus the average syllable rate for normal speed is in line with the documented syllable rates for German extracted from read texts, interviews and storytelling, which range from 5.40 syllables/second [13] (as cited in [12]) respectively 5.50 syllables/second [10] to 5.97 syllables/second [11].

However, our syllable rate differed very strongly between speakers and ranged from 4.44 syllables/second to 6.58 syllables/second at normal speed and from 6.18 syllables/second to 9.94 syllables/second in fast speech. Acceleration also varied greatly between 96% of the time for normal speech (almost no change) and 58% (almost halving the duration of the words in fast speech).

Figure 2 shows the distribution of the syllable rates for normal speed (*y*-axis) and fast speed (*x*-axis). Observations during the experiment point to the fact that some speakers were not able to produce speech faster, at least, in this experimental setup. We observed that instead of getting faster, some speakers (e.g. S3 and S4) were getting louder.

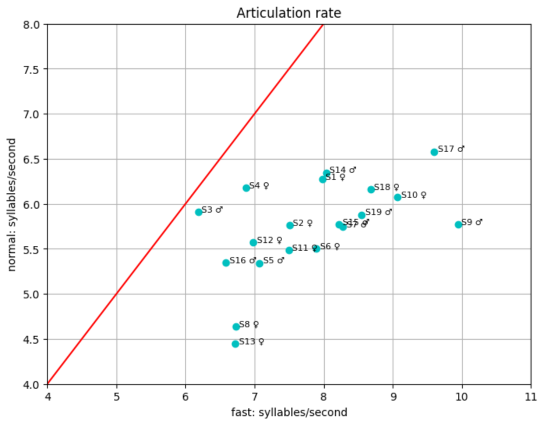

Figure 2: *Syllable rate of the speakers for fast (x-axis) vs. normal (y-axis) speech.*

4.2. Vowel duration

The average vowel duration of the corner vowels shows a systematic trend, in that the /a/-vowels (short as well as long and normal as well as fast variants) were the longest compared to the /i/-vowels and /u/-vowels (Figure 3). As different muscle groups are active in the production of these vowels, different timing could be expected. Comparable values are documented for German (e.g. [3]) and English (e.g. [4]).

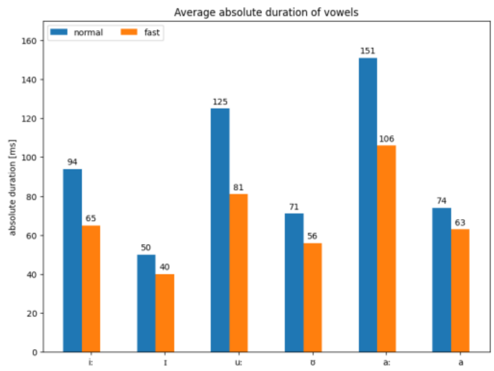

Figure 3: *Average absolute vowel durations.*

Setting the absolute durations of the normally spoken long vowels to 100% in each case, we get the relative durations, which remain relatively stable between the vowel groups (Figure 4). The /u/-vowels were affected the most by the acceleration, while the /a/-vowels were affected least. On average, the acceleration of the long vowels was considerably stronger (68.0%) compared to the short vowels (81.3%) in comparison with the duration of the normally spoken vowels. These values are relatively symmetrical around the average acceleration value for the whole phrase of 73.8%, which on the other hand corresponds approximately to the average of the acceleration of the other sounds in the minimal pair [C_1V(:)C_2ə] (75.2% for C_1 and 74.8% for C_2 as well as 75.1% for [ə]). Therefore, we find the following hierarchy of reduction of time caused by acceleration: long vowels > consonants & schwa > short vowels.

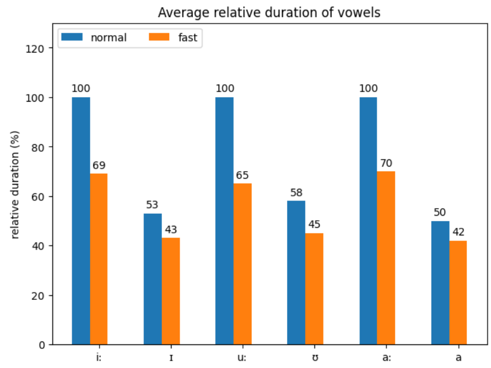

Figure 4: *Average relative vowel durations.*

Phonologically, it is primarily the duration contrast between long and short vowels which is of interest. This duration contrast is captured here as the long vowel to short vowel ratio. This was 1.90 on average at normal speed, and only 1.58 at fast speed. The relative duration contrast between long and short vowels was, therefore, reduced when speakers produced speech faster.

We also correlated the average articulation rate to the duration of the vowels (Figure 5). Since the articulation rates overlapped between normal and fast speaking rates (i.e. the fastest "normal" speaker was faster than the slowest "fast" speaker), we did not differentiate between fast and normal speaking rate in this regression analysis.

Regression analysis showed that the vowel contrast between long and short vowels reduced as the articulation rate increased.

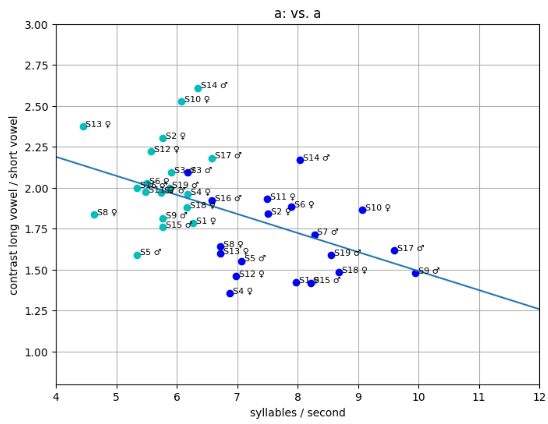

Figure 5: *Regression of vowel duration contrast (q = long vowel / short vowel) on syllable rate for normal and fast speaking rates in the* [a:] *vs.* [a] *vowel pair.*

Inspecting the displayed data in Figure 5 shows that beyond statistical significance of the linear regression (see Table 1) the realistically reachable minimal duration ratio of long to short vowels may be in the range of 1.25 which is obtained for some speakers already at a syllable rate of about 7 syllables/second irrespective of the vowel quality. Whether the contrast will neutralize effectively between 12 and 14 syllables/second as the regression analysis suggests may be disputable, since it is not clear if a speech rate of more than 12 syllables/second could be achieved. In any case, we see that the contrast for the /u/-vowels is the lowest compared to the others.

Table 1: *Ratio (q) of long vowel to short vowel at 6 and 10 syllables/second, determination coefficient R^2 of the linear regression and p-values*

Length contrast	***q* at 6 syl/s**	***q* at 10 syl/s**	**R^2**	***p*-value**
[a:] vs. [a]	1.956	1.491	0.252	0.0013
[i:] vs. [ɪ]	1.836	1.512	0.108	0.0442
[u:] vs. [ʊ]	1.705	1.278	0.228	0.0024

4.3. Extension of the length contrast

If the duration of all sounds in the minimal pairs with a long vowel [C_1V:C_2ə] is set to 100%, an effect is observed that goes beyond the expected duration contrast of the vowels. As Figure 6 shows there are significant contrasts for every sound in the minimal pairs of [C_1V:C_2ə] vs. [C_1VC_2ə]. That is, the contrast between long and short vowel is reflected not only in the ratio between long and short vowels (q = 1.90), but also in the ratio of the onset consonant C_1 of the long vowel minimal pair word in relation to the (phonologically identical) onset consonant C_1 of the short vowel minimal pair word (q = 1.08). The sounds of the 2nd syllable show an inverse but likewise significant contrast ratio of q = 0.96 for the consonant C_2 and of q = 0.85 for the [ə] (Figure 6).

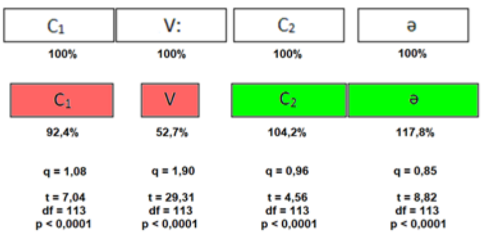

Figure 6: *Averaged relative duration differences in the sounds of the minimal pairs [C_1V:C_2ə] (top, white) vs. [C_1VC_2ə] (bottom, colored) for normal speech.*

We found similar differences between the speech sounds in the fast speech rate condition. (Figure 7).

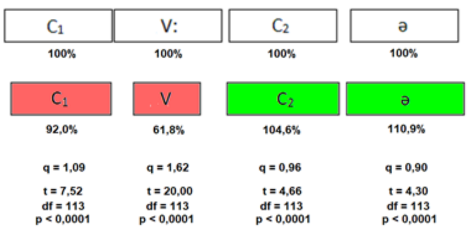

Figure 7: *Averaged relative duration differences in the sounds of the minimal pairs [C_1V:C_2ə] (top, white) vs. [C_1VC_2ə] (bottom, colored) for fast speech.*

According to paired *t*-tests (6 word pairs × 19 speakers), these differences were significant (Figure 6 & 7). These results show that the phonological opposition of the German vowels was also reflected in the relative durations of the surrounding sounds within the minimal pairs, regardless of the rate of speech. This means that the quantity contrast extends beyond the boundaries of the actual (short and long) vowels at hand, and lead to a different temporal organisation of the speech sounds in the entire minimal pair.

4.4. Vowel quality

The phonological contrast distinguishing the pairs /iː/-/ɪ/, /uː/-/ʊ/, and /aː/-/a/ in German is not only represented on the surface by a durational contrast but, in the case of the /i/- and /u/-pairs, by a spectral contrast (i.e. vowel quality) as well. Typically, to quantify vowel quality in the acoustic domain, the frequency the first two formants, F_1, and F_2 are estimated.

These measures were used in the present study, as well. The primary problem arising when we aim to compare and average formant frequency values is the fact that these frequency values are affected heavily by speaker characteristics, especially those that are different between male and female speakers (e.g., body height, vocal tract length, vocal fold length) [1,2].

Due to the specificities of our experimental dataset, in the present study we applied a new procedure which adapted to our data best. For the purposes of this method, we assume that the long vowels of a speaker at normal speech rate constitute the corners of this speaker's personal vowel triangle. Then, we put these long vowels in the corners of an equilateral triangle, with the origin as centre of gravity (Figure 8). The distances between the origin and the corners of this triangle are 1 (or 100%.).

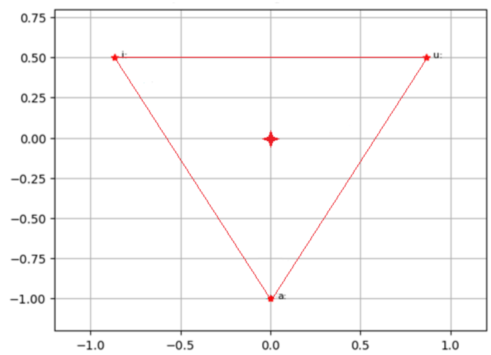

Figure 8: *Normalized reference vowel triangle*

As a next step, we calculate the position of every other vowel of the same speaker (i.e., short vowels at normal speech rates, and long and short vowels at fast speech rates) as linear combinations of the corner vowels (Figure 8). To be precise, the algorithm comprises 3 steps:

(1) We take the median values of the 6 or 12 repetitions for each vowel, and for each speaker. Then, we calculate the coefficients α, β, γ using the following equations:

$$\alpha * F1_{[i:]} + \beta * F1_{[u:]} + \gamma * F1_{[a:]} = F1_{vowel}$$
$$\alpha * F2_{[i:]} + \beta * F2_{[u:]} + \gamma * F2_{[a:]} = F2_{vowel}$$
$$\alpha + \beta + \gamma = 1$$

where $F1_{[i:]}$ means the F_1 frequency of the long vowel [i:] at normal speech rate etc.

(2) Next, we place the long corner vowels on the corners of an equilateral triangle with the centre positioned at (0,0):

$$P_{[i:]} = \left(-\frac{\sqrt{2}}{2}, \frac{1}{2}\right),\ P_{[u:]} = \left(\frac{\sqrt{2}}{2}, \frac{1}{2}\right),\ P_{[a:]} = (0,-1)$$

We calculate the position of the vowel (x_{vowel}, y_{vowel}) in the normalized vowel triangle for each vowel (median over 6 or 12 repetitions) of a speaker using the coefficients α, β, γ with the following equations:

$\alpha * x_{[i:]} + \beta * x_{[u:]} + \gamma * x_{[a:]} = x_{vowel}$
$\alpha * y_{[i:]} + \beta * y_{[u:]} + \gamma * y_{[a:]} = y_{vowel}$

(3) Lastly, we calculate spectral reduction, or “centralization” as the Euclidean distance of the projected vowel point from the origin using the following equation:

$$Dist = \sqrt{x_{vowel}^2 + y_{vowel}^2}$$

where Dist = 1 (or 100%) for the long vowels produced at a normal speech rate.

To get an average value, we average across the vowel locations of all speakers in their normalized vowel triangle (Figure 9).

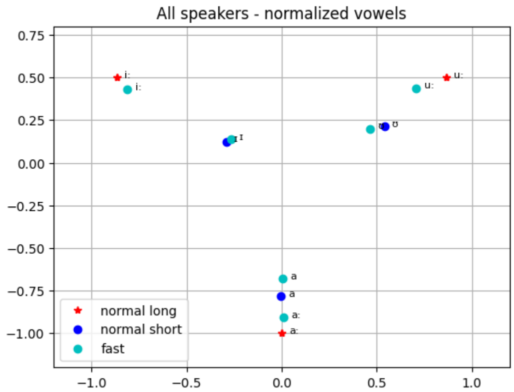

Figure 9: *Average relative position of the vowels in the normalized vowel space.*

Based on the above, relative expansion for the long vowels at normal speed is predefined as 100%, and we can see that the relative expansion of the short vowels differ considerably (Figure 10).

As was claimed by [5] (Figure 1) based on his auditory transcriptional impression, the short /i/-vowel is most centralized, followed by the short /u/-vowel. Contrary to his claim there is a quality difference between /a:/ and /a/ as well. The short /a/-vowel is centralized with respect to the long /a:/-vowel. This fact is also attested in the data of [7].

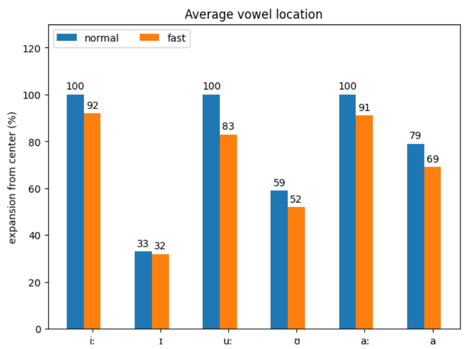

Figure 10: *Average relative vowel expansion (i.e., distance of vowels to the origin in the normalized vowel triangle shown in Figure 9).*

As in the case of length contrast, we can also perform a regression analysis here. Contrary to what we observed in the case of length contrast, there is no correlation between articulation rate and relative expansion (Table 2). That means that in speeding up, the speakers (on average) did not change their vowel quality contrast. For this reason, we should assume that spectral features are more stable cues of the vowel contrast at hand than durational cues.

Table 2: *Determination coefficient, Pearson’s correlation and significance for the quality contrast (relative vowel expansion).*

Expansion contrast	R^2	***r***	***p*-value**
[a:] vs. [a]	0.056	0.237	0.1512
[i:] vs. [ɪ]	-0.041	-0.203	0.2199
[u:] vs. [ʊ]	0.051	0.227	0.1700

4.5. Quality to quantity correlation

Although we have seen that the spectral contrast did not change as a function of speech rate, this does not mean, that the speakers did not change the quality of the vowels at all. Speaker S12, for instance, shifted the closed long vowels towards the midpoint of the vowel space, so that the formants of the fast long [u:] resembled the formants of the short [ʊ] produced in normal speech (Figure 11). By contrast, her fast [a:] production exceeded the distance of 100%, that is, she expanded her normal speech vowel triangle, similar to some other speakers observed previously by [6].

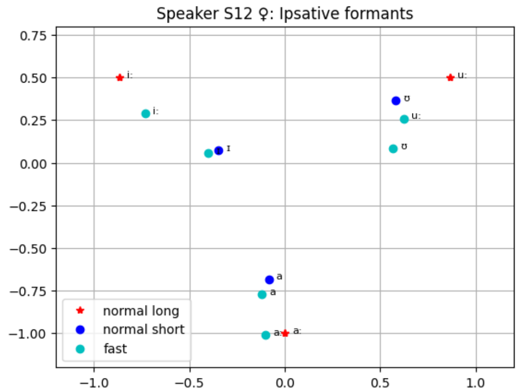

Figure 11: *Ipsative formant chart of speaker S12.*

By contrast, speaker S13 did not change the /i/-vowels at all, while all other vowels showed a tendency towards centralization while keeping the contrast distance (Figure 12).

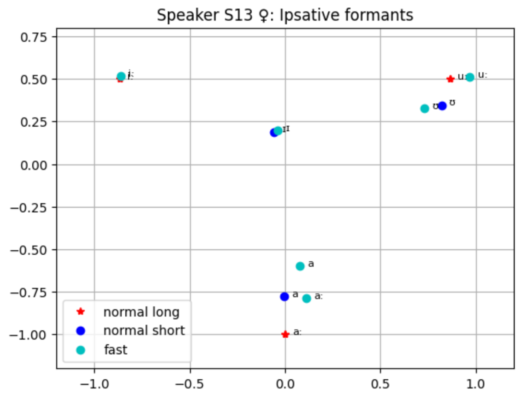

Figure 12: *Ipsative formant chart of speaker S13.*

Given this fact, it might be assumed that speaking faster either reduces the duration of the vowels or enhances centralization. To check this, we calculated the correlation between the derived normalized duration and spectral measures, comparing the averages for every speaker. In this case, we only compared the fast speech samples, since the long vowels at normal speech rates were set as reference to 100% in all cases.

Statistically, there was no correlation between duration reduction and relative expansion (with the exception of [i:]), which means that the identified strategies of reducing temporal and spectral features of vowels may not be considered complementary strategies to achieve faster articulation (Table 3). In fact, some speakers kept the spectral quality of their vowels almost unchanged while speaking faster. All the above suggests that the strategies speakers use to achieve faster speech rates seem to be idiosyncratic.

Table 3: *Determination coefficient, Pearson's correlation and significance for duration / expansion*

dur. vs. expans.	R^2	**r**	**p-value**
[a:]	0.035	0.188	0.4407
[i:]	0.304	0.551	0.0144
[u:]	0.161	0.401	0.0883
[a]	0.019	0.136	0.5783
[ɪ]	0.000	0.001	0.9979
[ʊ]	0.001	0.034	0.8886

5. Conclusions

We can conclude that with the method we chose to use in the present experiment, the articulation rate could actually be increased. However, the degree of acceleration was very much dependent on the ability of the speakers to move their articulators faster. The length contrast ratio between long and short vowels decreased when the speakers increased their speech rate, and was smallest for the /u/-vowel pair but did not neutralize. In fact, the length contrast was also observed to extend into the surrounding sounds of the minimal pairs in both speech rates conditions.

In the case of spectral distinction of the vowel length pairs we found that the contrasts between the vowel pairs varied in magnitude: we found the greatest difference between the /i/-vowel pair, and the smallest in the /a/-vowel pair. In contrast to the tendencies found in the durational differentiation of the vowel pairs, spectral features were not reduced when speakers produced faster speech rates.

The above indicates that the spectral features of the vocalic length contrast are more robust against an increase in speech rate compared to the durational features in German. On the other hand, the extension of the length contrast to the temporal structure of the entire minimal pair reflects the fact that the vocalic length opposition is an important feature of the German vocalic system.

It is important to note that the present results do not necessarily reflect the speakers' individual strategies in the experiment. Moreover, they may not reflect the strategies used by the speakers to speak faster in everyday situations, as speech rate differences between normal and fast speech were most likely greater here (and fast speech was much faster) than in normal communicative situations and the experimental design restricted the use of different acceleration strategies.

6. Acknowledgements

This study was conducted as part of the project funded by TKA and DAAD: "Phonetic aspects of fast speech in Hungarian and German" No. 177375, and was also supported by the NKFIH grant No. FK128814.

7. References

[1] P.M. Adank, R. Smits and R. van Hout, "A comparison of vowel normalization procedures for language variation research", *JASA* vol. 116, pp. 3099-3107 (2004)

[2] N. Flynn and P. Foulkes, "Comparing Vowel Formant Nomalization Methods" in *Proc. of the 17th Int. Congr. of Phon. Sciences*, Hongkong, pp. 683-686 (2011)

[3] K.H. Ramers, *Vokalquantität und -qualität im Deutschen*. Niemeyer, Tübingen 1988.

[4] T.H. Crystal and A.S. House, "The duration of American-English vowels: an overview", *Journal of Phonetics* vol. 16, pp. 263-284 (1988)

[5] K.J. Kohler, "German", *Journal of the International Phonetic Association*, vol. 20, pp. 48–50 (1990)

[6] B. Siebenhaar and M. Hahn, „Vowel Space, Speech Rate and Language Space", in S. Calhoun ea. (eds.): *Proc. of the 19th Int. Congr. of Phon. Sciences*, Melbourne, pp. 879-883 (2019).

[7] W.F. Sendlmeier and J. Seebode, „Formantkarten des deutschen Vokalsystems", *https://www.static.tu.berlin/fileadmin/www/10002019/Forschung/Formantkarten_des_deutschen_Vokalsystems_01.pdf* (2006).

[8] T. Becker, *Das Vokalsystem der deutschen Standardsprache*. Lang, Frankfurt a.M. 1998

[9] H. Pfitzinger, „Phonetische Analyse der Sprechgeschwindigkeit", *FIPKM Forschungsber. Inst. Phonetik u. Sprachl. Komm. Uni. München* vol. 38, pp. 117-264 (2001)

[10] G. Fenk-Oczlon and A. Fenk, "Measuring basic tempo across languages and some implications for speech rhythm", *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 2010)*, Makuhari, Japan, pp. 1537-1541 (2010)

[11] F. Pellegrino, C. Coupé and E. Marsico, E., "Across-language perspective on speech information rate", *Language*, vol. 87(3), pp. 539–558 (2011).

[12] S. Kowal, R. Wiese and D. O'Connell, "The use of time in storytelling", *Language and Speech*, vol. 26, pp. 377-392 (1983)

[13] S. Kowal and R. Wiese, *Pausenverhalten in experimentellen Erzählsituationen* Unpubl. Ms (1981)

[14] B. Lindblom, "Spectrographic study of vowel reduction" J. Acoust. Soc. Am. vol. 35. 1963, pp. 1773–l78l.

[15] B. Lindblom, "Explaining phonetic variation: a sketch of the H & H theory," Speech production and speech modeling, W. J. Hardcastle and A. Marchal, Eds. Dordrecht: Kluwer, 1990, pp. 403–439.

[16] G.E. Peterson and I. Lehiste. "Duration of syllable nuclei in English." *The Journal of the Acoustical Society of America* vol. 32.6 pp. 693-703 (1960)

[17] P. Boersma and D. Weenink, Praat: doing phonetics by computer. (2022, 6.3). [Computer program]. Available: http://www.praat.org/