INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

The Vocative Chant and Beyond: German Calling Melodies under Routine and Urgent Contexts

Sergio I. Quiroz[1], *Marzena Żygis*[2]

[1] Leibniz-ZAS & Berlin School of Mind and Brain, Berlin, Germany
[2] Leibniz-ZAS & Humboldt-Universitaet zu Berlin, Berlin, Germany

seisquiroz@gmail.com, zygis@zas.gwz-berlin.de

Abstract

This paper investigates calling melodies produced by 21 Standard German native speakers on a discourse completion task across two contexts: (i) routine context—calling a child from afar to come in for dinner; (ii) urgent context—calling a child from afar for a chastising. The intent of this investigation is to bring attention to other calling melodies found in German beside the vocative chant and to give an insight to their acoustic profile.

Three major melodies were identified in the two contexts: vocative chant (100% of routine context productions), urgent call (100% of male urgent context productions, 52.2% female productions), and stern call (47.8% female urgent context productions). A subsequent quantitative analysis was carried out on these calls across these parameters: (i) tonal scaling at tonal landmarks; (ii) proportional alignment of selected tonal landmarks with respect to the stressed or last vowel; and (iii) amplitude (integral and RMS) and (iv) duration of the stressed vowel, stressed syllable, and word. The resulting data were analyzed using a linear mixed model approach.

The results point to significant differences in the contours produced in the aforementioned parameters. We also proposed a phonological description of the contours in the framework of Autosegmental-Metrical Phonology.

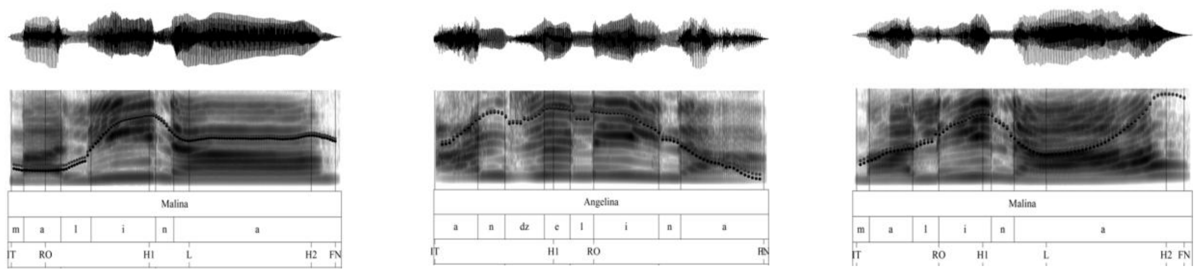
Index Terms: calling melodies/contours, vocative chant, German, acoustic/phonetic description

1. Introduction

Calling melodies are usually understudied in intonational literature, with a clear bias towards the vocative chant [1]. The vocative chant is a "sweet" melody associated with calling children, or intimate in-group members, from afar in routine contexts, such as coming to the table to dinner (see [2, pp. 116-119] for additional contexts). This melody is also known as the calling contour [2]. There are studies, however, that show the vocative chant is not the only calling melody available to speakers (see for Catalan: [3]; Turkish and Austrian German: [4]; Polish: [5], and German: [6], [7], [8], [9]). The main goal of this paper is to expand the understanding of calling melodies found in German across routine and urgent calling contexts. Unlike previous studies in German dealing with the vocative chant (see previous citations), this one contributes new acoustic measurements not yet applied to this chant. In particular this study reports on two calls not yet reported on for Standard German of Germany: the urgent call and the stern call. See Figure 1 for waveform and spectrogram versions of these three melodies.

2. Methods

Following [5] a discourse completion task (DCT) was used in which participants called eight names under a routine or urgent context (see section 2.2). Both contexts were bound by the overarching scenario that the speaker imagines being inside the house and has to call a child-cousin who is playing with a ball outside in the garden. Under (i) the routine context, the speaker calls the child in for dinner, while under (ii) the urgent context, the child is called in to be chastised for breaking a window. The DCT used picture-cues. Participants were taught to associate a black-and-white line drawing of a dinner plate and silverware with the routine context, while a line drawing of a broken window prompted the urgent context. Following the example set by [3], participants called names based on gender: female participants called only female names and male participants called only male names.



a. Vocative Chant (Routine Context) *b. Urgent Call (Urgent Context)* *c. Stern Call (Urgent Context)*

Figure 1: *Calling melodies in German (and context in which they occur) as produced by female speakers.*

http://dx.doi.org/10.21437/Interspeech.2017-1635

2.1. Participants

A total of 24 participants completed a full experiment session. Sessions from 3 participants were not used for further analysis due to completing an incorrect version of the experiment (N= 2 females) or due to heavy audio distortion during audio conversion (N= 1 male). We report herein only on the sessions from the other 21 participants.

All 21 participants (gender: N= 13 females, 8 males; mean age: x= 26.143 y.o., sd= 6.134) were German native speakers of Standard German who were living at the time of the experiment in Berlin, Germany. They reported having no hearing or speech impediments.

2.2. Stimuli

Two sets of 8 names were selected based on gender (Table 1). The names can be subdivided into 4 (mostly) sonorant names and 4 names with voiceless stops. The latter set of names was included to explore the interaction of intonation and Voice Onset Time not covered in the scope of this paper. Only the 3- and 4-syllable male names have antepenultimate stress, all other male and all female names have penultimate stress.

Table 1: *Female and male German names.*

Syll. Count	Female	Male
1	Linn [lɪn]	Finn [fɪn]
2	Mina [ˈmi.na]	Maron [ˈma.ʀɔn]
	Tara [ˈta.ʀa]	Thorsten [ˈtɔʀs.tɛn]
		Thorben [ˈtɔʀ.bɛn]
3	Malina [maˈli.na]	Florian [ˈflo.ʀi.jan]
	Tabea [ta.ˈbe.a]	Christopher [ˈkʀis.to.fer]
	Antonia [anˈto.nia]	Constantin [ˈkɔns.tan.tin]
	Viktoria [vikˈto.ria]	
4	Angelina [an.dʒɛˈli.na]	Immanuel [ɪˈma.nu.el]

2.3. Procedure

The recordings were made in the sound-proof booth of the Leibniz-Centre for General Linguistics in Berlin. Participants were seated about 40 cm away from a laptop computer screen which was used to display the experiment designed on MatLab [10] with the PsychToolbox package [11]. A microphone was placed about 20 cm away from a participant's mouth. Participants were given instructions and also briefly practiced what each picture cue meant by calling out the names "Linn/Finn" and "Mina/Maron", respective to participant's gender.

The experiment consisted of 64 trial runs (8 names * 4 repetitions * 2 contexts). Each trial run included a fixation cross slide, a picture-cue slide, and a stimuli slide in that order. The stimuli slide was displayed until the participant clicked the space bar to continue—all other slides were displayed automatically for 0.5 seconds. Sessions were recorded on DAT tapes and were then converted into WAV files using Adobe Audition 1.5.

2.4. Calling melody contours

A total of 1344 productions were reviewed (64 productions * 21 speakers) using Praat [12]. Productions were categorized into groups mainly based on the F0 pitch shape accompanied by a perceptual analysis. This resulted in three main categories (see Figure1): (a) the vocative chant (based on [6]): F0 has a high pitch followed by a sustained plateau and reduced rise on the ultima; (b) the urgent call (based on [5]): F0 starts low, rises to a high point and then falls back to a low pitch, and (c) the stern call (no previous report on this call could be found). Roughly, the stern call has an F0 contour which starts low, then rises high, falls back down low to almost the same level as the first low, and finally rises back up to the same level as the first high.

Dropped from further analysis were: (i) productions with one of the three F0 pitch shapes but which occurred in a context uncommon to the majority; (ii) productions with an F0 shape other than the main three so infrequent that not enough tokens could be gathered; (iii) productions which were corrupted due to digital conversion and noise during recording session; (iv) and productions with shifted H1 alignment (see section 3.1.1. for further discussion). This resulted in 938 productions (female: N= 711, male: N= 227) for further analysis.

2.5. Acoustic measurements across melodies

2.5.1. Tonal landmark measures

The measurements used here follow closely those set by [5]: (i) pitch scaling of the tonal landmarks; (ii) proportional alignment of the tonal landmark H1 with respect to the stressed vowel and (iii) of the tonal landmarks L and H2 with respect to the last vowel. Tonal landmarks are points occurring on the F0 pitch contour of an utterance. Table 2 expands on how they are defined in this study.

Table 2: *Tonal landmarks.*

Tone	Description
IT	Initial point of F0 contour
RO	Onset of the rise to the H1
H1	First maximum of F0
L	First minimum of F0 after H1
H2	Second maximum of F0 (if present)
FN	Final point of F0 contour

The tones in Table 2 were labeled manually across all productions of the three calls. If a production did not have a specified tone, the missing tone was grouped with a neighboring tone. In the case of the RO in vocative chants, the IT was used a surrogate in one and two-syllable words. For the L and the H2 of the urgent call, if not present, the FN value was used. Finally, for the stern call, the H2, when not present, was collapsed with the FN.

2.5.2. Additional measures

Productions were additionally analyzed across amplitude and duration by looking at the: integral, RMS, and duration of the stressed vowel, stressed syllable, and word. The RMS and duration of the stressed vowel were normalized with respect to the RMS and duration of the corresponding word. The integral is a summation of the duration and RMS amplitude (see [5] and work cited therein for further details).

2.5.3. Statistical analysis

A mixed effects model approach was used to analyze the extracted values. The R Studio environment [13] was used to carry out statistical analysis with additional use of the lme4 and lmer Test packages [14, 15].

Models were built for the tonal and other measures mentioned above and their significance was based on a Likelihood Ratio Test. Fixed factors for the acoustic measures are: gender [male, female], and call type [vocative chant, urgent call, stern call]; additionally for the tonal landmark models: tone type [IT, H1, L, H2, FN]. The random effects were maximized [16]. In the selected final model participant was used as a random effect. Importantly, we will focus on statistically significant differences.

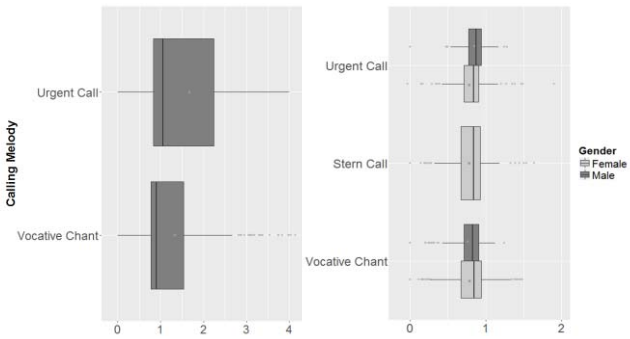
3. Results

3.1. Calling melody frequency

In the routine context (female: N= 369 productions; male: N= 197 productions), participants produced exclusively a vocative chant. Meanwhile, the urgent context (female: N= 347 productions; male: N= 182 productions) elicited an urgent call from all male speakers (100% from urgent context productions) and most female speakers (N= 181 productions, 52.2%). The alternative melody for females under the urgent context was the stern call (N= 166 productions, 47.8%).

3.2. Tonal landmarks

3.2.1. Proportional alignment



a. Male data (unadjusted) *b. All data (adjusted male data)*

Figure 2: *H1 proportional alignment across calling melodies.*

As seen from Figure 2a, the male data shows that H1 aligns within and beyond the stressed vowel in both urgent calls and vocative chants (note: a value between 0 (onset) and 1 (offset) indicates that the tone was realized within the stressed vowel). Upon close inspection of the data, the multisyllabic names (152 productions) are the cause of this variability with H1 aligning on vowels other than the stressed one.

Those cases are omitted from further analysis as mentioned in section 2.3. With this adjustment made, see Figure 2b, the male data (M) show that H1 aligns near the end of the stressed vowel across the vocative chant (VOC) and urgent call (URG) with no significant difference between the two melodies Similarly, female speaker data (F) show that the proportional alignment of the H1 occurs near the end of the stressed vowel across the vocative, urgent, and stern (STR) melodies without any significant difference between them, see Figure 2b.

The proportional alignment of the H2 and L tonal landmarks, Figure 3, show differing alignment patterns across the three melodies (note: a value between 0 (onset) and 1 (offset) indicates that the tone was realized within the final vowel). In the urgent call the L tone is produced slightly earlier than the H2 respectively for both genders, but the difference is not significant. The stern call's L is produced almost a quarter of the way after the final vowel onset while the H2 is significantly later towards the end of the last vowel (L-H2= -0.666, t= -12.885, p< 0.001). The L tone of the vocative chant is produced halfway into the last vowel. Meanwhile the H2 is produced at the end of the last vowel for females and slightly after the last vowel in male productions (M-F= 0.299, t= 4.265, p< 0.001).

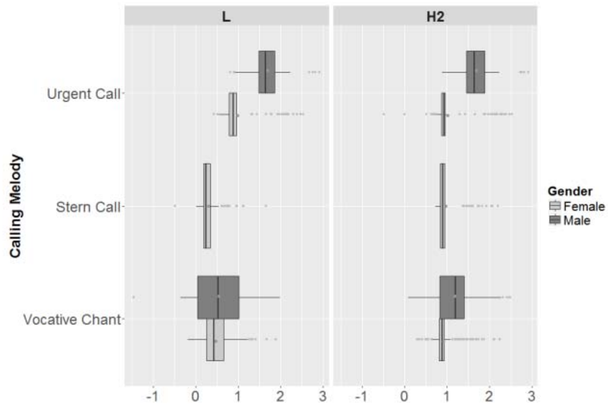


Figure 3: *L and H2 proportional alignment across calling melodies.*

3.2.2. Tonal scaling

Figure 4 shows a visual summary of the pitch scaling for the three identified calls across gender while significant values are reported in Table 3.

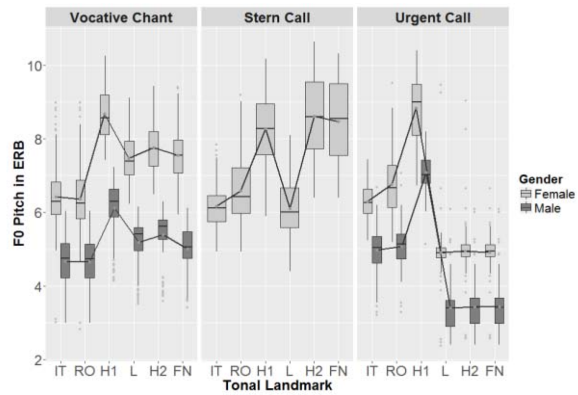


Figure 4: *Scaling of tonal landmarks*

Across both genders, the H2 of the vocative chant is significantly higher than its two surrounding tonal neighbors. While the FN is slightly higher than the L in female productions, this is not the case for males, where the FN is lower than the L. The scaling of the urgent call shows that L, H2, and FN are not significantly different from one another. Additionally, the IT is significantly higher than the FN across both genders. In the case of the stern call, the FN has a slightly lower pitch than the H2. The H1 is significantly lower than the H2. In this melody the L is slightly lower than the IT, but the difference is not significant.

Table 3: *Significant results for tonal scaling.*

Gender	Contour	Tones	B	t-val.	p-val.
Female	VOC	FN-H2	-0.202	-4.925	< 0.001
		L-H2	-0.284	-6.980	< 0.001
		FN-L	0.082	1.998	< 0.05
	URG	IT-FN	1.382	23.232	< 0.001
	STR	FN-H2	-0.145	-2.422	< 0.05
		H1-H2	-0.336	5.525	< 0.001
Male	VOC	FN-H2	-0.393	-5.691	< 0.001
		L-H2	-0.210	-3.080	< 0.01
		FN-L	-0.183	-2.658	< 0.01
	URG	IT-FN	1.532	18.745	< 0.001

3.3. Duration, integral and RMS measures

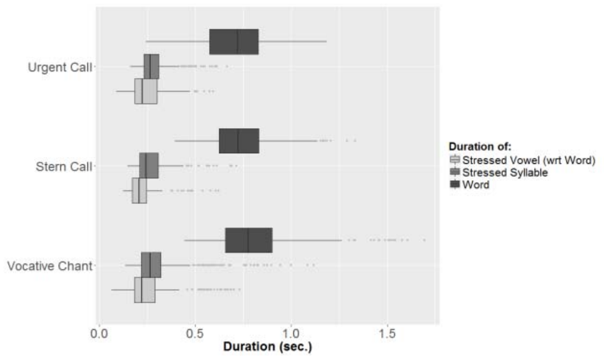


Figure 5: *Duration measures.*

At the word level, the urgent call is shortest (F: STR-URG= 0.034, t= 4.143, p< 0.001, VOC-URG= 0.0950, t= 16.316, p< 0.001; M: VOC-URG= 0.116, t= 14.389, p< 0.001) followed by the stern call as second shortest (F: VOC-STR= 0.061, t= 9.746, p< 0.001). Similarly, the stressed syllable is shortest in the urgent call (F: STR-URG= 0.023, t= 3.251, p< 0.01; VOC-URG= 0.030, t= 6.053, p< 0.001; M: VOC-URG= 0.011, t= 1.615, p= 0.106) with no significance in the difference between the longer vocative vs. stern stressed syllable (F: VOC-STR= 0.007, t= 1.351, p= 0.177). The pattern continues down the level of stressed vowel duration with a p< 0.001 significance, except in the difference between female urgent and stern calls (F: STR-URG= 0.002, t= 0.266, p= 0.791).

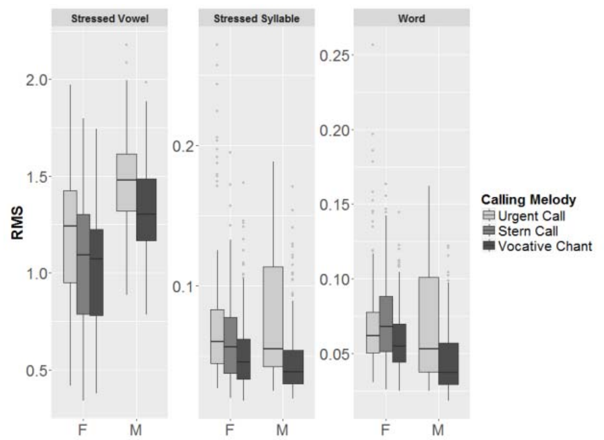


Figure 6: *RMS measures.*

The urgent call shows the highest RMS value (F: STR-URG= -0.006, t= -5.629, p< 0.001; VOC-URG= -0.0137, t= -19.064, p< 0.001; M: VOC-URG= -0.018, t= -18.249, p< 0.001) the stern call (F: STR-VOC= -0.008, t= -10.340, p< 0.001). This trend persists as so throughout the RMS for the stressed syllable and stressed vowel with p< 0.001. Results similar to the RMS have been obtained for the integral. The exception is that the integral of the stressed syllable is higher in the vocative chant than the stern call, but this was not significant (F: VOC-STR= 0.009, t= 1.307, p= 0.191).

4. Discussion

This paper reports three different Standard German calling melodies (vocative chant, urgent and stern call) excerpted in two contexts (routine and urgent). To the best of our knowledge, the urgent and stern calls have not been reported before.

Although our data shows that the stern call is limited to female speakers, further studies with differing contextual cues might tease out instances in male German registers. The F0 contour of the stern call hits the identified tonal landmarks at the upper and lower extremes of the speaker's pitch range. Duration measures indicate the stern call lasts longer than the urgent call, but shorter than the vocative. With regard to the amplitude, the stern call as a whole seems to have the same range as an urgent call. However, the amplitude values of the stressed syllable and vowel for the stern call show that the amplitude ranges between the urgent and vocative melodies.

The urgent melody is brief and has very few salient tonal landmarks so that other than the start and end, only a peak is essential.

Taking into account our results as well as the German ToBI system (GToBI; see e.g. [17]) we analyze the vocative chant melody as L+H* !H-% (see [6] for the same representation) and propose L+H* L-H% for the representation of the stern melody and L+H* L-% for the representation of the urgent melody. Further perceptual experiments should (dis)confirm our proposals.

From a cross-linguistic perspective, German vocative chant does not differ in its meaning from chants in most other languages investigated thus far, see e.g. [5] for an overview. However, it does differ, for instance, from Catalan [3] where such melodies are perceived as insisting calls. The phonology of German urgent call shares many similarities to that of Polish and other languages [see 5: Table 2].

5. Conclusions

This paper expands on the acoustic structure of under-reported German intonational contours: the urgent and the stern call. Additionally it has reported values that can contribute to the extant literature on German vocative chants ([4, 6-9, 17]). It also provides results of other measurements (duration, amplitude) which contribute to stating differences between the calls. Future studies should disclose whether the stern call is indeed limited to female speakers and reveal fine-grained differences between the melodies with respect to factors such as age, social status, situative context and others.

6. Acknowledgements

We would like to thank our participants for taking part in this study. This research has been supported by Bundes-ministerium für Bildung und Forschung (BMBF, Germany) Grant Nr. 01UG1411 to Marzena Żygis.

7. References

[1] Liberman, M. (1979). The Intonational System of English. New York: Garland.
[2] Ladd, D. R. (2008). Intonational phonology. Cambridge: Cambridge University Press. (Original plublication date 1996).
[3] Borràs-Comes, J., Sichel-Bazin, R., & Prieto, P. (2015). Vocative intonation preferences are sensitive to politeness factors. Language and Speech, 58(1):68-83.
[4] Göksel, A. & Pöchtrager, M. (2013). The vocative and its kin: marking function through prosody. In Sonnenhauser, B., & Hanna, P. N. A. (Eds.), Vocative!: addressing between system and performance (Vol. 261). Berlin: Walter de Gruyter.
[5] Arvaniti A., Żygis M., Jaskuła M. (2016). The Phonetics and Phonology of the Polish Calling Melodies. Phonetica, 73:338-361.
[6] Grice, M., Baumann, S., & Benzmüller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In: Jun, S-A. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing, 55-83. Oxford: Oxford University Press.
[7] Niebuhr, O. (2015). Stepped intonation contours: A new field of complexity. In: Skarnitzl, R. & Niebuhr, O. (Eds.) Tackling the complexity of speech. Prague: Charles University Press.
[8] Niebuhr, O. (2013). Resistance is futile - The intonation between continuation rise and calling contour in German. Proceedings of the 14th Interspeech Conference, Lyon, France, pp. 225-229.
[9] Schaeffler, F., & Biersack, S. (2003). Aspects of the timing of fundamental frequency in German chanted call contours. Phonum, 9, 129-132.
[10] MathWorks, (2012). Bioinformatics Toolbox: User's Guide (R2012a). Retrieved July 14, 2012 from www.mathworks.com/help/pdf_doc/bioinfo/bioinfo_ug.pdf
[11] Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. Perception, 36(14), 1.
[12] Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.22, retrieved 15 November 2016 from http://www.praat.org/
[13] RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.
[14] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1): 1-48. doi:10.18637/jss.v067.i01.
[15] Kuznetsova, A., Brockhoff, P. B., & Christensen, R.H.B. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. https://CRAN.R-project.org/package=lmerTest
[16] Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of memory and language, 68(3), 255-278.
[17] Grice, M. & Baumann, S. (2002). Deutsche Intonation und GToBI. Linguistische Berichte, 191: 267-298.