Modifications of tongue body kinematics as a focus marking strategy in German

Lena Pagel, Simon Roessig, Doris Mücke

University of Cologne, Germany

{lena.pagel, simon.roessig, doris.muecke}@uni-koeln.de

Abstract

The present study investigates articulatory modulations of prosodic prominence. Therefore, tongue body kinematics of 27 German speakers are analysed using Electromagnetic Articulography. Our data show that tongue positions and partly also peak velocities are systematically modulated in order to increase prominence. When examining the vertical and horizontal movement dimension, we observe a combination of different highlighting strategies involving the enhancement of sonority and place features for low and back vowels. The results suggest that the parameter modifications are not only a concomitant of accent but rather reflect a continuous increase in prosodic prominence from background through broad and narrow to contrastive focus. We therefore conclude that the tongue body is involved in prosodic strengthening in German and encodes varying degrees of prosodic prominence.

Keywords: focus marking, prominence, electromagnetic articulography, prosodic strengthening, tongue kinematics.

1. Introduction

In order to guide the listeners' attention, speakers can produce parts of an utterance as more prominent and therefore signal information structure. Those parts which are assumed to be the most informative for the listeners are in *focus* and distinguished from those in the *background* (Lambrecht, 1994). An enhancement of prominence, as in focus marking, can be achieved by the placement of a pitch accent which is attained by adaptations of the laryngeal system. Additionally, speakers can modulate their articulation through systematic adjustments of the supralaryngeal system. The latter has been subsumed under the term *prosodic strengthening* which describes the spatio-temporal expansion of oral vocal tract gestures in order to increase prosodic prominence on focussed words (Cho, 2002).

Two prevalent strategies of prosodic strengthening have been described, namely *sonority expansion* (Beckman, Edwards & Fletcher, 1992) and *localised hyperarticulation* (de Jong, 1995). Sonority expansion is connected to a greater opening of the vocal tract in vowels which allows for more acoustic energy to radiate from the mouth. This strengthens the syntagmatic contrasts between consonants and vowels (Beckman, Edwards & Fletcher, 1992; Cho, 2005; de Jong, 1991). Localised hyperarticulation, on the other hand, is associated with an enhancement of place features. Hyperarticulated vocalic targets lead to more distinct sound productions, strengthening the paradigmatic contrasts between different vowels. Tongue kinematics are modified in the vertical and the horizontal movement dimensions, i.e. the vocalic target can be moved up/down or front/back. Studies on prosodic strengthening report hyperarticulated tongue body movements in several vowels, e.g. a lower tongue body position in /a/, a higher or more fronted position in /i/ or a retraction in /ʊ/ (de Jong, 1991; 1995; Cho, 2002; 2005; Harrington, Fletcher & Beckman, 2000; Kim & Cho, 2011; Katsika, 2018). When approaching a more peripheric vocalic target under prominence, articulators have been observed to also move with higher peak velocities (Mücke & Grice, 2014; Katsika, 2018), though not all investigations come to the same results (Cho, 2002).

While a solid body of research reports differences between accented and unaccented words, only few studies have investigated detailed modifications in accented words with differing degrees of prosodic prominence. For German, Mücke and Grice (2014) examine strategies of prosodic strengthening in four focus conditions including one in which the target word is in the background and hence unaccented. In the remaining three conditions, the target word appears in broad, narrow or contrastive/corrective focus and receives the nuclear pitch accent in all three cases. The results reveal an increase of prosodic strengthening in the articulatory domain from broad to narrow and from narrow to contrastive focus. These findings provide evidence that articulatory cues are not only modulated as a function of accentuation but encode fine-grained degrees of prominence in speech production.

The investigation of Mücke and Grice (2014) is restricted to lip kinematics. Other studies, as already mentioned above, have analysed tongue kinematics but compared unaccented to accented words only (Cho, 2002; de Jong, 1995; Harrington, Fletcher & Beckman, 2000). The present study aims to fill this gap by examining tongue kinematics *across accentuation*, i.e. between unaccented vs. accented words, as well as *within accentuation*, i.e. different degrees of prosodic prominence between words that bear a nuclear pitch accent. On the one hand, the present work contributes to our understanding of the concomitants of accentuation by adding tongue kinematics data for German. On the other hand, it sheds light on the flexibility and the granularity of prosody-driven modulations of speech kinematics.

2. Methods

2.1. Participants

27 native speakers of German participated in the experiment. At the time of the recording, the subjects were between 19 and 35 years old and 17 identified as female, 10 as male. The participants grew up monolingually, did not report any speech or hearing impairments and did not have any special training in phonetics, phonology or prosody.

2.2. Speech material

The speech material consisted of 20 disyllabic German-sounding nonce target words. The first syllable always carried the lexical stress and served as the target syllable. It had a CV structure with a bilabial, labiodental or lateral consonant and /a/ or /o/ as the target vowel. The target words were embedded in controlled and consistent carrier sentences and could either be in the background (unaccented) or in broad, narrow or contrastive focus (accented). This made it possible to compare the realisation of target words across as well as within accentuation.

2.3. Recording procedure

Participants were recorded with Electromagnetic Articulography (Carstens AG501). Simultaneous acoustic recordings were carried out using a head-mounted condenser microphone

(AKG C520). To track the articulators' movements, sensors were placed on the lower and upper lip, chin, tongue tip, tongue blade and tongue body, alongside reference sensors. After the preparation of the subject, the recording session lasted about 45 minutes, including a training session. For the present paper, only the tongue body sensor will be analysed. Further data, including intonational data, are published in Roessig and Mücke (2019).

Participants were seated in front of a screen and engaged in an interactive game. They answered to a character's questions in order to retrieve hidden items. Four types of questions served as triggers to elicit the focus structure of the subjects' answers. Exemplary question-answer pairs are shown in table 1, in which the target word is underlined and the focus domain is marked by brackets.

Table 1: *Exemplary question-answer pairs eliciting four focus conditions.*

Background:	Q:	Hat er die Säge auf die Wohse gelegt?
		Did he put the saw on the Wohse?
	A:	Er hat [den Hammer]$_F$ auf die Wohse gelegt.
		He put [the hammer]$_F$ on the Wohse.
Broad:	Q:	Was hat er gemacht?
		What did he do?
	A:	Er hat [den Hammer auf die Wohse gelegt]$_F$.
		He [put the hammer on the Wohse]$_F$.
Narrow:	Q:	Wo hat er den Hammer hingelegt?
		Where did he put the hammer?
	A:	Er hat den Hammer [auf die Wohse]$_F$ gelegt.
		He put the hammer [on the Wohse]$_F$.
Contrastive:	Q:	Hat er den Hammer auf die Mahse gelegt?
		Did he put the hammer on the Mahse?
	A:	Er hat den Hammer auf die [Wohse]$_F$ gelegt.
		He put the hammer on the [Wohse]$_F$.

2.4. Measures

We investigated tongue body movements during the vowels /a/ and /o/ on the vertical and horizontal dimension by analysing two parameters: *mean position* during a selected window and *peak velocity*. The acoustic boundaries for the target syllable and the respective target vowel were labelled manually. For the mean position, all position values during the first half of the acoustic vowel were averaged, as depicted in figure 1. This measure was employed since local turning points in the trajectories were difficult to identify consistently in the data. The mean position is able to capture the spatial tongue movement during the vowel without depending on an extremum position. As a second parameter, the peak velocity of the articulatory gesture was manually labelled and analysed. All values were z-score normalised for each speaker and vowel and then compared between focus conditions for each vowel separately.

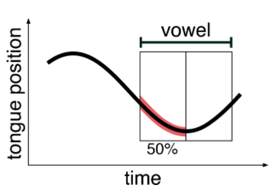

Figure 1: *Schema of mean tongue position measure.*

The results are analysed using Bayesian linear mixed models in R (R Core Team, 2018) with the package brms (Bürkner, 2018). Separate models for each articulatory parameter and vowel are constructed with the articulatory parameter as the dependent variable and focus type as fixed effect. All models include random intercepts for speakers and target words as well as by-speaker and by-target-word slopes for the effect of focus type. We used the default priors of the brms package. The estimated differences between focus conditions in terms of posterior means (Δ) are reported in addition to 95% credible intervals (CI) (table 2, 3). For the tongue body position, the probability of the estimated difference being smaller than zero is reported [Pr(Δ<0)], for the peak velocity, the probability of the estimated difference being greater than zero is reported [Pr(Δ>0)]. We consider probabilities from 0.95 upwards as representing convincing evidence for a difference between focus types.

3. Results

For an introduction to the data, the recorded tongue body trajectories are qualitatively examined. Position values are automatically extracted during the target syllable, normalised for each speaker and vowel and averaged across conditions. The mean trajectories are shown in figure 2 for syllables with /a/ and /o/ in four focus conditions and two movement dimensions. The data reveal that throughout the syllable, the tongue body travels towards the vocalic target position involving a lowering and retraction of the tongue in the vertical and horizontal dimension.

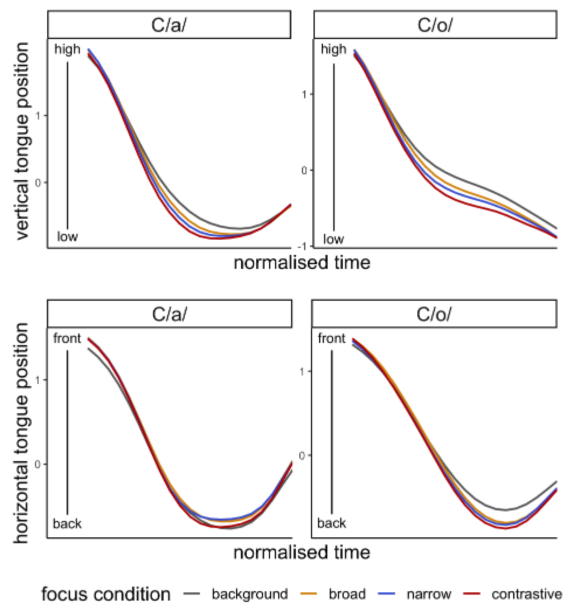

Figure 2: *Mean trajectories of normalised tongue body during target syllable.*

The visual inspection of the trajectories reveals subtle but consistent differences between focus conditions in three out of four cases. The low central vowel /a/ exhibits a lower vertical extremum position from background through broad and narrow to contrastive focus, revealing a gradient tongue lowering with increasing prominence. In the horizontal dimension, no clear trend can be observed for /a/. For the close-mid back vowel /o/, no extremum position can be identified in the vertical dimension. This may be due to the fact that the following vowel (schwa) requires and even lower tongue position. Nevertheless, the trajectories indicate a gradient tongue lowering between all focus types with increasing prominence. In the horizontal dimension, /o/ is produced with a retracted extremum from background through broad and narrow to contrastive focus.

The analysis of the mean position mirrors these qualitative observations. The following figures display the distributions with bars representing the means (left) and the means in more detail (right) of the mean positions. The results for each vowel and movement dimension are compared between background (bg), broad (br), narrow (na) and contrastive focus (co). Figure 3a shows the results for /a/, figure 3b for /o/. Taken together, the data reveal a systematic lowering of the tongue in both vowels and a retraction in /o/. These modifications occur across accentuation (background vs. broad) but also within accentuation between broad and contrastive, with narrow ranging in-between.

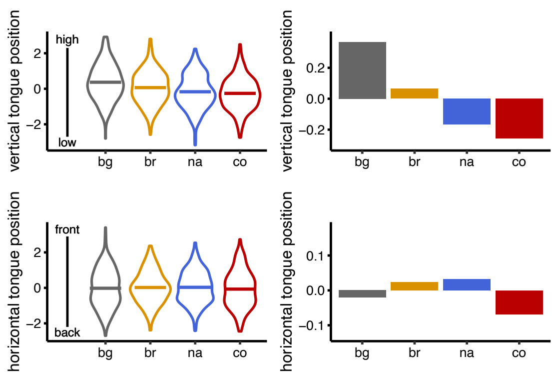

Figure 3a: *Mean normalised tongue position in /a/.*

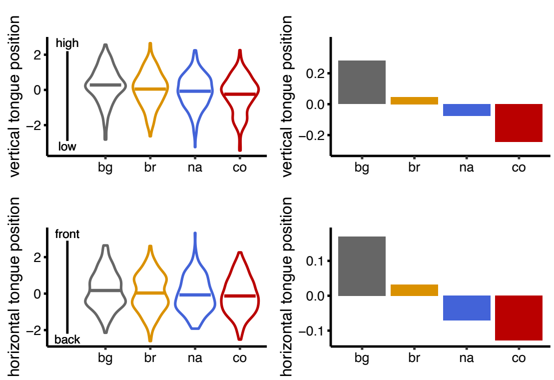

Figure 3b: *Mean normalised tongue position in /o/.*

In the vertical dimension, the Bayesian model provides evidence that the tongue is lowered in both vowels across accentuation (background vs. broad) and within accentuation (broad vs. contrastive). Narrow focus seems to take a middle position between broad and contrastive: While it patterns with contrastive in /a/, it patterns with broad in /o/ (with a probability of 0.95, the evidence for a difference between narrow and contrastive is generally a little weaker than between broad and contrastive). Unsurprisingly, the statistical analysis does not yield clear differences for the horizontal dimension in /a/. However, it provides evidence for differences in /o/ across accentuation (background vs. broad) and within accentuation (broad vs. contrastive). Narrow, again, takes a middle position. The model does not provide convincing evidence for a clear differentiation of narrow from either broad or contrastive.

Table 2: *Bayesian model results for mean position.*

vertical tongue body position in /a/			
	Δ	*95% CI*	*Pr(Δ<0)*
bg vs. br	-0.31	[-0.46, -0.16]	1.00
br vs. co	-0.32	[-0.49, -0.14]	1.00
br vs. na	-0.23	[-0.42, -0.05]	0.98
na vs. co	-0.08	[-0.26, 0.10]	0.77

horizontal tongue body position in /a/			
	Δ	*95% CI*	*Pr(Δ<0)*
bg vs. br	0.04	[-0.10, 0.19]	0.30
br vs. co	-0.08	[-0.24, 0.08]	0.80
br vs. na	0.02	[-0.14, 0.17]	0.43
na vs. co	-0.10	[-0.25, 0.06]	0.85

vertical tongue body position in /o/			
	Δ	*95% CI*	*Pr(Δ<0)*
bg vs. br	-0.23	[-0.39, -0.08]	0.99
br vs. co	-0.29	[-0.47, -0.11]	0.99
br vs. na	-0.30	[-0.30, 0.05]	0.88
na vs. co	-0.17	[-0.34, 0.00]	0.95

horizontal tongue body position in /o/			
	Δ	*95% CI*	*Pr(Δ<0)*
bg vs. br	-0.14	-0.28 -0.02	0.97
br vs. co	-0.17	-0.31 -0.02	0.97
br vs. na	-0.11	-0.25 0.04	0.89
na vs. co	-0.06	-0.20 0.09	0.74

The distributions and means of the normalised peak velocities are shown in figure 4a for /a/ and figure 4b for /o/. As they are based on absolute velocities, higher values indicate faster movements. In both vowels and dimensions, peak velocities seem to show a tendency towards an increase under prominence.

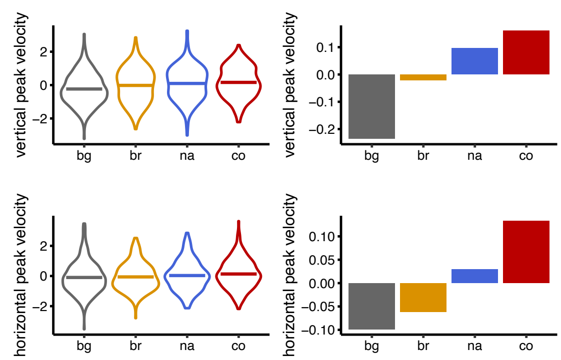

Figure 4a: *Normalised peak velocity of /a/-gesture.*

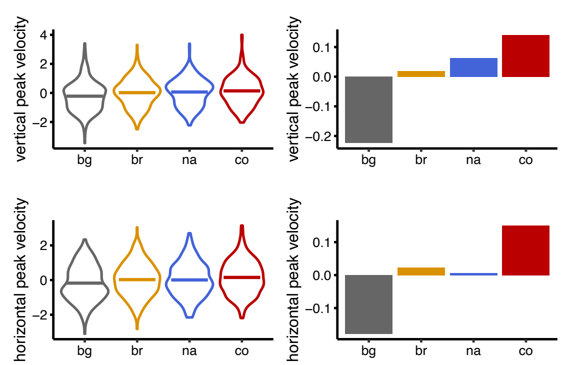

Figure 4b: *Normalised peak velocity of /o/-gesture.*

While the Bayesian analysis provides evidence for faster vertical movements across accentuation (background vs. broad) and within accentuation (broad vs. contrastive) in /a/, for /o/, there is only convincing evidence for a difference across accentuation. In addition, in both vowels, no compelling evidence for a differentation of narrow from broad or contrastive focus can be reported. For the horizontal peak velocity, the analyses only provide convincing evidence for faster movements across accentuation in the vowel /o/.

Table 3: *Bayesian model results for peak velocity.*

vertical tongue body peak velocity in /a/			
	Δ	*95% CI*	*Pr(Δ>0)*
bg vs. br	0.21	[0.07, 0.36]	0.99
br vs. co	0.19	[0.03, 0.35]	0.97
br vs. na	0.13	[-0.04, 0.29]	0.90
na vs. co	0.06	[-0.09, 0.22]	0.77
horizontal tongue body peak velocity in /a/			
	Δ	*95% CI*	*Pr(Δ>0)*
bg vs. br	0.06	[-0.09, 0.21]	0.74
br vs. co	0.17	[-0.01, 0.35]	0.94
br vs. na	0.06	[-0.13, 0.26]	0.69
na vs. co	0.11	[-0.11, 0.32]	0.82
vertical tongue body peak velocity in /o/			
	Δ	*95% CI*	*Pr(Δ>0)*
bg vs. br	0.27	[0.11, 0.42]	1.00
br vs. co	0.11	[-0.05, 0.27]	0.87
br vs. na	0.03	[-0.15, 0.20]	0.61
na vs. co	0.09	[-0.08, 0.25]	0.81
horizontal tongue body peak velocity in /o/			
	Δ	*95% CI*	*Pr(Δ>0)*
bg vs. br	0.23	[0.08, 0.38]	0.99
br vs. co	0.13	[-0.04, 0.30]	0.89
br vs. na	-0.02	[-0.21, 0.15]	0.41
na vs. co	0.15	[-0.03, 0.34]	0.91

4. Discussion and conclusion

The present study investigates how the tongue body is modified as a means of prosodic strengthening in German and whether it can encode focus structure across or also within accentuation.

The recorded articulatory data of 27 speakers provide evidence that tongue body kinematics are indeed modified in German when marking focus and that these modifications are not exclusively mediated by accentuation. Within *and* across accentuation, the tongue position is lowered in /a/ as well as /o/ and retracted in /o/. There is furthermore evidence that peak velocities of the vertical vocalic movements increase across accentuation in both vowels and even within accentuation in /a/. While the spatial differences within-accentuation are clearly observed between broad and contrastive, narrow seems to lie in-between the two. In some cases, it patterns with broad, in others with contrastive or is not distinguished consistently from either. Since the current analyses are not able to explain if this is because narrow is interpreted differently by individual speakers, the status of this condition has to be examined further in future research. All in all, the results suggest that prosodic strengthening by means of the tongue body has the potential to directly express focus structure. Therefore, the present study supports the findings of an earlier study by Mücke and Grice (2014) who observe modifications in lip kinematics between broad, narrow and contrastive focus for five German speakers. Notably, in contrast to the authors who do not find clear modulations between background and broad focus, the present data also reveal adaptations across accentuation. This might be due to the fact that the data set in the present study is considerably larger and therefore less sensitive to speaker-specific variability.

A closer look at the spatial modulations suggests that speakers apply both prosodic strengthening strategies described in the literature. Although the jaw and the lips are the primary articulators associated with sonority expansion, lowering the tongue body also contributes to a greater opening of the vocal tract. Indeed, the present data indicate a further lowering under prominence in both vowels /a/ and /o/. As /a/ is a low vowel, this can be interpreted as simultaneous sonority expansion and localised hyperarticulation strategies, enhancing the sonority and place feature at the same time. However, since the close-mid vowel /o/ is not generally produced with a low tongue body target, the tongue lowering under prominence is likely to be least partly aiming at expanding sonority. As for the horizontal movement dimension, the retraction of the tongue in the back vowel /o/ results in a more peripheric vocalic target and can therefore be associated with localised hyperarticulation to enhance the place feature. Note, that /a/ is produced as a central vowel in German and therefore does not need to be further contrasted to neighbouring vowels in the horizontal dimension. It is therefore expected that /a/ only exhibits a more extreme target in the vertical dimension under prominence. Taken together, the results suggest that primarily those parameters are modified under prominence which lead to either an expansion of sonority or to hyperarticulated target productions. The increase in peak velocity may be an effect of the greater distance that the tongue body travels towards the more peripheric target positions.

To conclude, the present study provides evidence for prosodic strengthening by means of the tongue body in German. Furthermore, it reveals that the articulatory cues of the tongue can encode several degrees of prominence.

5. Acknowledgements

This work was supported by the German Research Foundation (DFG) as part of the SFB1252 ”Prominence in Language” in the project A04 at the University of Cologne.

6. References

Beckman, M. E., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Gesture, Segment, Prosody* (pp. 68-89). Cambridge University Press.

Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1).

Cho, T. (2002). *The Effects of Prosody on Articulation in English*. Routledge.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ,i/ in English. *The Journal of the Acoustical Society of America, 117*(6), 3867-3878.

de Jong, K. (1991). *The oral articulation of English stress accent.* Dissertation, Ohio State University.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America, 97*(1), 491-504.

Harrington, J., Fletcher, J., & Beckman, M. E. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. B. Broe & J. B. Pierrehumbert (Eds.), *Acquisition and the Lexicon* (pp. 40.51). Cambridge University Press.

Katsika, A. (2018). The kinematic profile of prominence in Greek. *Proceedings of the 9th International Conference on Speech Prosody*, 13-16 June, Poznań, Poland.

Kim, S., & Cho, T. (2011). Articulatory Manifestation of Prosodic Strengthening in English /i/ and /ɪ/. *Phonetics and Speech Sciences*, *3*(4), 13-21.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge University Press.

Mücke, D. & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation: Is it mediated by accentuation? *Journal of Phonetics*, *44*(1), 47-61.

R Core Team. 2018. R: *A language and environment for statistical computing*. Vienna, Austria. http://www.r-project.org/.

Roessig, S., & Mücke, D. (2019). Modeling Dimensions of Prosodic Prominence. *Frontiers in Communication, 4*, 44.