VOICING IN POLISH: INTERACTIONS WITH LEXICAL STRESS AND FOCUS

Zofia Malisz[1,2], Marzena Żygis[3,4]

[1]Saarland University, [2]Bielefeld University, [3]Centre for General Linguistics, Berlin, [4]Humboldt University, Berlin
zmalisz@coli.uni-saarland.de, zygis@zas.gwz-berlin.de

ABSTRACT

We examine the dynamics of VOT in Polish stops under lexical stress and focus. We elicit real Polish words containing voiced and voiceless stop+/a/ syllables in primary, secondary and unstressed, as well as focus positions. We also correlate VOT with speech rate estimated on the basis of equisyllabic word length. Our results show that the relationships between prosody and VOT are consistent with the status of Polish as a *true voicing* language.
Keywords: voicing, aspiration, Polish, prosody

1. INTRODUCTION

The two-way voicing contrast in Polish is implemented by vocal fold vibration throughout closure in voiced stops and no vibration and low mean VOT values in voiceless stops [3], similarly to Russian [2, 10] and Croatian [11], as well as Spanish, French and Catalan [12]. The following list juxtaposes the main phonetic implementation patterns of stop voicing in Polish with those in English. The two languages respectively show:

- prevoiced (voicing lead) vs. short lag VOT patterns for the voiced phonemes,
- short lag vs. long lag VOT for the voiceless phonemes with little vs. much aspiration noise,
- no overlap vs. overlap in closure durations between voicing categories,
- absence vs. presence of a consonant voicing effect on preceding vowel duration.

The phonetic characteristics in Polish are consistent with a true *voicing* language encoded by the distinctive feature [±*voice*] [1, 2]. In *aspirating* languages such as English, the contrast is implemented by [±*spread glottis*].

It has been hypothesised that in contrast to aspirating languages, VOT values in true voicing languages are relatively less sensitive to, e.g. lexical stress and emphasis because "glottal spreading gestures for aspiration participate in prosodic strengthenings and weakenings, while voicing gestures do not" [4]. Additionally, speech rate proved to affect VOT differently in languages with a [±*voice*] and [±*spread glottis*] based contrast. In French or Spanish ([±*voice*]), prevoicing in voiced stops was found to be longer in slower speech, while in English, long-lag VOT, characterising fortis stops, increased as a function of decreasing speech rate [1, 12]. In both voicing contrast implementations, short-lag VOT (voiceless stops in Spanish, lenis stops in English) remained unchanged under speech rate variation, yielding symmetrical patterns, where only the phonologically specified feature is affected.

Polish has hitherto not been systematically studied with regard to prosodic prominence marking or speech rate effects on VOT. The stimuli used in most previous studies on Polish voicing placed the target consonants in bisyllabic words in a word-initial, focused and stressed position, either post-pausally, by using a word list [3, 4], or phrase-medially, by embedding in a carrier sentence [15]. The work by [4] and [13] additionally reported on VOT in all possible stop+vowel combinations.

The resulting values were found to range from ca. -160 to -10 milliseconds for voiced stops with the mean at -88.2 for /b/, -89.9 for /d/ and -66.1 for /g/ [4]. The prevoicing range was typically narrower in studies with a small number of samples from connected speech and word medially [3]. The VOT range for voiceless stops was found to lie between +10 and +80 miliseconds [4, 15] (overall mean = 34 ms) with the velar /k/ providing the largest mean positive value in [15], at +58 ms (in carrier sentences) and in [4], at +53 ms. The threshold between short- and long-lag VOT is taken to be located at 30 msec and most prototypical true voicing languages exhibit short-lag ranges with maxima at 40 ms [1].

The present work is restricted to stop+/a/ stimuli but instead, stress and focus positions are differentiated, utterance-medially. We study monolingual speakers of Polish from two dialectally homogenous regions. Our main objective is to assess the influence of strong prosodic positions on VOT, testing the predictions for a true voicing language such as Polish in this regard. We also aim to verify observations

Table 1: *A sample of study design: lexical stress position (S: primary = 1, secondary = 2, unstressed = 0) is varied for target syllables (voiceless onsets /pa/, /ta/, /ka/; voiced onsets /ba/, /da/, /ga/) in real Polish words. Focus (F) and IP boundary (IP) conditions are elicited by embedding in meaningful sentences in a Discourse Completion Task (context sentences not presented; target syllables were not boldfaced in the actual experiment).*

Elicited sentence	S	F	IP
Kamila kupiła tuli**pa**ny, a nie róże. `Kamila has bought tulips, not roses.`	1	✓	✓
Nie, to Kamila kupiła or**ga**nizator w supermarkecie, a nie Małgosia. `No, it was Kamila who bought a folder in the supermarket, not Malgosia.`	0	–	–
Kamila kupiła **ta**picerki w supermarkecie, a nie ogumienie. `Kamila has bought upholstering in the supermarket, not tyres.`	2	✓	–
Nie, to Kamila kupiła me**da**liki, a nie Małgosia. `No, it was Kamila who bought the medallions, not Malgosia.`	0	–	✓

that VOT values in Polish might be undergoing a diachronic change towards long-lag or that aspirated stops exist as variants under emphasis [15].

2. METHODS

2.1. Speakers

Speakers from the city (or region) of Wrocław (Lower Silesia) (N=7) and from the city (or region) of Poznań (Wielkopolska) (N=10) were recorded. Speakers were students and university clerks between 19 to 43 years old. They had not lived abroad or outside their native region in Poland for more than a year and were monolingual.

Proficient knowledge of English may influence lag duration in speakers of Polish [15] and is known to impact prevoicing [2]. Care was taken therefore, not to include speakers who had received or entered an intensive foreign language programme and/or pronunciation courses at the time of the recording.

2.2. Experimental design and procedure

Table 1 presents a sample of the study design and materials. The voicing contrast was elicited by placing syllables of the form /b, d, g, p, t, k/+/a/ in real Polish words (lemmata N=42, no compounds). Two sets of words were used: four- and five-syllables long, eliciting secondary stress that in Polish falls on the initial in words longer than three syllables. Each target syllable occurred once in the penultimate position for *primary* stress, in the initial position for *secondary* and in other, non-final position for *unstressed* (voicing contrast is neutralised in final position in Polish). Each target word was embedded in carrier sentences exemplified in Table 1, varying IP boundary (medial vs. final) and focus.

The stimuli list counted 168 sentences in Wrocław, where voiceless and voiced stops were elicited, and 84 in Poznań where only the voiceless subset was recorded. The experiment followed a Discourse Completion Task (DCT) procedure: participants were asked to listen to a pre-recorded context sentence in Polish and to read the stimulus sentence in response. One repetition per factor level combination was recorded per speaker.

2.3. VOT annotation

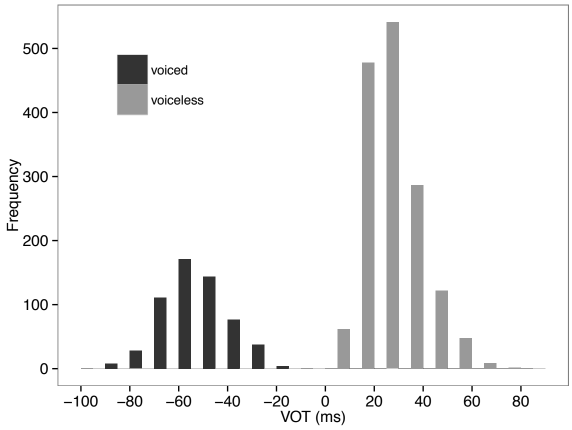
In voiceless stops, VOT was measured from the beginning of the stop release to the voicing onset of the following vowel /a/, marked at the zero-crossing before the first negative peak of pulsation, as in [3]. In voiced stops, VOT was always negative, i.e. voicing was present throughout the closure phase and continued through the usually very short burst to the next vowel. Consequently, all VOT values for voiced stops correspond to the duration of the voiced closure, measured from the offset of formant structure in the preceding vowel until the burst onset [3].

Nonetheless, some of the fully voiced closure phases in /b/, /d/ and /g/ presented a drop in voicing amplitude towards the stop release, visible in the waveform. In these cases, the onset of the drop was marked and the observations indexed.

3. RESULTS

Figure 1 presents the distribution of VOT values measured in voiced and voiceless target stops (N=2130). The respective values are dispersed without overlap, as found by [3, 4]. The distribution of positive VOT for the voiceless tokens is right-skewed with the median at 23.8 ms (SD=11.3). The mean negative VOT for the voiced tokens is -51 ms

Figure 1: The distribution of VOT values for voiced and voiceless stimuli.



(SD=13.2).

The data support cross-linguistic tendencies [6] in that a) the voiced bilabial exhibits the longest mean prevoicing (M=-56.6, SD=12.5) and b) with increasing backness, VOT lag duration increases in voiceless stops: the longest mean positive VOT is found for /k/ = 35.1 ms (SD=10.8), followed by /t/ = 23.7 ms (SD=7.3) and /p/ = 18.4 ms (SD=8.2).

3.1. VOT under focus and lexical stress

We analysed the effects of Stress, Focus, Syllable Count, Phrase Boundary, Place of Articulation (PoA), Gender and City on VOT in voiceless stops by fitting a linear mixed model with these variables as predictors, a random slope (Focus) and random intercepts (Speaker, Word Stimulus). We tested for several interactions. We did not find any significant differences between the dialectal regions that might express themselves in positive VOT values. We further report only significant main effects and interactions.

The results confirm the PoA effect on short-lag VOT in Polish (/k/: reference level; /t/: b=-8.5, SE=2.1, t=-4.1, p<.01; /p/: b=-13.2, SE=2, t=-6.6, p<.001). The model also reveals a significant interaction of Focus with Stress (*focus*primary*: b=2.7, SE=1, t=2.65, p<.05; *focus*secondary*: b=2.2, SE=1, t=2.1, p=.047). Figure 2 presents the subtle differences involved. Mean VOT in primary stress position, in focus, equals 26.3 ms (SD=13) vs. 23.9 ms (SD=10.7) when out of focus.

We also look at the impact of the same factors on the duration of prevoicing in the /ba/, /da/, /ga/ data with random intercepts.

A significant shortening effect of PoA on prevoicing for the alveolar /d/ (b=11.4, SE=2.6, t=4.3, p<.001) and the velar /g/ (b=6.8, SE=2.5, t=2.7,

Figure 2: VOT values for voiceless stops differentiated by lexical stress position and focus.

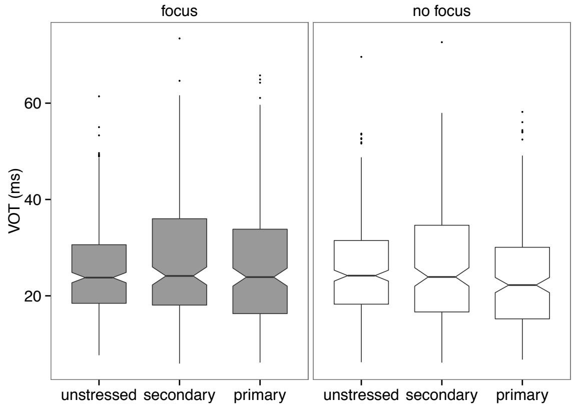
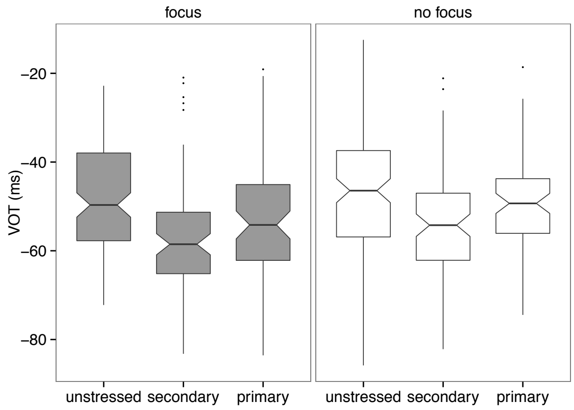


Figure 3: VOT values for voiced stops differentiated by lexical stress position and focus.

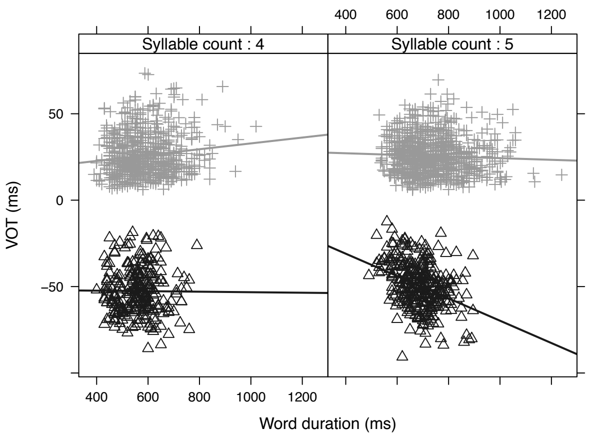


p<.05) relative to the reference bilabial stop, is found. Figure 3 suggests that prevoicing is longer in focus, in all stress positions. The mean VOT for all voiced stimuli in the focus condition is -51.8 ms (SD=13.4) vs. -50 ms (SD=13.1) when not in focus. An independent impact of the secondary stress position on VOT is also visible. The model confirms these observations with a significant main effect of Focus (b=-1.9, SE=0.9, t=-2.1, p<.05) and *secondary* stress (b=-7.6, SE=2.5, t=-3, p<.01).

3.2. Stability of voicing under stress and focus

48.2% (n=280) of voiced stops were voiced throughout the closure with no drop in amplitude. In the remaining cases (n=302), a drop in amplitude towards the end of the closure was observed (see section 2.3). 47.3% of obstruent targets that exhibited the change in amplitude involved the velar /g/, 27.8% the alveolar /d/ and 26.1% the bilabial /b/.

Figure 4: Correlations of word duration and VOT in voiceless (grey crosses) and voiced stops (black triangles) in 4- and 5- syllable words (in panels).



We compare the duration of the steadily voiced interval to the lower amplitude interval as a proportion of the entire closure duration. On average, at least half of the closure interval was filled with stable glottal vibration for all three places of articulation (60% of /b/, 66.5% of /d/ and 51.4% of /g/).

A generalised mixed model was formulated with presence/absence of steady glottal vibration during closure as the binomial response. We analysed the effect of the same predictors as in section 3.1 with random intercepts. We found a significant interaction of focus and /g/, suggesting that prosodic factors moderate the dynamics of glottal vibration during voiced velar closure. Specifically, unsteady vibration was 3.1 times more likely to occur in case of /g/ (OR=exp(1.1), SE=0.45, z=2.5, p<.01), in focus.

3.3. VOT in relation to speech rate

The design of the experiment did not include direct speech rate manipulation, e.g. similar to [1]. However, assuming that slower speech results in longer total word duration, we correlate word duration and VOT for four- and five-syllable long words separately, plotted in Figure 4. Observations below zero correspond to the duration of prevoicing in all voiced stimuli (black triangles), positive VOT observations come from the voiceless set (grey crosses). We find a weak, positive correlation (r(667)=.13, p<.001, r^2=.017) between decreasing speech rate, measured as equisyllabic word duration, and increasing VOT lag, in four syllable words. We also find a relatively stronger, negative correlation (r(318)=-.36, p<.001, r^2=.13) between slowing down and VOT in prevoiced stimuli in five syllable words.

4. DISCUSSION

Our interest lies in the interplay of prosodic factors with segmental detail [7, 8]. We conducted a prosodically differentiated study on VOT in Polish, varying lexical stress, focus and phrase boundary. We replicated previous results [3, 4] with a carefully controlled group of speakers.

First, our data follow universal tendencies regarding the effect of place of articulation on VOT [6, 10]. Furthermore, a lengthening of VOT in Polish short-lag stops occurs in stressed positions under focus. Positive main effects of focus and secondary stress on prevoicing were also found. Effect sizes for these factors are however small, ca. 2-3 ms. Hypotheses regarding the influence of strong positions on the voicing contrast in a true voicing language such as Polish, do predict that prosodic prominence (stress, focus) should not greatly affect VOT values.

The marginally significant result at p=.047 for the secondary stress effect on prevoicing has to be replicated. Also, possible influence of initial word position on VOT and its interaction with secondary stress needs to be considered, given the disputed status of secondary stress in Polish. Other work on this data is exploring these possibilities [7].

We also find that glottal vibration in the voiced velar stop /g/ drops in amplitude before the burst, especially under focus. As voicing in velars is generally hard to sustain [9], a longer closure under focus inhibits filling the interval with steady vibration.

Some authors pointed to emphatic realisations of sequences such as /ki/, that present noticeable aspiration [15], as possible evidence for a sound change towards long-lag in Polish. Our results show that positive VOT over the 30 ms threshold does occur. The surplus is mainly attributable to stops that physiologically trigger longer lags, such as /k/, rather than to strong prosodic positions. Such processes do not attest to a systematic change towards long-lag or prosodically conditioned variation [12]. However, variation that originates in the constraints of the vocal tract “may be reinterpreted by listeners as intended, and result in sound change” [12], we do not however find evidence for such change in Polish.

In true voicing languages positive VOT tends to be constant across rates, while prevoicing increases in slower speech [1, 12]. Correlations between speech rate and VOT in Polish implicate the predicted pattern. [1] note that VOT variation under speech rate resembles the one found for strong prosodic positions. The exact impact of these co-dependent factors on Polish voicing will be addressed in further research (see also [5, 14]).

5. ACKNOWLEDGEMENTS

We would like to thank our speakers, as well as Sergio Quiroz, Katarzyna Dziubalska-Kołaczyk and Joanna Błaszczak. This research has been supported by Bundesministerium für Bildung und Forschung (BMBF, Germany) Grant Nr. 01UG1411 to Marzena Żygis.

6. REFERENCES

[1] Beckman, J., Helgason, P., McMurray, B., Ringen, C. 2011. Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics* 39(1), 39–49.

[2] Helgason, P., Ringen, C. 2008. Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36(4), 607–628.

[3] Keating, P. 1979. *A phonetic study of a voicing contrast in Polish*. PhD thesis Brown University, Providence, RI.

[4] Keating, P. A., Mikoś, M. J., Ganong III, W. F. 1981. A cross-language study of range of voice onset time in the perception of initial stop voicing. *The Journal of the Acoustical Society of America* 70(5), 1261–1271.

[5] Konopska, L., Sawicki, J. 2013. Studies of the VOT parameter in realizations of Polish. Voiceless and voiced plosive phonemes. *Logopedia* (42), 101–122.

[6] Maddieson, I. 1997. Phonetic universals. In: Hardcastle, W., Laver, J., (eds), *Handbook of phonetic sciences*. Oxford: Blackwell 619–639.

[7] Malisz, Z., Wagner, P., Żygis, M. 2015. Acoustic correlates of prosodic structure in Polish. *presented at the 37. Jahrestagung der DGfS* Leipzig. Deutsche Gesellschaft für Sprachwissenschaft.

[8] Malisz, Z., Żygis, M., Pompino-Marschall, B. 2013. Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Journal of the Association for Laboratory Phonology* 4(1), 119–158.

[9] Ohala, J. J. 1983. The origin of sound patterns in vocal tract constraints. In: MacNeilage, P. F., (ed), *The production of speech*. New York: Springer Verlag 186–192.

[10] Ringen, C., Kulikov, V. 2012. Voicing in Russian stops: Cross-linguistic implications. *Journal of Slavic Linguistics* 20(2), 269–286.

[11] Smiljanić, R., Bradlow, A. R. 2007. Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America* 118(3), 1677–1688.

[12] Solé, M.-J. 2007. Controlled and mechanical properties in speech: A review of the literature. In: Solé, M.-J., Beddor, P. S., Ohala, M., (eds), *Experimental approaches to phonology*. Oxford University Press chapter 18, 302–322.

[13] Święciński, R. 2012. *Acoustic aspects of palatalization in English and Polish*. PhD thesis Maria Curie-Skłodowska University Lublin.

[14] Trochymiuk, A. 2008. *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna* [Articulation in hearing-impaired children. An auditive and acoustic analysis]. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

[15] Waniek-Klimczak, E. 2011. Aspiration and style: A sociophonetic study of the VOT in Polish learners of English. *Achievements and perspectives in SLA of speech. Frankfurt: Peter Lang* 303–316.