

THE PHONETICS AND PHONOLOGY OF THE POLISH VOCATIVE CHANT

*Amalia Arvaniti*¹, *Marzena Żygis*², *Marek Jaskula*³

¹University of Kent, UK

²Centre for General Linguistics (ZAS) & Humboldt University, Berlin, Germany

³West Pomeranian University of Technology, Szczecin, Poland

¹a.arvaniti@kent.ac.uk, ²zygis@zas.gwz-berlin.de, ³Marek.Jaskula@zut.edu.pl

ABSTRACT

The Polish vocative chant is a melody used, as in other languages, to call someone for a routine reason. Three repetitions of 12 Polish names (1-4 syllables long, 3 names per syllable count) were elicited from 16 speakers using a Discourse Completion Task. The results, based on data from 11 speakers (6 female), showed that the contour ends in a small rise (not a plateau as in other languages), while the initial rise aligns with the end of the stressed syllable. In polysyllabic names, the contour starts with low F₀ and the rise begins immediately before the (always penultimate) stress. Based on these data and comparisons with contours used in wh-questions and in other types of Polish vocatives, we analyse this contour as LH* !H-H%. The implications of this analysis for autosegmental models of intonation are discussed.

Keywords: vocative chant, intonation, Polish.

1. INTRODUCTION

The vocative chant, a type of melody described as *stylized* by Ladd [11], has been investigated in many languages, including English [4], Hungarian [13], Dutch [9], Catalan [3] and Portuguese [6]. Here we examine the Polish vocative chant. As in other languages, the vocative chant is used in Polish to call a person for a routine reason; cf. [11]. Impressionistically, the Polish melody sounds similar to comparable melodies in English [4], German [8], and Dutch [9]. Preliminary inspection, however, suggested essential differences with respect to the melody's end. As can be observed in Figure 1, the F₀ contour remains low until the beginning of the stressed syllable and also shows a final rise, unlike similar melodies in other languages which end in a plateau [4], [8], [9]. We were thus interested in investigating the following questions. Is the initial rise associated with the stressed syllable and if so, what form does this association take? Is the final rise simply a return to a default mid-level? If so, one would expect it to be dramatically reduced under tonal crowding, possibly leading to plateaux in crowding contexts.

2. METHODS

2.1. Speakers

Data were elicited from 16 native speakers of Standard Polish (8 female). The speakers' age ranged from 18 to 23. They were all volunteers and naïve as to the purposes of the experiment; none reported a history of speech or hearing disorders.

2.2. Materials

The materials were Polish names 1 to 4 syllables long; see Table 1. Polysyllabic names were stressed on the penult. Three names per syllable count were selected; some included mostly sonorants, e.g. *Natalia*; others included either voiceless fricatives, e.g. *Jaś*, or voiceless stops, e.g. *Piotr*. The purpose of incorporating such differences was to observe the effect of tonal crowding in Polish and the impact of the contour on the acoustics of fricatives in particular (in comparison with fricatives in the scolding contour discussed briefly in sections 2.3. and 4. below; these results are not reported here).

Table 1: The names used in the study.

No of sylls	Names
1	<i>Jan Jaś Piotr</i> [jan] [jaɕ] [pjɔtr]
2	<i>Daniel Bartek Patryk</i> [ˈdajɛl] [ˈbartɛk] [ˈpatrɨk]
3	<i>Natalia Marianna Małgośka</i> [naˈtalja] [maˈrjana] [mawˈgɔɕka]
4	<i>Magdalena Alexandra Małgorzata</i> [magdaˈlɛna] [alɛkˈsandra] [mawgɔˈzɨtka]

2.3. Procedures

Recordings were conducted in a sound-proof room at West Pomeranian University of Technology, Szczecin, Poland. Three repetitions were elicited using a Discourse Completion Task (DCT). The speakers were asked to imagine calling a child (i) to dinner, (ii) to get a present, or (iii) because she had

broken a vase. Contexts (i) and (ii) elicited the vocative chant, discussed here; context (iii) elicited the “scolding” contour, the data from which are discussed in detail in [2]. Speakers generally found the DCT easy, but some were not consistent in their use of the vocative chant which they replaced with other melodies (cf. [3]). For this reason, data from 5 speakers were discarded and the results reported here are based on 11 speakers (6 female).

2.4. Measurements and statistical analysis

F0 was measured in ERB to reduce differences between male and female speakers while retaining audible differences between the two. Measurements were taken at specific points in the contour (barring obvious microprosodic perturbations; see Figure 1). The points measured were:

- *IT*: the F0 at the onset of the contour
- *RO*: the F0 at the onset of the rise to the first peak; in 1- and 2-syllable words, *IT* and *RO* coincided
- *H1*: the F0 of the first peak
- *L*: the F0 minimum in the contour’s dip
- *H2*: the F0 maximum at the contour’s end

In addition, the distance of *RO*, *H1* and *L* from the following segmental landmarks was measured:

- The distance of *RO* from the onset of the stressed vowel (SVO): $t_{SVO}-t_{RO}$
- The distance of *H1* from the onset of the stressed vowel expressed as a proportion of the vowel’ duration: $(t_{H1}-t_{SVO})/dur_{SV}$
- The distance of *L* from *H1* and *H2*: $t_{H1}-t_L$ and $t_{H2}-t_L$ respectively
- The distance of *L* from the onset of the last vowel (which coincided with the stressed vowel in monosyllables), as a proportion of the vowel’ duration: t_L-t_{LVO}/dur_{LV} .

Since the data from our female and male speakers differed substantially in their distributions, statistical calculations were conducted separately for each group. We used linear mixed-effects models with syllable count as a fixed effect and participant and item as random effects. All p-values reported in the paper are pMCMC values.

3. RESULTS

Preliminary inspection of the data indicated that in polysyllabic words the vocative chant is realized on the last foot (i.e. the last two syllables), with low F0 before the stressed syllable, rising F0 on the stressed penult and a reduced rise on the ultima; see Figure 1. F0 comparisons in 3- and 4-syllable words confirmed this observation. *IT* and *RO* were similar in scaling (see Figure 2), and stable across syllable

count [for *IT*: males, $t = 1.51$, females: $t = -1.68$; for *RO*: females, $t = 0.33$). The only significant difference was found in male speakers for *RO* (mean *RO* in 3.34 ERB 3-syllable and 3.66 ERB in 4-syllable words; $t = 3.053$, $p < 0.05$). The alignment of *RO* was also unaffected by syllable count [males: $t = -0.463$, females: $t = -0.044$], though somewhat variable: it occurred on average 82 ms (s.d. = 35.1) before the stressed vowel onset for males, and on average 93 ms (s.d. = 39.4) for females.

Figure 1: Illustration of measurements.

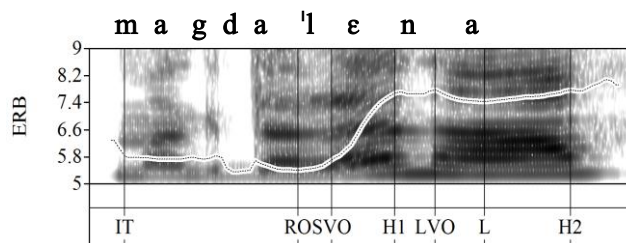
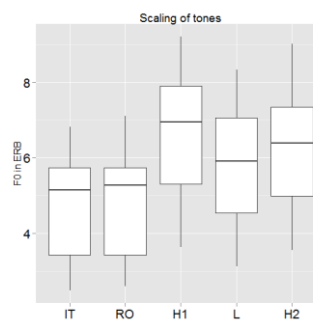
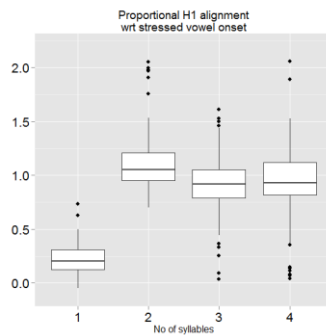


Figure 2: Scaling of Initial Tone, Rise Onset, H1, L and H2; data pooled over words and participants.



The data also confirmed that the Polish vocative chant involves a final rise; see Figure 2. *H2* was scaled significantly lower than *H1* but still significantly higher than *L* [for *H2* vs. *H1*: females, $t = -2.8$, $p < 0.01$; males, $t = -3.004$, $p < 0.01$; for *H2* vs. *L*: females, $t = 14.6$, $p < 0.001$; males, $t = 14.2$, $p < 0.001$). These differences between *H1* and *L* and *L* and *H2* were largely unaffected by tonal crowding: they held at all syllable counts for both male and female speakers [all pairwise comparisons were significant at $p < 0.001$], indicating that the melody is not realized as a plateau under tonal crowding. Rather, undershoot under tonal crowding involved *H1* instead which showed lower scaling in monosyllables compared to longer words [1 syll. vs. 2 sylls: males, $t = -2.492$, $p < 0.01$; females, $t = -4.841$, $p < 0.001$; 1 syll. vs. 3 sylls.: males, $t = -5.155$, $p < 0.001$; females, $t = -6.167$, $p < 0.001$; 1 syll. vs. 4 sylls.: males, $t = -5.705$, $p < 0.001$; females, $t = -4.085$, $p < 0.001$].

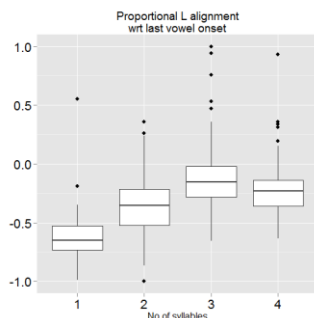
Figure 3: Proportional alignment of H1 with respect to the stressed vowel onset.



Regarding alignment, H1 was consistently aligned with the end of the stressed vowel, except in monosyllables where it occurred close to the vowel's onset [males: $p < 0.001$ for all comparisons, except 3- vs. 4-syllables, n.s.; females: $p < 0.001$ for all comparisons involving 1-syllable words; all other comparisons were n.s.]; see Figure 3.

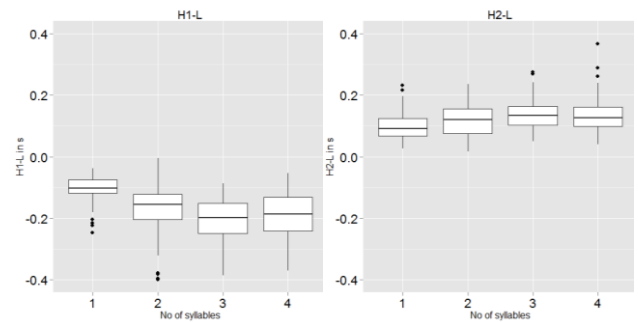
L alignment was also affected by syllable count. L aligned later in 1- and 2-syllable words (roughly in the middle of the last vowel) than in 3- and 4-syllable words where it occurred close to the last vowel's onset [males: $t = 2.662$, $p < 0.001$; females: $t = 3.468$, $p < 0.001$]; see Figure 4.

Figure 4: Proportional alignment of L with respect to the last vowel onset (this vowel coincides with the stressed vowel in monosyllables).



Results also show that L was realized closer to H2 than to H1 (males: on average, 121 ms for tH2-tL vs. 146 ms for tL-tH1; females: on average, 121 ms vs. 188 ms respectively). The temporal distance tH2-tL was significantly shorter in monosyllables as opposed to 3- and 4-syllable words in male speakers (1 vs. 3, $t = -1.701$, $p < 0.05$; 1 vs. 4, $t = -1.649$, $p < 0.05$). No differences were found for female speakers. The difference tL-tH1 was larger for monosyllabic words in comparison to 2-, 3- and 4-syllable words for female speakers (1 vs. 2, $t = 2.2$, $p < 0.05$; 1 vs. 3, $t = 2.8$, $p < 0.01$; 1 vs. 4, $t = 2.5$, $p < 0.01$) and in comparison to 3- and 4-syllable words for male speakers (1 vs. 3, $t = 3.9$, $p < 0.001$; 1 vs. 4, $t = 3.15$, $p < 0.001$). See Figure 5.

Figure 5: Distance of L from H1 (left) and H2 (right).



4. DISCUSSION

Overall, the results confirmed the preliminary conclusions drawn from the inspection of the data. The results from IT and RO confirmed that the contour remains low until the onset of the stressed syllable, if there is a sufficient number of syllables to do so (3 or more); if not, the contour still starts low but rises quickly. The ensuing pitch peak is aligned with the stressed vowel and is followed by a dip and a final rise of reduced range (male speakers 0.5 ERB; female speakers: 0.33 ERB on average) and high in the speaker's frequency span. Tonal crowding had a small but significant effect limited to monosyllables which showed earlier alignment and lower scaling of H1, and later alignment of L. The difference between L and H2 was present even in monosyllables, indicating that the final rise is deliberate and not a return to some default.

Based on the above, the proposed autosegmental analysis of the Polish vocative chant is LH* !H-H%. The initial rise, as it is aligned with the stressed syllable, is analysed as a LH* pitch accent. The use of a bitonal representation is necessary here as the melody remains consistently low at the beginning. This contrasts with the "scolding" contour which clearly lacks the initial low F0; cf. Figures 1 and 7.

The dip and rise after the accentual peak are analysed as a sequence of a downstepped !H- phrase accent and a H% boundary tone. The phrase accent is analysed as !H-, rather than L-, as it is scaled much higher than the uncontroversially L tone of the pitch accent (cf. Figure 2). In addition, evidence from other Polish melodies, like the wh-question melody in Figure 8, suggests that Ls are not upstepped between Hs even in extreme tonal crowding. Thus, the dip in the vocative chant cannot be plausibly attributed to an upstepped L tone. Finally, the use of H% reflects the small but consistent final rise of the melody.

Figure 7: The “scolding” contour.

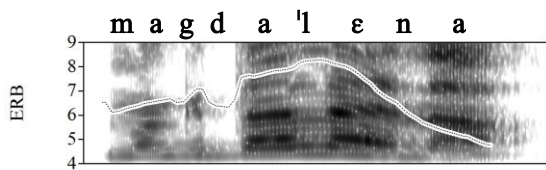
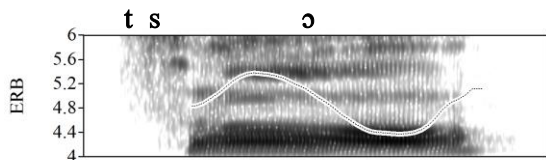


Figure 8: The wh-question contour; *co* “what?”



The analysis of the contour’s final part as consisting of two distinct tonal events is based on the following observations. First, it is clear from the results on the temporal and F0 distance between the L target (!H-) and the two H tones that the L is not strongly connected to the accentual H (H1); thus it is unlikely to be a part of the melody’s pitch accent. Second, inspection of other Polish melodies (e.g. the wh-question melody as realized in long utterances; see [2]) indicates that tones between the last pitch accent and the boundary tone spread. This behaviour is consistent with a phrase accent [7]. If phrase accents are a necessary part of the Polish intonational system, parsimony dictates a uniform analysis involving phrase accents for all melodies.

The autosegmental representation of the Polish vocative chant differs from analyses of similar melodies in other languages shown in Table 2. As this kind of diversity has been deplored, e.g. [11], [12], it is worth examining these differences in some detail. Inspection of Table 2 quickly shows that differences are of three types: (i) some are a result of differences between intonational systems; (ii) others reflect distinct theoretical positions; (iii) others are a matter of convention. The use of LH* for Polish reflects the first type: as indicated above, Polish clearly has a contrast between an accent best analysed as H* and a bitonal LH* and uses LH* in the vocative chant. This can be juxtaposed to English, where a similar contrast obtains [4] but a H* pitch accent is used instead in the vocative chant. Similarly, the analysis of the Dutch vocative chant as H*!H% reflects the fact that the melody applies in steps to successive feet when possible [7], [9]. An analysis of this sort, though suitable for Dutch, would be clearly inappropriate for Polish in which the entire melody is confined to the last foot even in long vocatives (see [2] for details). Further, as shown, the final part of the Polish contour is rising; thus the representation of this part of the melody cannot be the same as the plateau of Dutch or

German. On the other hand, the use of two tonal events to represent the final rise !H- and H% reflects the second type of disagreement about whether phrase accents are needed as an entity (cf. [4] and [9]; for a discussion see [12]). Finally, the use of LH* instead of L+H* has been largely a notational difference, i.e. a matter of convention (though [10] have recently proposed to reserve the use of representations with the plus sign for accents showing a loose connection between their tones).

Table 2: AM representations of the vocative chant.

Language	Pitch Accent	Phrasal Tones
Arabic [5]	L+H*	H-L%
Catalan [3]	L+H*	!H%
Dutch [9]	H*!H*	%
English [4]	H*	!H-L%
German [8]	L+H*	H-%
Greek [1]	L*H	!H-!H%
Hungarian [13]	H*	!H-0%
Polish	LH*	!H-H%
Portuguese [6]	(L+)H*	!H%

In short, while some differences in notation can mask cross-linguistic similarities of form (e.g. % and 0% in Table 2 both represent a lack of F0 change), not all differences can be attributed to notational vagaries. This in turn suggests that any crosslinguistic standardization of autosegmental modelling can only be confined to differences in formal representation, not to unresolved theoretical disagreements or differences between phonological systems.

5. CONCLUSION

In conclusion, the phonetics of the Polish vocative chant together with a comparison with other Polish melodies lead to an autosegmental representation of the former as LH* !H-H%. This representation reflects the unambiguously low onset of the accentual rise, the lack of copying of the melody ([2]; cf. [7], [9]), and its clear (if typologically unusual) final rise. The realization of the vocative chant in Polish leads to clear differences in its representation relative to representations used in other languages for functionally similar melodies. Such differences apply across the board, not just to Polish, and are due to a number of reasons, including both differences in phonological systems and theoretical assumptions, on the one hand, and differences in the form of representation (rather than the analysis per se), on the other. Standardization of the latter is viable and clearly desirable, but systemic differences will always lead to principled differences in analysis.

Acknowledgments

This research has been supported by University of Kent research funds to Amalia Arvaniti and Marzena Żygis, and by Bundesministerium für Bildung und Forschung (BMBF, Germany) Grant Nr. 01UG1411 to Marzena Żygis.

5. REFERENCES

- [1] Arvaniti, A., Baltazani, M. 2005. Intonational analysis and prosodic annotation of Greek spoken corpora. In: Jun, S. (ed), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 84-117.
- [2] Arvaniti, A., Żygis, M., Jaskuła, M. The phonetics and phonology of Polish vocatives. Forthcoming.
- [3] Borràs-Comes, J., Sichel-Bazin, R., Prieto, P. Vocative intonation preferences are sensitive to politeness factors. *Language and Speech* (Special issue on the Experimental approaches to the production and perception of prosody).
- [4] Brugos, A., Shattuck-Hufnagel, S., Veilleux, N. 2006. *Transcribing Prosodic Structure of Spoken Utterances with ToBI* (MIT Open Courseware). <http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/>
- [5] Chahal, D., Hellmuth, S. 2014. The Intonation of Lebanese and Egyptian Arabic. In: Jun, S. (ed), *Prosodic Typology II*. Oxford: Oxford University Press, 365-404.
- [6] Frota, S., Gruz, M., Fernandez-Svartman, F., Vigário, M., Collischonn, G., Fonseca, A., Serra, C., Oliveira, P. in press. Intonational variation in Portuguese: European and Brazilian varieties. In: Frota, S., Prieto, P. (eds), *Intonation in Romance*. Oxford: Oxford University Press.
- [7] Grice, M., Ladd, D. R., Arvaniti, A. 2000. On the place of "phrase accents" in intonational phonology. *Phonology* 17. 143-185.
- [8] Grice, M., Baumann, S., Benzmüller, R. 2005. German Intonation in Autosegmental-Metrical Phonology. In: Jun, Sun-Ah (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press, 55-83
- [9] Gussenhoven, C., Rietveld, T., Kerkhoff, J., Terken, J. 2003. ToDI second edition. <http://todi.let.kun.nl/ToDI/home.htm>
- [10] Jun, S., Fletcher, J. 2014. Methodology of studying intonation: from data collection to data analysis. In: Jun, S. (ed), *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 493-519.
- [11] Ladd, D. R. 2008a. *Intonational Phonology*. Cambridge: Cambridge University Press.
- [12] Ladd, D. Robert. 2008b. Review of S. Jun (ed) (2005). *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press. *Phonology* 25, 372-376.
- [13] Varga, L. 2008. The calling contour in Hungarian and English. *Phonology* 25(3): 469-497.