Syllable cut and energy contour: a contrastive study of German and Hungarian

Katalin Mády[1], Krisztián Z. Tronka[1] & Uwe D. Reichel[2]

[1] Institute of German Studies, Pázmány Péter Catholic University, Hungary
[2] Department of Phonetics and Speech Communication, University of Munich, Germany
{mady|tronka}@btk.ppke.hu, reichelu@phonetik.uni-muenchen.de

Abstract

Syllable cut is said to be a phonologically distinctive feature in some languages where the difference in vowel quantity is accompanied by a difference in vowel quality like in German. There have been several attempts to find the corresponding phonetic correlates for syllable cut, from which the energy measurements of vowels by Spiekermann proved appropriate for explaining the difference between long and short vowels. On this basis, we intended to compare German as a syllable cut language and Hungarian where the feature was not expected to be relevant. However, the phonetic correlates of syllable cut found in this study do not entirely confirm Spiekermann's results. It seems that the energy features of vowels are more strongly connected to their duration than to their quality.

1. Introduction

The German vowel system is characterised by a correlation of vowel quantity and vowel quality: long vowels are normally tense, while short vowels are lax, cf. /iː/ – /ɪ/: *Miete* 'rent'– *Mitte* 'centre', /eː/ – /ɛ/: *Weg* 'way' – weg 'away' etc. It has long been discussed whether one of both features is predictable from the other and can therefore be regarded as redundant.

One group of phonologists treats the quantity as the primary phonological (or even the only phonologically relevant) feature in this opposition. However, quantity is an accent-phenomenon in German, i.e. long vowels occur mainly in stressed position. An appropriate description must thus assume a set of rules shortening an underlying long vowel in an unstressed syllable in order to provide the correct surface forms, cf. ***Musik*** [u] [iː] 'music' – ***Musiker*** [uː] [i] 'musician' – ***musikalisch*** [u] [i] [ɑː] 'musical' – ***Musikalität*** [u] [i] [ɑ] [eː] 'musicality'. Other phonologists propose that the distinctive feature is rather tenseness. Since this feature remains intact in the alternation above, such an analysis can describe it in a more plausible way without assuming rules changing an underlying feature in the surface representation. However, the assumption of distinctive tenseness is in one respect unsatisfactory: there are several connections between the vowel opposition and prosodic phenomena (quantity, stress, phonotactic equivalence between long vowels, diphthongs and short vowel + consonant combinations etc.) – indicating that this opposition is probably not a segmental one.

Another solution of the problem is based on the assumption of a syllable cut opposition in Standard German. The basic idea of this concept is that stressed short lax vowels are somehow "unperfect" in the sense that they require a postvocalic segment in the same syllable, while short (if unstressed) or long (if stressed) tense vowels do not. The described problems of the other two concurring theories are avoided in this concept since (1) the opposition of abrupt cut (*scharfer Schnitt*) with a lax vowel and smooth cut (*sanfter Schnitt*) with a tense vowel is clearly a prosodic one and (2) temporal differences between the two vowel classes are just concomitant phonetic phenomena (or even side effects) of this higher suprasegmental contrast. Despite of its phonological plausibility, this concept was often rejected in the second half of the 20th century – because of the lacking phonetic correlate of the syllable cut in Contemporary German.

In his study, Spiekermann [1] discussed and investigated all phonetic correlates for vowel segments that had been assumed so far by phonologists from Sievers through Trubetzkoy up to Vennemann and Maas & Tophinke (for references, see [1]). Spiekermann found that the parameters used to describe energy contours were highly relevant for the contrast abrupt vs. smooth cut: 'E-Max' (Germ *E-Zahl*, number of energy peaks), 'E-Pos' (position of the energy maximum) and 'E-Hold' (Germ *E-Halt*, difference between energy minimum and maximum divided by the maximum). According to Spiekermann's results, smoothly cut (i.e. tense and long) vowels had more energy peaks that were located further back in the segment, and smoothly cut vowels had a higher intensity level throughout the entire segment than abruptly cut vowels. The tendency for the energy maximum to be located further back in smoothly cut and earlier in abruptly cut vowels lead Spiekermann to the assumption that the main characteristics of the syllable are not to be found in the nucleus-coda transition as proposed by Sievers, but in the onset-nucleus transition.

Spiekermann also tested vowel oppositions in Finnish and Czech that primarily make use of a quantitative opposition and thus are not regarded as a syllable cut language. He found that in all languages, longer durations are associated with a higher E-Max, while E-Pos and E-Hold were more or less indifferent for duration. These values were either located between smooth and abrupt cut in German or were closer to the measures for abrupt cut.

While Spiekermann's results are impressing, there are two main shortcomings in the experimental setup. Firstly, he relied on a relatively small corpus (n = 225) that involved all VC combinations of German uttered only once, thus, no statistic analysis could be undertaken. Secondly, his analysis was carried out manually, and the parameters were expressed in three categories instead of metric (i.e. percent) values.

There are strong phonological arguments for the assumption that syllable cut is not crucial for the Hungarian vowel system [2]. First, while a German syllable including a short vowel is only well formed if the vowel is followed by a consonant, short vowels can be syllable final (i.e. they do not require a coda) in Hungarian (eg. ***falu*** /a/, /u/, 'village'). Second, the relevance of syllable cut was primarily restricted to accented syllables, as it is the only position where vowel quantity is distinctive in German (and most Germanic languages). In Hungarian, however, vowel quantity is independent of word stress (that is always on the first syllable in the word).c.f. *falat* /a/ 'mouthful' – *falát* /aː/ 'his/her wall'.

Like German, Hungarian involves seven vowel classes [3], (/i, y, u, e, ø, o, a/), of which all can be realised long or short. The main vowel opposition in Hungarian is durational, while long and short /e/ and /a/ also differ in quality. There is a smaller quality opposition in /o/ and /ø/, where the laxness of the short vowel is mostly explained by dynamic effects, and which most speakers of Hungarian are not aware of [4].

Based on the assumption that syllable cut plays a central role for German vowels but it is not relevant for Hungarian, it was hypothesised that the features E-Max, E-Norm, E-Pos, and E-Hold were relevant for the distinction between long (smoothly cut) and short (abruptly cut) vowels in German. At the same time, long and short vowels were not expected to differ significantly for Hungarian along their energy patterns but behave similar like in Finnish and Czech.

Our comparative study was based on slightly modified measuring methods based on metric instead of ordinal scales (see below). For this reason we first tested these measures for German and Hungarian separately. On the basis of our findings, the results from the two languages will be compared and discussed.

2. Material and methods

2.1. Speech material

Both the German (4 speakers) and the Hungarian (2 speakers) corpora included /i, y, u, e, ø, o, a/ as short and long vowels in nonsense words embedded in a carrier sentence (including 6 syllables in the German corpus and 9 in the Hungarian one). German words had the structure /C_1VC_2ə/ where C_1 and C_2 were stops and had the same place of articulation (PoA) (labial or velar), while C_1 was voiced and C_2 unvoiced. The structure of the Hungarian stimuli was slightly different, as the last vowel was /a/ and C_1 and C_2 were identical. Consonants were varied for PoA (palatal and velar) and voicing. Both corpora were balanced for vowel duration and quality and consonant PoA plus voicing in Hungarian (n(germ) = 1076, n(hung) = 1228).

2.2. Methods

Smoothed energy contours were calculated by applying overlapping Hanning windows of 25 ms on the rectified oscillogram in order to remove glottal closure peaks.

Following measures, based on Spiekermann [1] but not in full accordance with this study, were then derived from these contours (see also Fig. 1):

- the absolute number of maxima 'E-Max',
- this value normalised to the length of the contour 'E-Norm' (number of peaks divided by number of samples),
- the relative position 'E-Pos' of the last maximum within the contour, and
- the ratio 'E-Hold' of the difference between the global maximum and minimum with respect to the global maximum.

In contrast to Spiekermann we did not only calculate 'E-Max' in absolute terms, because a positive relation between contour length and the number of maxima within this contour is somewhat self-explaining, and as vowels in smoothly cut syllables tend to be longer than in abruptly cut ones, the former will trivially show more energy peaks than the latter.

In order to cancel out this durational effect, we divided the energy peak number by the length of the energy contour. Furthermore we avoided the loss of information due to data quantification Spiekermann carried out for E-Pos, for which he divided the vowel into 9 segments, and for the quotient E-Hold which had been categorised in 3 different classes. Instead of categorising E-Pos, we directly calculated the relative position of the last maximum with respect to vowel length, and also for E-Hold no classification was done. Therefore in our study the features E-Pos and E-Hold are not ordinally but metrically scaled.

The measure E-Pos was also modified: in Spiekermann's analysis, only vowels with exactly one energy maximum were included in the analysis. However, if the distinctive character of syllable cut is based on the state of the energy level in the moment when the vowel is cut by the following consonant, then the position of the last energy maximum is relevant, no matter how many peaks were counted before. Thus, we calculated E-Pos as the position of the last maximum, but for reasons of compatibility, E-Pos was also calculated for vowels with one maximum.

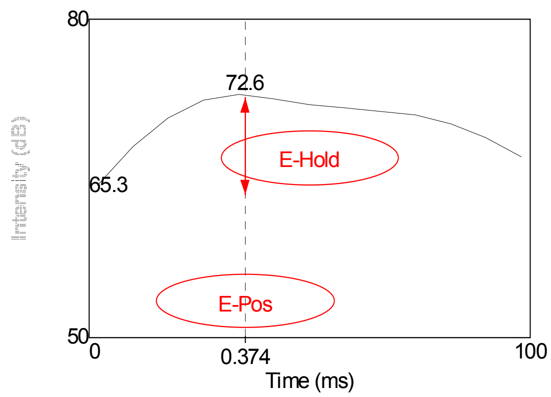

Figure 1: Energy contour of a vowel segment with the parameters E-Max = 1, E-Norm = 0.0007, E-Pos = 0.374, and E-Hold = 0.095, duration = 100 ms.

3. Results

3.1. German vowels

3.1.1. *Vowel length*

First, the parameters duration, E-Max, E-Norm, E-Pos and E-Hold were tested for correlation. None of the energy measures showed a linear correlation. Therefore, Spearman's Rho was calculated over all parameters. No strong correlation (higher than $\rho = 0.6$) was found between any of the parameters. Duration was correlated positively with E-Max, E-Pos and E-Hold, but negatively with E-Norm, i.e. longer vowels had relatively less energy peaks than short ones.

The significance of duration and the energy measures was tested by a T-test ($\alpha \leq 0.05$, two-tailed). Most units of data did not meet the condition of a normal distribution for an ANOVA, but all were large enough to perform a Welch test (n > 50) that does not require normally distributed and homogenous samples.

The difference for all tested variables between smoothly and abruptly cut vowels was highly significant. Long vowels had more energy peaks (E-Max). However, the relative number of energy maxima (E-Norm) was smaller for longer vowels, i.e. they were less compact in long vowels than in short ones. The last energy maximum (E-Pos) was located further back in the vowel segment, as was proposed by Spiekermann.

Our results differ from those of Spiekermann at a further point: while he found a high intensity level throughout the entire vowel segment (E-Hold), our results show exactly the opposite: longer vowels have a larger difference between intensity maximum and minimum than short ones (Fig. 2).

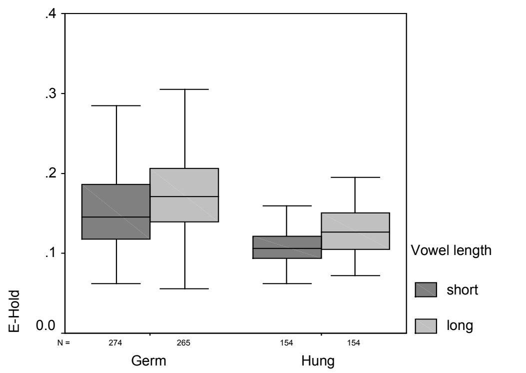

Figure 2: E-Hold of short and long vowels in German and Hungarian.

3.1.2. Vowel classes

In order to test the overall results for vowel quality, the T-test was performed for each German vowel. Although the tendency described above could be observed in most vowels, almost no vowel (except for /ø/) suited the expected sample, at least on a 5% level.

All long vowels had a significant greater duration and a larger number of energy maxima. Most vowels (except for /a/ and /u/) had a higher value for E-Norm.

The least reliable parameter was E-Pos. Three vowels did not show a difference at all (/i, o, u/), and in /e/, the difference was not significant. The tendency in E-Hold was not much clearer: 4 vowels matched the overall pattern, while three (/i, y, a/) did not. If E-Pos and E-Hold were calculated according to Spiekermann's method, the pattern was even less clear.

3.2. Hungarian vowels

3.2.1. Vowel length

For Hungarian, correlations between duration and the energy measurements were approximately identical with German. Also the T-test revealed the same tendencies between long and short vowels: a higher number of energy peaks in long vowel, but their higher density in short vowels. E-Pos lied further back for long vowels which also had a larger maximum-minimum difference of intensity.

3.2.2. Vowel classes

As already seen with the German data, duration was significantly different between short and long vowels. While most long vowels had more energy peaks than their short pendant (not true for /e/), E-Norm showed the same distribution as for the total sample: smaller percent values for long vowels than for short ones.

E-Pos had again an unclear tendency (while /y, u, a/ fitted the overall pattern) according to which E-Pos was higher for E-Hold vowels. On the other hand, E-Hold met our expectations regarding vowel difference.

3.3. Comparison of German and Hungarian vowels

As said in 2.1, both the German and the Hungarian corpora involved two different consonant contexts. Although both consonant PoA and voicing have an impact on vowel duration, this could be ignored in the previous sections because each corpus was balanced for these factors. However, a comparison between German and Hungarian vowels required a corpus where at least the consonant following the vowel (that is the more influential one for vowel length) was identical for both languages. Therefore, only stimuli with a V + /k/ sequence were considered for the contrastive corpus. It was still large enough to meet the condition for the Welch test (n(germ) = 539, n(hung) = 308) for general comparison. But, as the Hungarian corpus involved an additional factor to the German one (voicing contrast for consonants), the number of repetitions for the single vowel classes was rather small (n=22). Therefore, the non-parametric H-test (by Kruskal-Wallis) was performed instead of the Welch test.

German and Hungarian vowels (without any specification) differed significantly for duration, E-Max, E-Pos, and E-Hold, but not for E-Norm. The energy parameters were related to duration in the same way as in the separate languages: German vowels, that were significantly longer than Hungarian ones (m(germ) = 113 ms, m(hung) = 94 ms, see Fig. 3), had a higher number of absolute but a lower number of relative maxima, which were located further back in a vowel segment, in which the overall energy level was lower.

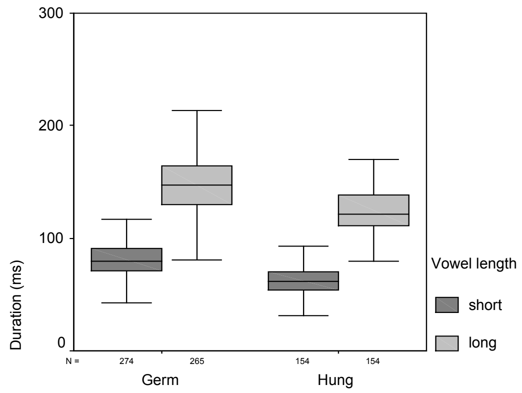

Figure 3: Duration of short and long vowels in German and Hungarian.

The comparison for long vowels and short vowels between the two languages revealed a similar tendency as was found for German long and short vowels on the one hand and Hungarian long and short vowels on the other hand: while

German long and short vowels were longer than their Hungarian counterpart, they also included more absolute energy peaks which occurred further back in the vowel and had a larger minimum-maximum difference in the intensity level. However, this tendency is only clear-cut for short vowels. Against our expectation, the relative number of energy peaks is somewhat smaller in the shorter Hungarian vowels (the difference is not significant, though), and the measure E-Hold does not differ significantly for the long vowels of the two languages.

Finally, German and Hungarian vowels were compared class-wise (here the H-test was used). A significant longer duration was found for each vowel class, except for /yː/. Surprisingly, an interlingual comparison of the vowel classes (German /a/ with Hungarian /a/, German /aː/ with Hungarian /aː/ etc.) showed that none of the 14 vowels had four significantly different energy parameters. In other words, the general interlingual tendency could not be shown for any of the vowel classes.

When German and Hungarian vowels were compared separately according to their length (see 3.1.2 and 3.2.2), E-Norm and E-Hold seemed to be relatively stable parameters for differentiating between long and short vowels within one language. If only these two parameters are being taken into account, then three vowels fit the general pattern: short /i/, /y/ and /u/.

4. Discussion and conclusions

The theory of syllable cut is based on languages where (1) long vowels are associated with tenseness and (2) stressed short vowels always require a consonantal syllable coda. In these languages, vowel length is only distinctive in stressed syllables. The main idea is that smoothly cut (long and tense) and abruptly cut (short and lax) vowels are not only different by length and tenseness, but there is a hidden, a more general category that governs the other two features.

After several decades of unsuccessful search for reliable acoustic correlates of the hypothetical syllable cut, Spiekermann [1] showed that the difference relies on the different intensity patterns in vowels with smooth and abrupt syllable cut. He also proposed that non-syllable-cut languages like Finnish and Czech did not show the regularities that were found for German. The same tendency was expected for Hungarian which is a non-syllable-cut language.

In this study, the energy and duration parameters proposed by Spiekermann were retested on a larger German speech corpus. The results showed a significant difference between long and short vowels for all parameters he used, but the intensity pattern we found was different: long (smoothly cut) German vowels had in fact less energy maxima than short vowels if normalised by vowel duration, and the intensity levels were characterised by larger differences than in short (abruptly cut) vowels. Thus, the image that tense vowels are characterised by a straightforward and high energy level did not prove to be appropriate. On the contrary, short (lax) vowels seem to have a more compact energy distribution (relatively more energy maxima and a smaller decrease of the intensity level during the vowel segment) according to their higher E-Norm and lower E-Hold values. E-Pos did not seem to be a reliable measure, no matter whether all vowels were considered or only those with one peak (as in [1]). Furthermore, E-Hold is a problematic parameter due to the non-linearity of the dB scale, the quotient (energy maximum – energy minimum)/energy maximum still depends on the respective energy level.

Against our hypothesis, exactly the same energy patterns were found between long and short Hungarian vowels that are not governed by syllable cut supposed. The analysis of single vowel classes showed that Hungarian vowels fitted the general pattern even better than German ones (there were less exceptions and non-significant relationships). Although there are hints that vowels differ according to vowel height, the tendencies are not clear-cut enough to allow for further conclusions.

The difference found between long and short vowels in each language also appeared when German and Hungarian vowels were compared with each other. German vowels, which were longer than their Hungarian counterparts: had a higher absolute number but a lower relative number of energy maxima, the last peak was located further back in the segment and the intensity difference was larger. Therefore, it is possible that the parameters examined by Spiekermann are of more general nature and are inherently connected to vowel duration. It is evident that a longer vowel will probably contain more energy peaks and a larger range of increase and decrease of intensity level than a shorter segment. On the other hand, a smaller number of samples (= a shorter duration) will automatically lead to higher E-Norm values. Thus, the observed tendencies do not necessarily reflect anything else than side effects of duration.

For a better understanding of this phenomenon, several further aspects should be taken into account: what pattern can be found in (1) unstressed vowels, (2) diphthongs, and (3) reduced vowels? Do unstressed vowels, that are long and tense underlyingly but short on the surface, show the characteristics of smooth or abrupt cut? Is duration a crucial factor for reduced vowels that are not marked for tenseness? How would the tenseness difference of Hungarian mid vowels influence the energy parameters in a larger corpus? The answers to these question may bring us closer to the question whether syllable cut is a relevant feature in the German vowel system.

5. Acknowledgements

We would like to thank to Christian Geng, ZAS, Berlin, for providing us with the German and Hungarian data, and to Phil Hoole, Department of Phonetics and Speech Communication, University of Munich, for his helpful comments.

6. References

[1] Spiekermann, H., 2000. *Silbenschnitt in deutschen Dialekten.* Tübingen: Niemeyer.

[2] Tronka, K., in press. Die Vokallänge im Deutschen und Ungarischen. In: *Junge Germanisten stellen sich vor,* I. Szigeti (ed.). Frankfurt/Main: Lang.

[3] Mády, K., 2001. Kontrastive Phonetik Deutsch-Ungarisch in Hinblick auf zu erwartende Interferenzphänomene. *Linguistische Beiträge Pasmaniensis 1*, 29–51.

[4] Siptár, P.; Törkenczy, M., 2000. *The Phonology of Hungarian*. Oxford: University Press.