2014 ACTA UNIVERSITATIS CAROLINAE PAG. 21–31
PHILOLOGICA 1 / PHONETICA PRAGENSIA XIII

VARIABILITY OF CZECH ALVEOLAR PLOSIVES: A LOCUS EQUATION PERSPECTIVE

DAVID LUKEŠ, DITA FEJLOVÁ,
RADEK SKARNITZL

ABSTRACT

Locus equations are linear regressions of (F2-at-vowel-midpoint, F2-at-consonant-release) pairs, derived from CV sequences with one C (usually an obstruent stop) and a range of V's. This article uses them primarily to assess the synchronic variability of Czech /d/ × /t/ as an indicator of possible sound change in the case of /d/, but also as a means for a preliminary evaluation of their speaker-distinguishing potential, on the premise that changing sounds may yield better results than stable ones. While evidence for the variability of /d/ has been found in the reduced linearity (r^2) of its characteristic LE, conclusive proof of its usefulness for FSI, especially as compared with /t/, still remains to be adduced. The articulatory dimensions of the variation of Czech /d/ (alveolar × dental, apical × laminal) are also discussed.

Key words: locus equations, sound change, FSI, Czech, plosives

1. Introduction

In the field of forensic speaker identification (FSI), it has become clear that no single acoustic parameter will serve as a silver bullet, discriminating one speaker from all others. That is why researchers have been attempting to find parameters which may contribute to speaker identification in a multidimensional space. In the segmental domain, one strand of research has been devoted to the speaker-discriminating power of sounds undergoing historical change. De Jong et al. (2007) contrasted the usefulness for FSI of changing Southern Standard British English vowels /æ, ʊ, uː/ with that of stable /iː, ɑː, ɔː/. While the outcomes have to be interpreted with some care, and it is certainly not the case that all evolving speech sounds are by definition better identifiers than their comparatively static counterparts (de Jong et al., 1813), the achieved results warrant further investigation, in particular on other languages and/or other types of speech sounds.

In the case of Czech, there is a slight problem in that synchronic variation with respect to individual speech sounds is seldom systematically researched, which means that convincingly identifying sounds undergoing diachronic change is not a trivial task. However, it is well known that synchronic variation often turns out to reflect diachronic change (see, e.g., McMahon, 2000: 3; Harrington, 2012). Since non-systematic observations indicate that the synchronic variability of Czech /d/ might be increasing, this could speak for incipient sound change. In canonical realizations, the Czech alveolar plosives /t/ and /d/ differ slightly in terms of their place of articulation: while /d/ is fully alveolar or even partially postalveolar in some speakers, its voiceless counterpart tends to be articulated in the addental region (Skarnitzl, this volume). Even though this form of articulation remains prevalent, some speakers, including professional ones from the media, appear to exhibit a tendency to unify the place of articulation of both stops in favour of the addental position.

In order to assess the variability of /t/ and /d/ in Czech, we decided to use locus equations (LE's), which characterize the relationship between the given plosive and the following vowel, or more precisely, between F2 at the consonantal release and at the midpoint of the following vowel. Technically speaking, they are "[first-order polynomial] linear regressions of the frequency of the second formant transition sampled at its onset (F2-onset) on the frequency of the second formant sampled in the middle of the following vowel (F2-vowel)" (Sussmann et al., 1997: 2826). In terms of a graphical representation, this means that F2-vowel on the x-axis is plotted against F2-onset on the y-axis, and the resulting dataset is fitted with a line which best approximates it.

The general equation of a line being $y = ax + b$, the constant a is the *slope* of the regression line (how steeply it rises or falls) and the constant b, its *y-intercept* (the ordinate at which the line intercepts the y-axis). For each stop, a LE of this form is generated, with a pair of values (a, b). In the past, these values have been successfully used as indicators of the place of articulation of the stop used to generate the given LE (Sussmann et al., 1991), but recent research into the articulatory mechanism underlying the regularity of the LE relationship has shown that more circumspection is needed. Thus, according to Iskarous et al. (2010: 2031), LE's are most probably the acoustic counterpart of a corresponding articulatory linear relationship involving the horizontal position of the tongue body (at the previously defined points of a CV sequence), which does interact with the place of articulation (i.e. the tongue tip / blade position), but still remains too indirect an indicator thereof. Nevertheless, Manuel (1995) has successfully used measures of F2 for capturing place variation in the dental-alveolar range: she found that contextually nasalized [ð] in phrases like win those has a markedly lower F2 at release than proper alveolar [n] in win noes (*ibid*.: 461). Consequently, we considered F2 a priori as a viable indicator of place of articulation differentiation between Czech /t/ and /d/, with the reservation that in the case of plosives, F2 transitions might be less useful because potentially masked by the explosion.

In this study, the variability in the /t/ and /d/ data will be evaluated based on the coefficient of determination (r^2) of the LE regression line, which is a statistic describing the goodness of fit of the line; it quantifies how much of the variance in the original data is explained by the linear model. The relation embedded in LE's has repeatedly been shown to be highly linear (see, e.g. Lindblom and Sussmann, 2004), i.e. the regression lines have a high r^2, and when this linearity is compromised, it is generally for the reason that the datapoints belong to several different populations. A case in point are velars, for which, in general, two LE's are derived: one in the context of front vowels, the other when the following vowel is back.

The present article thus strives to determine whether impressionistic observations regarding /d/, as described above, have a measurable acoustic basis, in other words whether there really is a degree of variability in different speakers' realizations of /d/ which would make it a potentially useful identifier. Some variability has already been confirmed in the articulatory data presented by Skarnitzl (this volume), but articulatory variability does not always translate into acoustic (or perceptual) variability. Greater variability of the articulation place of the Czech /d/ – a greater non-homogeneity in the population – would be indicated by lower r^2 values and may, in turn, reflect an ongoing sound change. If this main hypothesis is confirmed by our data, /d/ would, according to the subsequent hypothesis, serve as a valuable source of speaker-specific information for speaker identification purposes.

2. Method

2.1 Materials & Subjects

Seventeen adult male native speakers of Czech were asked to read 35 sentences in a natural manner. The sentences contained 76 [tV] and [dV] sequences, where V is a vowel from the set /ɪ, iː, ɛ, ɛː, a, aː, o, oː, u, uː, ou̯/. The preceding segmental context was not controlled in any special way.

The sentences were recorded at the sampling frequency of 32,000 Hz with 16-bit quantization and segmented automatically using the Prague Labeller (Pollák et al., 2007). The boundaries of target speech sounds were subsequently adjusted manually in accordance with the guidelines in Machač and Skarnitzl (2009). It should be pointed out that, for the sake of consistency with the Prague Phonetic Corpus (Skarnitzl, 2010), this segmentation may have differed slightly from that usual in LE studies, in which the boundary of the CV transition is placed at the first faintest sign of appearance of F2; in our approach it is the onset of *full* formant structure which is the primary criterion (*ibid.*: 30).

By implication, this meant that we could not work with the traditional LE-related anchor points for F2 measurement, which are defined in terms of vowel boundaries: F2 at onset of transition and F2 at mid-vowel. After some experimentation, placing the first anchor at −5 ms and the second at +20 ms relative to *our* boundary yielded satisfactory results in terms of the derived LE's. Thus, in connection with our study, these will be the actual physical points in time referred to when mentioning the concepts of F2-onset and F2-vowel. The apparent leeway in the exact placement of these anchor points is an argument in favour of the robustness of the LE relationship, a point already made in Sussman et al. (1997), who showed that an LE relationship, albeit somewhat less linear, can also be derived from VC sequences.

2.2 Formant Analysis

Using a script in Praat (Boersma and Weenink, 2012) with standard settings for male and female voices, we extracted F2 formant trajectories from each CV sequence, starting 10 ms before the CV boundary and ending 20 ms after it, with an increment of 5 ms. This yielded seven equidistant measurement points, F2-onset and F2-vowel being the second and the final value, respectively. We went through the formant tracks and noticed that part of them were intuitively implausible (values jumping back and forth, or improbably high/low). Attempts to hand-correct them were not always successful, as the broadband spectrograms in Praat often did not offer any additional clues, the F2 frequency being "drowned" in surrounding noise. Hand correction is also prone to be affected by the corrector's bias. To sift through the F2 contours and flag "suspect" ones, we therefore devised a method which attempts to simulate post-hoc the effects of a formant tracker.

The F2 contours were each individually fitted with a curve based on quadratic polynomial regression, the rationale for this being that formant transitions reflect smooth and gradual transformations of the vocal tract configuration; it is a reasonable assumption (which is implemented by formant trackers) that they should not "jump" excessively on a local scale. In FSI, such parameterization of formant trajectories was pioneered by McDougall (2006). In a following paper, she and Nolan conclude that with their material, "cubic polynomials provide a better fit to the F1 and F2 contours, [but] it appears that a worthwhile amount of speaker-distinguishing information can be captured with […] quadratic approximations" (2007: 1828). Since we only fitted a window from the first half of vowels, we regarded quadratic polynomials as sufficient for capturing the relevant dynamic patterns.

r^2 values were computed as an estimator of goodness of fit and contours were classified into several subsets according to threshold values of r^2. For the sake of clarity, we shall hereafter distinguish the two uses which we have made of the r^2 estimator: while in relation to the goodness of fit of an LE regression line, we shall continue to call it r^2, in its second, auxiliary capacity as "post-hoc formant tracker", we shall refer to it as "threshold". In our subsequent analyses, we worked primarily with tokens with the highest threshold value possible.

3. Results and Discussion

3.1 The variability of Czech /d/

The different subsets of the initial data set, picked according to increasingly stringent threshold criteria, are characterized in Table 1. The r^2 values of the LE's for [tV] and [dV] derived from each subset are also tabulated. As we have mentioned, our original assumption motivating the use of the post-hoc formant tracker was that we would be able to exclude mismeasurements, allowing the underlying linearity of the LE relationship, demonstrated by previous studies, to manifest itself more extensively, thus maximizing r^2. However, while the r^2 for the [tV] LE shows a more or less steady increase from 0.566 (no threshold applied) to 0.711 (threshold: > 0.95), the r^2 for the [dV] LE shows a more or less steady *de*crease from 0.653 (no threshold applied) to 0.599 (threshold: > 0.95). The data sets with a threshold value of 0.95, along with their LE regression lines, are plotted in Figures 1 and 2, for [tV] and [dV] sequences respectively.

At first glance, this may seem as an invalidation of the usefulness of our post-hoc formant tracker, but only until one realizes what this says about the data, and how it

Table 1. Summary of the effects of the post-hoc formant tracker (see text for explanation). From top to bottom, increasingly restricted subsets of the initial data set with increasingly strict post-hoc formant tracker thresholds are presented (the first row describes the entire original data set). The r^2 of the LE's for [tV] and [dV] sequences derived from each subset are also tabulated.

post-hoc formant tracker threshold	# of [tV] tokens	# of [dV] tokens	r^2 of [tV] LE	r^2 of [dV] LE
N/A	707	572	0.566	0.653
> 0.70	609	531	0.609	0.632
> 0.80	553	503	0.596	0.613
> 0.90	422	420	0.643	0.598
> 0.95	269	324	0.711	0.599

fits together with our previously mentioned informal observations on the variability of contemporary Czech /d/. At this point, we might perhaps draw a parallel with a long established practice connected with LE's: that of computing, for velar stops, not one but two LE's, one with a set of front vowel contexts, the other with a set of back vowel contexts (Iskarous et al., 2010: 2028). The motivation for this separation was that when computed from velar stop tokens pooled across all vowel contexts, the resulting LE's yielded a much lower r^2, and even a cursory inspection of the scatterplots of F2-onset against F2-vowel made it clear that the dataset was not strictly linear – it would be much better accommodated by one regression line in its first half and a different one in its second, since velar plosives are articulated in two different ways depending on whether a front or back vowel follows.

In the case at hand, it is not obvious that we would be able to increase the r^2 of the [dV] LE by splitting it into two (or more) according to some vowel-context rule (*cf.* Fig. 2 – no

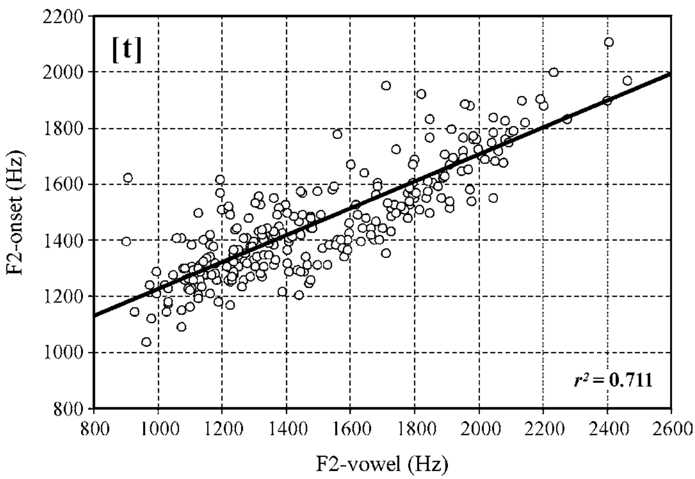

Figure 1. Scatterplot of [tV] tokens across speakers, satisfying a threshold value of 0.95 (see text), fitted with LE regression line ($y = 0.5542x + 688.3$).

breakpoint is apparent from which the data would seem to follow lines with different slopes). However, the intuition that lower r^2 is caused by articulatory variation can be carried over from the velar stop scenario. If this variation, as it seems, is not conditioned by the vocalic context, it can result either from speaker-dependent specificities, or it may even be randomly varying from token to token. This increased, apparently phonologically non-conditioned variability is an important first shred of evidence for the case that /d/ may indeed be undergoing sound change in Czech, making it eligible for further scrutiny as a potentially valuable source of speaker-specific information.

A reader who is acquainted with LE literature will have remarked that our r^2 values are rather low. The main causes are probably threefold. Firstly, one of the observations of Sussmann et al. (1997), who made a comparison of CV-based (traditional) and VC-based LE relations, was that the latter have a significantly lower goodness of fit (a mean r^2 of 0.44 compared to 0.72), for which phenomenon they offered the explanation that CV sequences may require "greater articulatory precision" (*ibid.*: 2834). If linearity and articulatory precision are thus related, the high r^2 values achieved in most LE studies are probably influenced by the fact that the target words are either isolated or embedded in carrier sentences which make them stand out, and thereby elicit a more careful articulation; by contrast, our target words were seamlessly integrated in the surrounding text. Secondly, the slightly different approach to segmenting the CV sequences (see section 2.1) may have rendered the LE relationship more fuzzy. Finally, LE's are most often computed on a speaker-by-speaker basis (e.g. Sussmann et al., 1997; Brancazio and Fowler, 1998), which of course removes interspeaker variability. But since our goal was on the contrary to assess this interspeaker variability, it made sense to pool all the speakers together. Compared with Iskarous et al. (2010: 2025), who have also computed LE's for data pooled across subjects and obtained an r^2 for alveolar stops of 0.635 (contrast with 0.858 for /p,b/ or 0.853 for /k,g/), our results appear much less atypical.

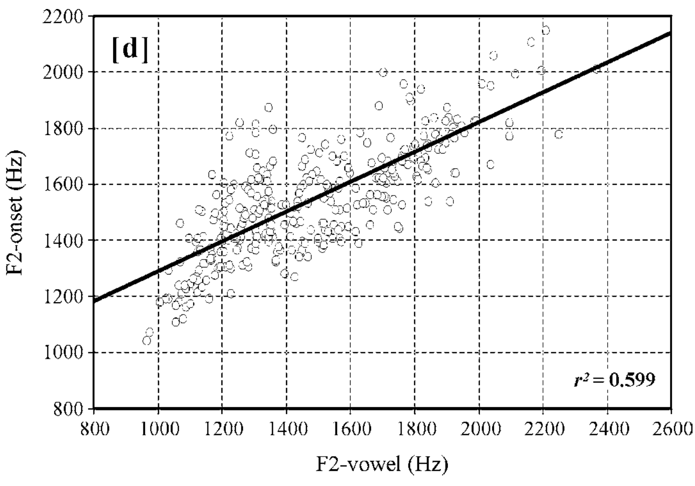

Figure 2. Scatterplot of [dV] tokens across speakers, satisfying a threshold value of 0.95 (see text), fitted with LE regression line ($y = 0.5307x + 757.5$).

3.2 LE's and Articulation

What might be the articulatory reality underlying our LE's for [tV] and [dV]? We mentioned that Iskarous et al. (2010) have demonstrated the correlation between the F2 values used as input for LE's and the horizontal tongue body position. We have also seen that the Czech "alveolars" are in fact addental [t̪] and alveolar-to-postalveolar [d/d̠], with the variability of the place of articulation of /d/ possibly increasing (cf. the electropalatographic displays in Skarnitzl, this volume, Fig. 1). The most natural conclusion is that the position of the tongue tip/blade affects, to a certain extent, the position of the tongue body and thereby also F2, which is reflected by the locus equations. The varying place of articulation for /d/ would thus neatly dovetail with the observed lower r^2 for the [dV] LE. This would seem to be corroborated by the findings of Dart (1991: 89), who found that F2-onset of French and English coronal consonants indeed depends on their place of articulation: it is higher in transitions to and from dental than alveolar stops.

The comparison of F2-onsets in /t/ and /d/ in our data is illustrated in Figure 3. We can see that the mean is significantly higher for alveolar /d/ (t-test for independent samples: t (591) = 4.73; $p < 0.001$), which is in stark disagreement with Dart's observations. How to reconcile this contradiction? Dart also mentions that in the case of constrictions at the front of the mouth, the frequency of F2 is greatly affected by the tongue body height (higher position leading to a higher F2). She goes on to say that "there has been some dispute whether it is the apical or the laminal articulation which has the higher tongue body" (*ibid.*: 63). In other words, lingual configuration (apicality *vs* laminality) also plays a role, even though perhaps unclear. In this respect, Maddieson and VanBik's (2003) research on coronals in Hakha Lai is a case study of sorts. Having determined by linguo- and palatographical means that Hakha Lai contrasts a lamino-dental stop with an apico-alveolar one (*ibid.*: 233–234), they proceeded to study their acoustics and obtained a significantly higher F2-onset mean for the alveolar (*ibid.*: 235), much like we did. This indicates that unlike the French and English coronals studied by Dart, Czech coronals most likely share the same combinations of place of articulation and lingual configuration features as those of Hakha Lai: /t/ tends towards addental and laminal, and /d/ towards alveolar and apical. These features presumably conspire to generate the observed F2 pat-

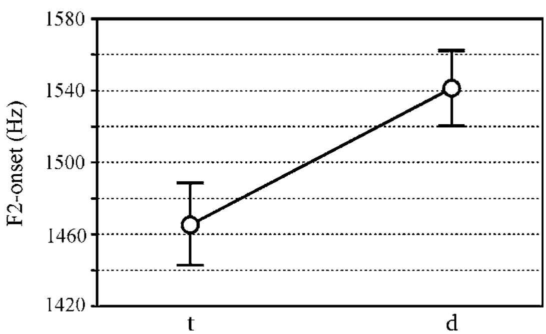

Figure 3. F2-onset means for /t/ and /d/. Computed from the token set with a threshold of 0.95. The vertical bars indicate a 95% confidence interval (±1.96 SE).

terns, clarifying at least part of the dispute referred to by Dart. Let us also briefly recall that, in contrast to Dart, Manuel (1995) found the same patterning of F2 and place of articulation as we did for English dental *vs* alveolar *nasal* stops (*cf.* Introduction).

3.3 LE's as an Indicator of Speaker Specificity

As we mentioned in section 3.1, the increased variability observed in the realizations of /d/ in Czech may be due to speaker-specific differences, as well as due to random variation. In a preliminary attempt to verify the potentially higher speaker-discriminating potential of /d/, as compared to that of /t/, we took the same coefficients as above from *each speaker's* LE, and plotted them one against the other in the two-dimensional *slope* × y-intercept space. Figure 4 illustrates the [tV]-based and [dV]-based LE's. In both cases, the speakers seem to spread out in the available space, but it is obvious that the points derived from [dV] LE's are much less clustered together than those derived from [tV] LE's. Furthermore, speakers perform differently, as can be seen e.g. in the case of PATJ, whose [tV] tokens are quite extreme, and [dV] tokens average. These results indicate that our putatively changing /d/ should indeed do a better overall job at distinguishing speakers.

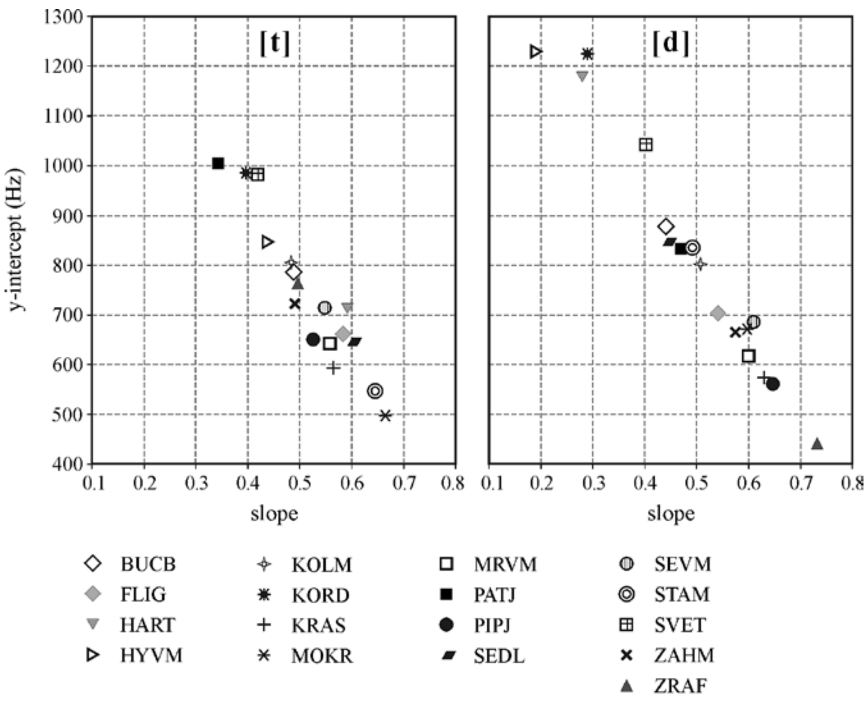

Figure 4. Scatterplot of y-intercept against slope from each speaker's LE derived from his [tV] (left) and [dV] (right) tokens satisfying a threshold value of 0.9.

4. Summary and conclusions

The purpose of this paper was to compare the variability of Czech /t/ and /d/. Previous informal observations, as well as articulatory data, hinted at more variable realizations of /d/ than of /t/: on the one hand, /d/ appears to be produced with a more retracted place of articulation, but on the other hand some speakers seem to produce /d/ as a dental sound. As greater synchronic variability can be regarded as an indicator of ongoing sound change, we hypothesized, based on results reported by de Jong et al. (2007), that the apparently changing /d/ would manifest greater capacity to differentiate speakers as compared with that of the relatively stable /t/.

Locus equations (LE's) were computed from [tV] and [dV] tokens whose reliability had previously been established by a post-hoc formant tracking simulation. For the most reliable subset of tokens (threshold = 0.95), the goodness of fit of the [dV] LE was substantially lower than that of the [tV] LE. When the linearity of LE's is thus compromised, it generally indicates that the source tokens come from several different populations (*cf.* the case of velars). In our case, it has been taken to mean that there is greater variation in the articulation of /d/, perhaps indeed indicative of sound change.

We also examined some of the articulatory implications of our findings. In general, F2 patterns are conditioned by tongue body configuration, but it remains partially unclear how tongue tip/blade positioning affects the tongue body and thereby F2. With the help of data from a study on Hakha Lai, we suggested that the overall higher F2-onset of /d/ is triggered by a combination of alveolar and apical articulation, whereas /t/ rather tends towards laminality in addition to its addental place (see also the discussion in Skarnitzl, this volume). This means that the kind of variability exhibited by /d/ should indeed be reflected by LE's.

Finally, we also tested the speaker-discriminating power of LE parameters derived from [tV] and [dV] sequences. In what must be regarded as a preliminary analysis, we plotted (slope, y-intercept) pairs taken from each speaker's LE, and it could be seen that the way they spread out, particularly the [dV] ones, should allow for discriminating at least between some of the speakers. Confirming the greater spread of [dV]-based LE coefficients and ascertaining their within-speaker consistency, as well as effectively implementing between-speaker comparison with a chosen distance measure, is a subject for future work.

ACKNOWLEDGEMENTS

A project output of the Internal grants 2012 VG130 solved at FF UK. This research was also supported by the project GAČR 406/12/0298 granted to the third author and by the Programme of Scientific Areas Development at Charles University in Prague (PRVOUK), subsection 10 – Linguistics: Social Group Variation.

REFERENCES

Boersma, P. & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.09, retrieved on March 15, 2012 from <http://www.praat.org>.

Brancazio, L. & Fowler, C. A. (1998). On the relevance of locus equations for production and perception of stop consonants. Perception & Psychophysics, 60/1, pp. 24–50.

Dart, S. N. (1991). UCLA Working Papers in Phonetics No. 79, March 1991: Articulatory and Acoustic Properties of Apical and Laminal Articulations. Los Angeles, UCLA Phonetics Laboratory.

de Jong, G., McDougall, K., Hudson, T. & Nolan, F. (2007). The Speaker-Discriminating Power of Sounds Undergoing Historical Change: A Formant-Based Study. In: Proceedings of the 16th ICPhS. Saarbrücken: ISPhS, pp. 1813–1816.

Harrington, J. (2012). The relationship between synchronic variation and diachronic change. In: A. Cohn, C. Fougeron & M. K. Huffman (Eds.), The Oxford Handbook of Laboratory Phonology. Oxford: Oxford University Press, pp. 321–332.

Iskarous, K., Fowler C. A. & Whalen, D. H. (2010). Locus equations are an acoustic expression of articulator synergy. Journal of the Acoustical Society of America, 128/4, pp. 2021–2032.

Lindblom, B. & Sussmann, H. M. (2004). Articulatory and acoustic bases of locus equations. In: P. Branderud & H. Traunmüller (Eds.), Proceedings of FONETIK 2004: The XVIIth Swedish Phonetics Conference, Stockholm, Sweden, pp. 8–11.

Machač P. & Skarnitzl R. (2009). Fonetická segmentace hlásek. Praha: Nakladatelství Epocha.

Maddieson, I. & VanBik, K. (2004). Apical and Laminal Articulations in Hakha Lai. Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 30/1, pp. 232–243.

Manuel, S. Y. (1995). Speakers nasalize /ð/ after /n/, but listeners still hear /ð/. Journal of Phonetics, 23, pp. 453–476.

McDougall, K. (2006). Dynamic features of speech and the characterization of speakers: Towards a new approach using formant frequencies. Speech, Language and the Law, 13/1, pp. 89–125.

McDougall, K. & Nolan, F. (2007). Discrimination of Speakers Using the Formant Dynamics of /uː/ in British English. In: Proceedings of the 16th ICPhS. Saarbrücken: ISPhS, pp. 1825–1828.

McMahon, A. (2000). Lexical Phonology and the History of English. Cambridge: Cambridge University Press.

Pollák P., Volín J. & Skarnitzl R. (2007). HMM-Based Phonetic Segmentation in Praat Environment. In Proceedings of the XIIth International Conference "Speech and computer – SPECOM 2007". Moscow: MSLU, pp. 537–541.

Skarnitzl, R. (2010). Prague Phonetic Corpus: status report. AUC Philologica 1/2009, Phonetica Pragensia, XII, pp. 65–67.

Skarnitzl, R. (2013). Asymmetry in the Czech alveolar stops: An EPG study. This volume.

Sussman, H. M., Bessell, N., Dalston, E. & Majors, T. (1997). An investigation of stop place of articulation as a function of syllable position: A locus equation perspective. Journal of the Acoustical Society of America, 101/5, pp. 2826–2838.

Sussman, H. M., McCaffrey, H. A. & Mathews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization Journal of the Acoustical Society of America, 90/3, pp. 1309–1325.

VARIABILITA ČESKÝCH ALVEOLÁRNÍCH EXPLOZÍV Z POHLEDU LOKUSOVÝCH ROVNIC

Resumé

Lokusové rovnice představují způsob, jak úsporně charakterizovat soubor sekvencí CV, kde C je jeden daný konsonant (většinou explozíva) a V náleží do řady vokálů napříč vokalickým prostorem. Každá realizace je reprezentována bodem, jehož x-ovou souřadnicí je F2 uprostřed vokálu a y-ovou souřadnicí

je F2 konsonantického lokusu. Lokusovou rovnici na základě bodů vyvodíme lineární regresí s polynomem prvního stupně. Takováto data totiž obvykle vykazují vysoký stupeň linearity, a dají se proto zjednodušeně popsat přímkou, resp. jejími koeficienty (směrnicí a úsekem na ose y). Interpretace těchto koeficientů se různí (viz článek). Stupeň linearity je možno kvantifikovat pomocí determinačního koeficientu (r^2), a pokud je v rámci srovnatelných dat u vzorku některého konsonantu nižší než u ostatních, je snaha vysvětlit takový rozdíl tím, že je zdrojová sada nehomogenní. Např. u velárních sad se nižší linearita vysvětluje tím, že vykazují výraznou alofonickou variaci v závislosti na vokalickém okolí (přední × zadní V).

Na základě našich dat se zdá, že sady s českým /d/ se od /t/ liší právě sníženou linearitou, což ukazuje na zvýšenou synchronní variabilitu místa artikulace /d/, neboť platí, že F2 je v koronální oblasti citlivý na změnu místa. Z toho plynou dvě hypotézy, jež je třeba dále ověřovat: a) je možné, že u českého /d/ dochází k posouvání hlásky; b) pokud platí a), pak má /d/ potenciál být ukazatelem, který může přispět k efektivnímu rozlišení mluvčích na základě akustických parametrů. Zároveň naše akustická měření potvrzují čerstvá elektropalatografická data o rozdílu v místě artikulace českého /d/ (primárně alveolární/postalveolární) a /t/ (addentální). Ve světle akustických a artikulačních poznatků z jiných jazyků se navíc zdá, že F2 v této oblasti interaguje i s nastavením jazyka (apikální × laminální).