



# “Flat pitch accents” in Czech

Tomáš Duběda

Institute of Translation Studies, Charles University in Prague

dubeda@ff.cuni.cz

## Abstract

In this paper we investigate a particular type of stress marking in Czech, in which the syllable perceived as prominent is not accompanied by any clearly audible change in the overall pitch course. The paper gives a perceptual, phonotactic and acoustic account of these “flat pitch accents”. No positional effects or semantic correlates of words bearing this type of accent were found. Flat accents have significantly reduced intonational variability, as expected, and their durational and dynamic correlates are partly different from other accent types. However, none of these findings speaks in favour of compensation between prosodic parameters.

**Index Terms:** pitch accents, neutralized accents, Czech.

## 1. Introduction

When transcribing Czech intonation, one observes cases where the syllable perceived as prominent is not accompanied by any clearly audible change in the overall pitch course, i.e. has nearly the same pitch as the preceding and the following syllables. The present study gives a perceptual, phonotactic and acoustic account of these “flat pitch accents”.

Czech is a Western Slavonic language with distinctive vowel length both in stressed and unstressed syllables and only postlexical intonation. Word stress, located on the first syllable of stressable words, has demarcative potential. While earlier descriptions of Czech stress relied mostly on its dynamic nature (e.g. Hála 1962: 303), more recent studies showed that intensity peaks are not a systematic correlate of stressed syllables (Duběda 2006). Accentual lengthening is said to be very limited, mainly due to phonological relevance of vowel length (Hála 1962: 303; Palková 1997: 279), and there seems to be no tendency for unstressed vowels to have a more reduced articulation than the stressed ones (Palková 1997: 279). For the rest, the phonetic characterization of stress is mainly a matter of intonation.

Two alternative views of Czech intonation have been proposed:

a) Some authors (e.g. Palková 1987) hold the position that Czech accents are perceived in the framework of the whole stress unit, which is the relevant building stone of the intonation phrase. This view is also supported by the fact that both native and non-native listeners often report that they perceive Czech accents as very weak, and, in the case of non-native listeners, may hear the accent on a different syllable. The perception of accent in languages with fixed stress may also be influenced by the listeners’ expectations: they “hear” the accent simply because they recognize the word (Ladd 2008: 52).

b) Another concept of the intonation phrase in Czech is based on the assumption that modelling accents as local phenomena is not impossible in the domain of intonation (though it would be less obvious for duration and intensity). First, when inspecting  $f_0$  curves, one can see that most prenuclear melodic changes occur in the vicinity of stressed

syllables. Of course, this argument is less relevant for short stress units, where all syllables are close to at least one stressed one, but even in longer stress units, intonation changes seem to be connected with the stressed syllable. Second, the alignment of nuclear pitch contours with stress units of different length is not linear, but sensitive to stress location (Palková 1997: 312). Third, other languages with initial stress and distinctive vowel length have already been successfully described by means of tonal events (e.g. Estonian: Asu & Nolan 2003). Fourth, strings of tonal events and contours may be mutually convertible (Ladd 2008: 46). Fifth, an experiment based on automatic prediction of Czech intonation by a neural network has shown that tonal stylization may be as successful as contour-based approaches (Duběda & Raab 2008).

A point should be made on the use of the terms “stress” and “accent” in this paper. “Stress” denotes a potential property of a syllable in the word to become prosodically prominent (accented). “Accent” stands for a perceived prominence triggered by variability in prosodic features.

## 2. “Flat pitch accents”

### 2.1. Occurrence of S\* in read Czech

A set of five pitch accents (L\*H, HL\*, H\*, H\*L and S\*) permitting intonational stylization of read Czech was proposed by Duběda (submitted – a). The analysis primarily aimed at describing prenuclear parts of intonation phrases. The material underlying this study was a recording of a newspaper text containing 410 words, made by five semi-professional speakers (3 men and 2 women, speaking standard Czech). The instruction was to deliver a careful yet lively reading without excessive stylization. The total number of accents analyzed was 1,508.

An illustration of this tonal transcription is displayed in Figure 1.

The third accent, labelled as S\*, is aligned with the syllable [ˈspu:] of the stress unit [ˈspu:vɔdɲi:fɔ]. There is no audible pitch change in the 3-syllable window around this syllable ([na.ˈspu:vɔd]), the small valley between [u:] and [ɔ] being micromelodic. There are no apparent reasons for deaccenting: the word *přívodního* (‘former’ – genitive singular) is a 4-syllable content word, and the syllable [ˈspu:] is by no means exposed to stress clash. Five trained listeners (including the author) evaluated the syllable as perceptually prominent, i.e. accented. There are no other clearly identifiable cues to accent (intensity increase or audible lengthening of the previous syllable [na]).

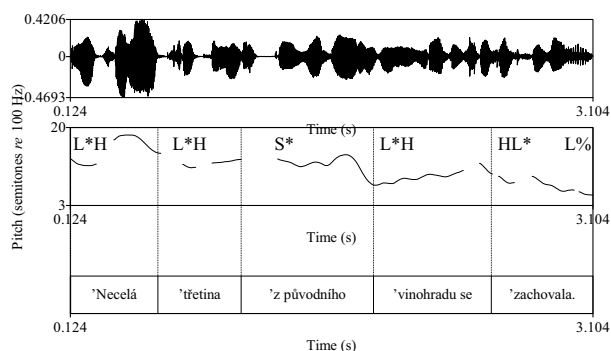


Figure 1: Example of tonal stylization of a Czech sentence ('Less than a third of the former vineyard was preserved.'). Vertical lines delimit stress units.

The tonal labelling of the recordings, carried out by the author, was based on inspecting the  $f_0$  curve and listening to the sound. The S\* label was assigned to syllables perceived as prominent, but showing no intonational change with respect to adjacent syllables which would allow to assign any other pitch accent type.

Such "flat pitch accents" make up about 8% of all pitch accents in the material studied. They occur both in prenuclear (9%) and nuclear (4%) positions. If we consider the five speakers separately, the frequency of S\* accents is 4%, 5%, 6%, 9% and 16%, respectively. In the last speaker, the high percentage of S\* accents is correlated with a global intonational flatness.

## 2.2. Evidence for "flat pitch accents" from other languages

Information about flat or reduced pitch accents may be found in literature devoted to focus and parentheses: in Pierrehumbert (1980: 223), post-nuclear pitch accents are characterized as reduced ("echo" accents). In Vaissière (1999), parentheses are described as having reduced  $f_0$  range and lower register. Furthermore, overall pitch span reduction is given as a correlate of declination and downstep (Vaissière 1983), implying that the probability of finding a reduced pitch accents grows towards the end of the phrase. Martin (2009) claims that intonational contours associated with prosodic words may become neutralized in specific syntactic positions.

Accounts on the phonetics of stress often indicate that there may be interaction and compensation between prosodic parameters, which includes cases where one prosodic parameter is neutralized and some other compensates for it. For instance, Sluijter et al. (1996) argue that in "stress accent languages" (i.e. languages which mark stress not only by intonation; note that the use of the term "stress" differs from ours), a syllable without a pitch accent may still be perceived as prominent due to other parameters (duration, intensity, spectral balance, absence of reduction).

## 3. Analysis and results

In the light of what has been said above, we shall investigate the S\* accents from the following points of view:

- (i) Inter-labeller agreement will be verified so as to confirm whether the syllables which have been assigned an S\* accent by the author are consistently perceived as prominent.
- (ii) Positional and semantic aspects of S\* accents will be analyzed.

(iii)  $F_0$  correlates will be examined.

(iv) Duration, intensity and spectral balance of syllables bearing S\* accents will be investigated.

### 3.1. Inter-labeller agreement

Accent location in the speech material was evaluated by four independent listeners (3 pregradual and 1 postgradual students of phonetics, native speakers of Czech, familiar with auditory analysis). Approx. 10% of S\* accents assigned by the author were not confirmed by the listeners. For the other accent types, this value is only 5%. This would mean that (putative) S\* accents are twice as likely to remain unnoticed than other accent types. However, a  $\chi^2$  test has shown that this difference is not significant ( $\chi^2 = 3.077$ ;  $df = 1$ ;  $p = 0.079$ ). Despite that, we decided to exclude from our analysis all S\* accents which have been identified by less than 3 out of the 5 listeners (including the author). The total number of S\* accents analyzed is thus 112, i.e. approx. 7% of all pitch accents in our material.

### 3.2. Positional and semantic aspects

A detailed positional analysis of all accent types in the material studied was described in Duběda (submitted – b). It has been found that the distribution of S\* accents is not influenced by phrase length or position within the phrase. The only systematic feature is that they do not occur phrase-initially (flat configurations in these positions typically receive the H\* label). This free distribution may seem counter-intuitive, since one would expect a greater frequency of reduced forms of accents towards the end of the phrase, which would be correlated with progressive pitch span narrowing.

As for the semantic characterization of words which have been assigned the S\* accent, 26% of them turned out to have semantic predisposition for deaccenting (see Table 1 for their classification). In the remaining cases, S\* accents are located on content words with no semantic propensity to deaccenting.

Table 1. Words having semantic predisposition for deaccenting which were realized with an S\* accent.

Category	Occurrences
part of compound numeral (e.g. 'ninety' in 'ninety three')	2
postnuclear word (e.g. 'average' in 'twice the average' with narrow focus on 'twice')	1
preposition	3
adverb	7
pronoun	3
auxiliary or modal verb	6
repeated content word (e.g. repetitions of the word 'vineyard', which appears frequently in the text)	7
TOTAL	29

To put this problem in a broader perspective, we marked all words in the text which are of the types displayed in Table 1, to see what the proportion of S\* accents in these words would be. We found 221 such words in the material, out of which only 29 (i.e. 13%) received the S\* accent.

None of the percentages (26% and 13%) speaks in favour of the hypothesis that the distribution of flat pitch accents is systematically influenced by semantic properties of the words.

### 3.3. $F_0$ correlates

To characterize intonational variability in the 3-syllable window around each S\*, we measured the  $f_0$  in the centre of each of the three vowels. To determine which contours are acoustically “flat”, a threshold value is needed. Since there is little unanimity as to this value, we decided to test two different thresholds: we set the “strict threshold” at 1 semitone, and the “loose threshold” at 3 semitones (cf. the discussion in Rosen & Fourcin 1986, and the claim of ‘t Hart, 1974, that linguistically stable intonational changes usually have a minimum span of 3 ST). If the whole  $f_0$  contour in the 3-syllable window fits into the interval of  $\pm 1$  ST around the stressed vowel, the accent can be qualified as “strictly flat”; if it fits into the interval of  $\pm 3$  ST, it is “loosely flat”. Table 1 gives percentages of S\* and non-S\* accents which meet these two conditions.

Table 1.  $F_0$  variability around accents (S\* accents vs. other types of accents; “strictly flat” are accents fitting into the interval of  $\pm 1$  ST around the stressed vowel, and “loosely flat” are those which fit into the interval of  $\pm 3$  ST).

	S* accents	Other accents
“Strictly flat accents”	65%	7%
“Loosely flat accents”	100%	65%

65% of accents labelled as S\* meet the condition of fitting into the “strict” interval of  $\pm 1$  ST, while the remaining 35% have an  $f_0$  difference greater than 1 ST with respect to the syllable on the left or right (only 5% have a difference greater than 1 ST on both sides). An informal analysis of part of them showed that they could often be classified as a different accent type (i.e. other than S\*) if finer perceptual criteria were applied. Others are clearly perceived as flat despite an objective acoustic difference: this is mostly due to syllable-internal  $f_0$  changes which are captured by the acoustic algorithm, but which sound flat when considered globally. On the other hand, only 7% of other accent types (i.e. non-S\*) fulfil the condition of the  $\pm 1$  ST interval in the 3-syllable window around the stressed syllable. Most of these cases can be explained by syllable-internal changes which were not captured by the acoustic algorithm, but which give an impression of inter-syllabic pitch change. Some errors of the labeller were found as well (though in a very limited number).

The “loose” condition of intonational flatness, as proposed by ‘t Hart (1974), seems inadequate: not only would all accents perceived as “flat” meet it, as expected, but 65% of the remaining accents would be classified as “flat”, too.

### 3.4. Duration, intensity and spectral balance

The temporal change between the pre-accentual and accented syllable, as well as that between the accented and post-accentual syllable was measured for each accent in terms of ratio between both syllable nuclei. These were normalized for intrinsic duration (i.e. values for intrinsically long segments such as open vowels were reduced proportionally, and *vice versa*). The results are displayed in Table 2.

Table 2. *Durational variability around accents (S\* accents vs. other types of accents; positive values correspond to lengthening, negative values to shortening).*

	S* accents	Other accents
Difference left vs. accented	−1%	8%
Difference accented vs. right	8%	14%

Syllables bearing an S\* accent show almost no durational variability with respect to the preceding syllable, while they contrast with the following syllable, which is by 8% longer, on the average. Syllables bearing other accent types are typically longer than preceding syllables, and shorter than following syllables. It is interesting to note that all observed variables except for one express lengthening. In other words, there is lengthening between the last syllable of the preceding stress group and the stressed syllable (though not for S\* accents), and another lengthening between the accented syllable and the following one. This cumulative lengthening is probably compensated by shortening of final syllables in longer stress units (otherwise the tempo of the phrase would be progressively slowed down with every new accent). Since this kind of behaviour is somewhat surprising, we validated it by using rough (not normalized) durations of vocalic nuclei: this procedure returned the same kind of results.

S\* are specific in that they are not accompanied by lengthening with respect to the preceding syllable. The difference −1% vs. 8% is significant (t-test,  $p = 0.033$ ). On the other hand, the difference in the lengthening of the following syllable (8% vs. 14%) is not significant (t-test,  $p = 0.130$ ).

Table 3 gives data on the variability of syllable nuclei with respect to intensity. Values obtained by measuring average intensity over each nucleus (in dB) were normalized for intrinsic intensity (i.e. values for intrinsically intense segments such as open vowels were reduced proportionally, and *vice versa*).

Table 3. *Dynamic variability around accents (S\* accents vs. other types of accents; values greater than 1 correspond to intensity increase, values smaller than 1 correspond to intensity decrease).*

	S* accents	Other accents
Difference left vs. accented	−0.02 dB	0.75 dB
Difference accented vs. right	−0.10 dB	−0.09 dB

S\* accents show very little dynamic variability, considering that the just-noticeable difference in intensity is roughly 1 dB. Other accent types, by contrast, are accompanied by an intensity increase of 0.75 dB, on the average. However, this is not an intensity peak (the following syllable has almost the same intensity), but rather a dynamic “reset”. The difference left vs. accented is significant (t-test,  $p = 0.018$ ), the difference accented vs. post-accentual is not (t-test,  $p = 0.098$ ).

According to Sluijter et al. (1996), accents may be realized by different distribution of energy in the spectrum: frequencies above 500 Hz were found to be emphasized in accented syllables in Dutch. Following this proposal, we measured the

average intensity of each nucleus after high-pass filtering at 500 Hz. The results are displayed in Table 4.

Table 4. *Dynamic variability around accents after high-pass filtering at 500 Hz (S\* accents vs. other types of accents; values greater than 1 correspond to intensity increase, values smaller than 1 correspond to intensity decrease).*

	S* accents	Other accents
Difference left vs. accented	−0.76 dB	0.25 dB
Difference accented vs. right	0.17 dB	0.56 dB

None of the differences is significant (t-test for left context:  $p = 0.109$ ; t-test for right context:  $p = 0.182$ ). Interestingly, the difference left vs. accented for S\* accents is negative, i.e. there is no dynamic peak in higher frequencies as suggested by Sluijter et al. (*op. cit.*), but rather a drop.

## 4. Conclusions

Perceived accents with “flat” pitch realization (S\*), despite their reduced form, seem to be perceptually stable, as was shown by high inter-labeller agreement. They occur in all positions in the phrase, and are not correlated with words of reduced semantic content.

The acoustic “flatness” of accents labelled as S\* was confirmed for approximately two thirds of them by applying a relatively strict just-noticeable-difference threshold of 1 semitone. The remaining one third contains cases which could be indeed classified as a different accent type, and cases of  $f_0$  patterns which sound flat despite rapid syllable-internal changes.

In the domain of duration, S\* accents share post-accentual lengthening with other accent types, but differ significantly in the amount of accentual lengthening (i.e. the difference left vs. accented).

In the domain of intensity, S\* accents are not accompanied by a dynamic reset, which property makes them significantly different from other accent types. The analysis of spectral balance showed no significant difference between S\* and other pitch accents. Surprisingly, S\* show a dynamic valley in higher frequencies, rather than a peak.

In summary, S\* accents share some properties of other accent types (no positional constraints, no constraints with respect to the semantic character of the word, very fine intonational changes for some of them, post-accentual lengthening), but differ from other accent types (apart from their reduced intonational form) in the absence of lengthening between the pre-accentual syllable and the accented one, and in the absence of a dynamic reset.

None of the differences which were found between S\* accent and other accent types speaks in favour of a trade-off relation between prosodic parameter, by virtue of which reduced pitch would be compensated by some other parameter. The perceptual stability of S\* accents is thus probably based on fine variations in  $f_0$  and duration, as well as on the perceptual stereotype of hearing the accent word-initially.

Although no stylization method, to our knowledge, includes S\* in the inventory of pitch accents, it is not inappropriate to adopt them into the Czech system, as they correspond to perceived prominence and have non-negligible frequency (8% in our material). They might be a typical

feature of Czech prosody, which is sometimes reported as globally “flat”, and their occurrence may characterize personal prosodic styles (cf. the variation 4–16% for our speakers).

## 5. Acknowledgements

This research was carried out and presented under the GAČR P406/10/0101 grant.

## 6. References

- [1] Hála, B., Uvedení do fonetiky češtiny na obecně fonetickém základě [Introduction to Czech phonetics with elements of general phonetics], Nakladatelství Československé akademie věd, Prague, 1962.
- [2] Duběda, T., “Intensity as a macroprosodic variable in Czech”, in *Speech Prosody 2006*, R. Hoffmann, H. Mixdorff (eds.), Dresden, 185–188, 2006.
- [3] Palková, Z., *Fonetika a fonologie češtiny* [Phonetics and phonology of Czech], Praha, Karolinum, 1997.
- [4] Palková, Z., “Einige Beziehungen zwischen prosodischen Merkmalen im Tschechischen”, in *XIV<sup>th</sup> Congress of Linguists*, Vol. I, Berlin, 507–510, 1987.
- [5] Ladd, D. R., *Intonational Phonology*, Cambridge University Press, Cambridge, 2008.
- [6] Asu, E. L. and Nolan, F., “Testing a model of Estonian intonation”, in *Proceedings of ICPH 2003*, Barcelona, 1249–1252, 2003.
- [7] Duběda, T. and Raab, J., “Pitch accents, boundary tones and contours: Automatic learning of Czech intonation”, in: Sojka, P., Horák, A., Kopeček, I. and Pala, K. (eds.), *Text, Speech and Dialogue 2008*. Springer, 293–302, 2008.
- [8] Duběda, T. (submitted – a) “Towards an inventory of pitch accents for read Czech”.
- [9] Pierrehumbert, J., *The phonology and phonetics of English intonation*, MIT, 1980.
- [10] Vaissière, J., “Utilisation de la prosodie dans les systèmes automatiques: un problème d’intégration des différentes composantes” in *Faits de Langues, Oral – écrit: Formes et théories*, Ophrys, 9–16, 1999.
- [11] Vaissière, J., “Language-Independent Prosodic Features”, in *Prosody. Models and Measurements*, ed. A. Cutler, 53–66, 1983.
- [12] Martin, Ph., *Intonation du français*, Armand Colin, Paris, 2009.
- [13] Sluijter, A. M. C. and van Heuven, V. J., “Spectral balance as an acoustic correlate of linguistic stress” in *J. Acoust. Soc. Am.* Vol. 100/4, 2471–2485, 1996.
- [14] Duběda, T. (submitted – b) “Positional variability of pitch accents in Czech”.
- [15] Rosen, S. M. and Fourcin, A. J., “Frequency Selectivity and the Perception of Speech”, in B. C. J. Moore (ed.), *Frequency selectivity in hearing*. Academic Press, 373–487, 1986.
- [16] ’t Hart, J., “Discriminability of the size of pitch movements in speech”, in *IPO Annual Progress Report 9*, 56–63, 1974.