Rhythm influences the tonal realisation of focus

Nadja Schauffler[1], *Katrin Schweitzer*[2]

[1]Institute of English Linguistics, University of Stuttgart, Germany
[2]Institute for Natural Language Processing, University of Stuttgart, Germany
nadja.schauffler@ifla.uni-stuttgart.de, katrin.schweitzer@ims.uni-stuttgart.de

10.21437/Interspeech.2015-21

Abstract

Several studies suggest that rhythm affects different aspects in speech production and perception. For example, in German, discourse structure is normally marked by pitch accent placement and pitch accent type, however, there is variation that cannot be explained by purely semantic or syntactic factors. Prosody-inherent factors, like rhythm, can contribute to this variation. This becomes evident in prosodically more complex environments: while the prosody of utterances containing one focused constituent is well investigated and rather clear-cut, the prosodic organisation of multiple contrastive foci is less clear. In double-focus constructions, for example, two focused constituents demand prominence, possibly resulting in the realisation of two pitch accents. If these pitch accents are required on adjacent syllables they conflict with rhythmic preferences. We present a sentence reading experiment investigating the tonal realisation of two focused constituents and how their contours affect each other in different rhythmic environments. Specifically, we tested whether a potential pitch accent clash in a sentence with two corrective foci influences the pitch excursion and the absolute peak height of the accented syllables. The results demonstrate that rhythmic constraints affect the organisation of the tonal marking of corrective focus.

Index Terms: rhythm, prosody, contrastive focus, melodic effects, F_0 parametrisation

1. Background

It is well accepted that information structure and prosody interrelate. For instance, in German, information structure can be marked by both accent placement and accent type (e.g. [1, 2, 3]). Focused constituents are (typically) prosodically marked by a pitch accent, like in “Does Mike eat cakes? – No, he BAKES cakes.” – where we expect a contrastive accent on “bakes” (capital letters indicate focus expected to be marked by pitch accent throughout the paper). However, the actual choice of pitch accent type as well as the distribution of pitch accents in a phrase can vary beyond of what discourse structure can explain. Several studies suggest that prosody-inherent factors, like rhythm, contribute to this variation in pitch accent placement. Speakers apparently use deaccentuation in order to avoid two accented syllables (e.g. [4, 5]), or two stressed syllables directly following each other (e.g. [6, 5, 7]). Therefore pitch accents may be omitted, or shifted in order to avoid a clash. For instance, [4] found cases in a German radio news corpus where pitch accents are shifted from the noun to the adjective in order to prevent an accent clash due to another accented syllable following the noun.

The purpose of the present experiment is to investigate the prosodic realisation of two adjacent contrastive accents in double-focus environments, like e.g. “Does Mike eat cakes? - No, he BAKES BREAD”. In particular, we test how double-accent realisations are influenced by rhythmic factors, namely the avoidance of pitch accent clashes. Given the findings by the aforementioned studies, we hypothesise that speakers will avoid the realisation of a pitch accent when its production would lead to a clash. To test our hypothesis, an experiment based on a sentence-reading task was designed to investigate whether double-focus constructions always surface with two pitch accents regardless of rhythmic constraints. It is important to note that for this data, there are no pitch accent labels available yet – the current analysis solely concentrates on acoustic features. The parameters we investigate, absolute peak height and pitch excursion, have been shown to be cues used to convey prominence and to correlate with contrastiveness (cf. e.g. [8, 9, 10]). We expect that rhythm affects the use of these cues. In double-focus constructions two foci are expected to be prosodically marked, potentially by two pitch accents. We expect that it plays a role how close together they are.

We are aware that we cannot infer from purely acoustic parameters to perception. Perception will be included in future work in order to get a comprehensive picture of contrastiveness. Nevertheless, the current analysis gives insight into how the rhythmic environment influences the tonal organisation of prosodic focus marking. In our test-cases the foci were on sentential objects which were corrections to previously mentioned objects. It has been debated how an accent marking contrastive focus looks like phonetically and phonologically (see e.g. [11, 12]). Our study investigates if prosody-inherent factors, like rhythm, need to be taken into account in addressing such questions.

2. Data elicitation

The data for the current study was elicited via a reading production experiment.

2.1. Participants

Sixteen (5 men, 11 women) German native speakers participated in the experiment. Their mean age was 27.25 years (range: 19 to 33) and none of them had known speech or reading disorders. All participants were naïve as to the purpose of the experiment. They were paid for their participation.

2.2. Stimuli

In order to test the prosodic realisation of double-focus environments, a set of stimuli was constructed containing question-answer pairs. The mini-dialogues were designed in such a way that two objects introduced in the question needed to be corrected in the answer. To investigate influences of the rhythmic environment, two conditions per sentence type were included:

one eliciting pitch accents on successive syllables, which therefore directly follow each other (**clash** condition) and one condition in which the potential pitch accents are separated by an unaccented intervening syllable (**no clash** condition). Examples are given in (1); lexically stressed syllables are underlined. Since phrase length matters in the distribution of accents (cf. [13]), the stimuli were controlled for number of syllables (8 words and 13 syllables starting from the embedded clause). In the double-focus conditions, the first object's final syllable was the one bearing lexical stress. The second object always had four syllables with lexical stress on the initial syllable in the clash condition and lexical stress on the second syllable in the no-clash condition. In order to avoid a segmental influence on the tonal marking, which would be especially expected for stops, there were only continuants in the coda of the first object and no voiceless stops in the onset of the second object. Additionally, all "contrasting pairs" (first objects and second objects) were controlled for word form frequency which were taken from the Leipzig Wortschatz corpus [14].

(1) **Context**: Hat Melli gesagt, dass Tobi das Schlagzeug Schülerinnen gegeben hat?
Did Melli say that Tobi has given the drums to pupils?

clash: Nein, sie hat gesagt, dass Tobi das [KlaVIER]$_{OBJ1}$[LEHrerinnen]$_{OBJ2}$ gegeben hat.
No, he said that Tobi has given the piano to teachers.

no clash: Nein, sie hat gesagt, dass Tobi das [KlaVIER]$_{OBJ1}$[StuDENtinnen]$_{OBJ2}$ gegeben hat.
No, he said that Tobi has given the piano to students.

Two control conditions were added to each sentence type, in order to see the tonal realisations of single-focus constructions. The control conditions also allow us to get an idea whether the participants generally processed the context question and understood the task. In condition F1 only the first object was corrected while in condition F2 only the second object was corrected (see examples in (2)). These conditions were also included in the frequency analysis and matched in number of syllables and segmental make-up according to the description above.

(2) **Context:** Hat Melli gesagt, dass Tobi das Schlagzeug Schülerinnen gegeben hat?
Did M. say that Tobi has given the drums to pupils?

F1: Nein, sie hat gesagt, dass Tobi das [KlaVIER]$_{OBJ1}$ [Schülerinnen]$_{OBJ2}$ gegeben hat.
No, he said that Tobi has given the piano to pupils.

F2: Nein, sie hat gesagt, dass Tobi das [Schlagzeug]$_{OBJ1}$[LEHrerinnen]$_{OBJ2}$ gegeben hat.
No, he said that Tobi has given the drums to teachers.

2.3. Procedure

Twenty sentences per condition (clash, no clash, F1, F2) were distributed over 4 lists using a Latin Square Design. The experimental sentences in each list were pseudo-randomised for each participant so that the first 3 mini-dialogues were fillers and that sentences of the same condition were not successive. One list contained 20 experimental sentences and 40 filler sentences. The context questions of each question-answer pair had been previously recorded spoken by a female speaker who was instructed to read the questions neutral as well as natural. The recordings took place in a sound attenuated chamber. The mini-dialogues (both question and answer) were presented on a screen, preceded by instructions and a context story that was designed to make the question-answer pairs more plausible. The instructions were also given verbally. The participants were asked to click on a symbol which triggered playing of the question on loudspeakers. Participants were instructed to first silently read the dialogue, listen to the context question and then produce the answer. They controlled the appearance of each new dialogue themselves by pressing a key on a keyboard. They were instructed to repeat their productions in case of misreadings.

3. Data processing and analyses

The recordings were automatically segmented into words, syllables and phonemes [15]. The analyses were carried out on the stressed (and potentially accented) syllable of the target words (underlined syllables in examples (1) and (2)).

3.1. PaIntE: Parametric Intonation Modelling

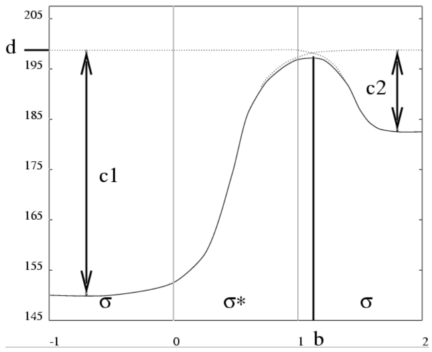
We captured details of the shape of the tonal contour on the syllables of interest using the PaIntE model [16], which approximates a peak in the smoothed F_0 contour by employing a model function operating on a 3-syllable window. There are 6 free parameters in the function term which are set by the model so that the actual F_0 shape is fit best. They are linguistically meaningful: parameter *b* locates the peak within the 3-syllable window, parameter *d* encodes its absolute height. The remaining parameters specify the amplitude and (amplitude normalised) steepness of the rise before, and the fall after the peak (parameters *c1* and *c2* for the amplitude and *a1/a2* for the steepness). Figure 1 illustrates the function.

As mentioned before, we expect potential differences in pitch excursion and absolute peak height. For pitch excursion, the relevant PaIntE parameter varies according to the pitch contour: In rising contours, c1 is expected to be higher than c2, whereas in falling contours, c2 is expected to be higher. Thus, the higher value of either of them was taken. That is, the analysis captures changes in pitch excursion independent of pitch accent type. For absolute peak height we directly employed parameter *d*. Since we are looking at adjacent syllables in *clash* and *F1*, the size of the PaIntE approximation window (3 syllables) is potentially problematic. If a peak is found on any of these syllables, its values will be encoded for each of the syllables: the values capturing the peak's shape are expected to be the same, only the temporal alignment within the 3-syllable window will differ. Therefore, we reduced our dataset in such a way, that only syllables were retained, where the *b* parameter was between 0 and 1.2 ensuring that the peak is not realised further into the next syllable than 20% of that syllable's duration. This reduced the data set considerably (148 of the 638 test items, i.e. 23% were removed). However, the reduction ensures that the *c* and *d*-values employed match the syllable in question.

3.2. Statistical analyses

For each parameter under investigation, we performed a linear mixed effects analysis to investigate the relationship between that parameter and the rhythmic and semantic context. In all analyses, we tested the following factors as fixed effects: the experimental condition (*clash*, *no clash*, *F1, F2*), the type of the object (first object vs. second object), the interaction of these two factors, and the trial number (to control for possible learning effects). As random factors, we tested intercepts

Figure 1: *The PaIntE model function operating on a 3-syllable window with the syllable for which the parametrisation is currently carried out ($\sigma*$) in the context of its immediate neighbours. The x-axis indicates time (normalised for syllable duration, i.e. the current syllable spans from 0 to 1) and the y-axis displays the fundamental frequency in Hertz. Parameters a1 and a2 are not displayed.*



for subjects, word and item and by-subject random slopes for conditions. For each of the analyses, we determined the best fitting linear mixed model by carrying out model comparisons using likelihood ratio tests (cf. [17, 18, 19]). All factors were tested for their significance by comparing the model including the effect in question to the model without it. Only significant factors ($p < 0.05$) were retained in the final models. When a significant predictor in the model has multiple levels, the variable levels are only compared to the intercept, not amongst each other, since only the differences to the intercept are encoded. Therefore we re-levelled the respective variable, so that the reference level (i.e. the intercept) was changed and all potential significances could be detected. We determined the effect of the different variable levels by means of the t-values. We assume that t-values of $> |2|$ indicate significance. For the statistical analyses we used R 3.1.2 [20] with the package lme4 [21].

4. Results

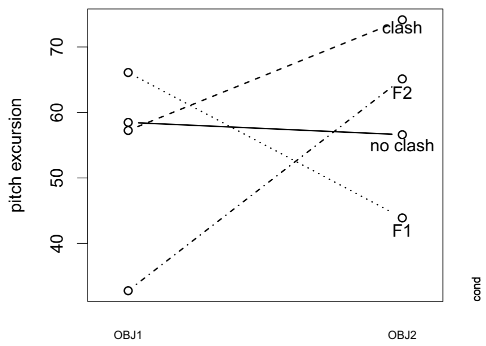
The final models for both parameters under investigation had the same structures in terms of fixed and random effects: they included the interaction of *condition* (with the levels *clash, no clash, F1* and *F2*) and *object* (with the levels *OBJ1* and *OBJ2*) as fixed effects. They had random intercepts for subject and word. The by-subject random slope for *condition* did not significantly improve the models, neither did the trial number. We can therefore assume that there was no habituation effect throughout the experiment.[1]

4.1. Pitch excursion

Figure 2 plots the effect of condition on pitch excursion (y-axis) as obtained from the linear mixed model (using the *languageR* package [22]). The relevant syllable of the first object (OBJ1) and the relevant syllable of the second object (OBJ2) are on the x-axis. Condition *no clash* on *OBJ1*, which is the model's intercept (i.e. the reference condition), is about 58 Hz. The difference between the two objects is encoded in the trajectory of the lines in the plot. The two conditions with one focus (*F1* and *F2*) display a greater pitch excursion on the respective focused object. *F1* (dotted line) has a greater excursion on the first object (which is focused), compared to the second (non-focused) object (β= -20.264, SE=7.775 t=-2.606). That is, pitch excursion in object 2 is about 20 Hz smaller than in object 1 where it is about 66 Hz. *F2* (dot-dashed line) has a greater excursion on object 2 (β = 34.221, SE = 8.020, t = 4.267) compared to object 1, i.e. pitch excursion in object 2 is about 34 Hz greater than in object 1 where it is about 32 Hz. In the double-focus constructions, i.e. conditions *no clash* and *clash*, only for the clash condition the pitch excursion differs significantly between OBJ1 and OBJ2 (β = 18.775, SE = 7.492, t = 2.506). That is, in the presence of a potential clash, the pitch excursion is about 19 Hz greater on OBJ2 than on OBJ1. For *no clash*, the objects are realised with a similar excursion (β = -1.874, SE = 5.160, t = -0.363).

[1]Following a reviewer's comment, we included gender as fixed factor in a subsequent analysis in order to additionally control for gender differences. This did not alter the effects.

Figure 2: *The effect of* condition *and* object *on pitch excursion. Pitch excursion (y-axis) is given in Hz.*



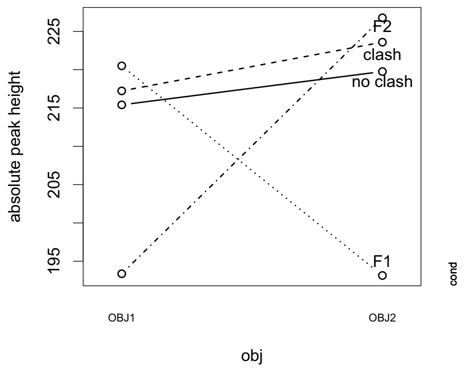
For the first object (OBJ1, circles on the left side of the plot), the results show that all conditions having a focus differ significantly from the unfocused condition (F2): *no clash* by about 26 Hz (SE = 5.976, t = 4.301), *clash* by about 24 Hz (SE = 6.218, t = 3.935), and *F1* by about 33 Hz (SE = 6.036, t = 5.520), while they do not differ from each other. That is, peak excursion on OBJ1 reflects the difference between unfocused first objects (F2) and focused ones (F1, clash, no clash).

For the second object (OBJ2, circles on the right side of the plot), the unfocused condition F1 differs from all other conditions significantly in that pitch excursion is greater in all the other cases: by about 13 Hz (SE = 5.923, t = 2.14) for *no clash*, by about 21 Hz (SE = 6.042, t = 3.514) for *F2*, and by about 30 Hz (SE = 6.053, t = 4.998) for *clash*. Interestingly, among the focused conditions on OBJ2, only the *clash* and the *no clash* condition differ significantly from each other with respect to their pitch excursion. In the *clash* condition, pitch excursion is about 18 Hz higher than in the *no clash* condition (SE = 5.352, t = 3.278). That is, peak excursion on OBJ2 reflects the difference between the two conditions differing only in their rhythmic make-up: If the two potential focus markings are on syllables directly following each other (condition *clash*), pitch excursion is greater than when there is an intervening syllable.

4.2. Absolute peak height

Figure 3 illustrates the effect that *object* and *condition* have in the model predicting absolute peak height of the tonal contour realised on relevant syllable. The difference between the two

Figure 3: *The effect of* condition *and* object *on peak height. Absolute peak height (y-axis) is given in Hertz.*



objects is encoded in the trajectory of the lines in the plot. The two single-focus conditions, F1 and F2, realise OBJ1 significantly different from OBJ2: compared to the focused OBJ1 (where the peak reaches about 220 Hz) the peak on the unfocused OBJ2 is realised about 31 Hz lower in condition *F1* (SE = 6.754, t = -4.694). In cases where OBJ2 is focused (*F2)*, the peak is realised about 29 Hz higher than in the unfocused OBJ1 (SE = 7.876, t = 3.690) where it reaches about 193 Hz. Both double-focus conditions realise OBJ1 and OBJ2 with similar peak height.

For OBJ1 (circles on the left side of the plot), the results show that absolute peak height only significantly differs between the unfocused condition *F2* and the focused ones (*F1, clash, no clash*): the peak in condition *F2* is at about 193 Hz and thereby about 22 Hz lower than in the *no clash* condition (SE = 5.779, t = 3.814), about 24 Hz lower than in *clash* (SE = 5.926, t = 4.027) and about 27 Hz lower than in *F1* (SE = 5.811, t = 4.668). The two double-focus conditions *clash* and *no clash* do not differ from each other (t = 0.481).

For OBJ2 (circles on the plot's right side), the results show that, as in OBJ1, only the condition in which OBJ2 is unfocused (*F1*), differs from the ones, where OBJ2 is focused. The peak on OBJ2 is at about 193 Hz and thereby about 27 Hz lower than in *no clash* (SE = 5.704, t = 4.667) and than in *clash* (SE = 5.721, t = 4.782), and about 34 Hz lower than in *F2* (SE = 5.502, t = 6.113). The double-focus conditions *clash* and *no clash* do not differ significantly from each other (t = 0.717).

To sum up, while absolute peak height is generally influenced by the focus structure, the differences in rhythm have no effect.

5. Discussion and Conclusion

We presented stimuli elicited from a production study where we compared different focus constructions. The parameters investigated were pitch excursion and absolute peak height, derived from a parametric intonation model which approximates potential peaks in the contour. The results have to be interpreted carefully for two reasons. Firstly, the analysis is purely acoustic, i.e. perception is not taken into account and we cannot automatically infer from acoustics to perception. It has been claimed, however, that a Hertz scale reflects listeners' prominence judgements the best [10]. Secondly, the analysis is based on purely tonal features, i.e. other parameters known to influence prominence and, most importantly, pitch accenting, such as intensity and duration are not taken into account. Nevertheless, this study presents an insight of how our data is realised and how double-focus constructions differ, when their rhythmic pattern is manipulated. The results also clearly show, that focused and unfocused expressions differ in the analysed parameters. Focus is reflected by a greater pitch excursion, i.e. greater amplitude of the rising or falling contour on the syllable (compared to its immediate context), and by a higher peak height, possibly reflecting a pitch accent. That is, the acoustic parameters employed are suitable to investigate the tonal realisation of focus.

The two conditions that were designed to test rhythm as a prosody-inherent factor potentially influencing the tonal realisation of focus, differed in pitch excursion. If accentuation of both foci in double-focus constructions would lead to a clash, i.e. if the second focus semantically requires prosodic focus marking on a syllable directly following the stressed syllable of the first focus, pitch excursion was higher on the second object. That is, in cases of a rhythmic clash, speakers alter the tonal realisation of the focus. Presumably, when the two required prominences to mark focus are directly adjacent, speakers add prominence-lending cues to the second one, in order to mark it distinctly. The difference between the pitch excursions on the first and second object in the *clash* condition (about 19 Hz) is remarkable: it is about the same difference as between the two objects in the single focus condition *F1* (about 20 Hz). It is highly likely that *F1* was realised with a pitch accent on *OBJ1* (induced by the focus), which results in the difference to the unfocused (and therefore unaccented) *OBJ2*. The question is how the same difference is perceived when it is realised on two foci, in the *clash* condition. First results from a prominence-judgement experiment using the recordings of this study show that listeners judge the first object in the *clash* condition sig. more often as not prominent compared to the *no-clash* condition. This suggests that the increased excursion on the second object in the *clash* condition perceptually removes prominence from the preceding object.

In the *no clash* condition, speakers produce the two objects with similar pitch excursions. The intervening unaccented syllable seems to be enough to realise two distinct contours. In the above mentioned prominence-judgement task, listeners perceived prominence on both objects more often than in the other conditions. Further perception experiments are planned to investigate whether rhythm will affect the acceptability of the productions in the double-focus constructions.

The results for absolute peak height clearly show that it is used to mark focus tonally, as has been shown before (e.g. [23]). This parameter is employed in single as well as double-focus conditions. However, it does not seem to be influenced by the rhythmic environment. That is, the peak height of the contour on the two objects is the same. Future work will investigate whether speakers preferably produce both foci in one or in two phrases by analysing intonation boundaries and the length of potential pauses. Additionally, we will investigate the timing of the peak and how it is affected by rhythmic manipulations. Preliminary results suggest that peaks are realised later in the *clash* condition than in the *no clash* condition which corresponds to the current findings, since delayed peaks are also prominence lending parameters [24].

To conclude, our study demonstrates that the rhythmic environment needs to be taken into account in mapping semantic categories to phonological categories and their phonetic implementation.

6. References

[1] S. Baumann, M. Grice, and S. Steindamm, "Prosodic marking of focus domains - categorical or gradient?" in *Proceedings of the third International Confrerence on Speech Prosody*, Dresden, Germany, 2006, pp. 301–304.

[2] S. Baumann and M. Grice, "The intonation of accessibility," *Journal of Pragmatics*, vol. 38, no. 10, pp. 1636 – 1657, 2006, prosody and Pragmatics. [Online]. Available: http://www.sciencedirect.com/science/article/B6VCW-4JMVHR2-3/2/4c31f695b09fb445e16f753bf47384e8

[3] K. Schweitzer, A. Riester, M. Walsh, and G. Dogil, "Pitch Accents and Information Status in a German Radio News Corpus," in *Proceedings of Interspeech 2009*, Brighton, UK, 2009, pp. 877–880.

[4] A. Riester and J. Piontek, "Anarchy in the NP. when new nouns get deaccented and given nouns don't," *Lingua*, to appear.

[5] S. Tilsen, "Utterance preparation and stress clash: planning prosodic alternations," in *Speech production and perception: planning and dynamics*, S. Fuchs, P. Perrier, M. Weirich, and D. Pape, Eds., 2012.

[6] G. Kentner, "Linguistic rhythm guides parsing decisions in written sentence comprehension," *Cognition*, 2012.

[7] K. Bohn, R. Wiese, and U. Dohmas, "The status of the rhythm rule within and across word boundaries," in *Proceedings of the 17th International Conference of the phonetic sciences 2001, Hong Kong*, 2011.

[8] S. Baumann, J. Becker, M. Grice, and D. Mücke, "Tonal and Articulatory Marking of Focus in German," in *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, 2007.

[9] C. Gussenhoven, *The Phonology of Tone and Intonation*. Cambridge University Press, 2004.

[10] A. C. M. Rietveld and C. Gussenhoven, "On the relation between pitch excursion size and prominence," *Journal of Phonetics*, vol. 13, no. 3-4, pp. 299–308, 1985.

[11] M. Grice, S. Baumann, and R. Benzmüller, "German intonation in autosegmental metrical phonology," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun, Ed. Oxford: OUP, 2005, pp. 55–83.

[12] F. Kugler and A. Gollrad, "Production and perception of contrast in german: Evidence for left-headed pitch accents," 2014, submitted.

[13] D. Ladd, *Intonational Phonology*, 2nd ed. Cambridge, UK: Cambridge University Press, 2008.

[14] "Leipzig wortschatz corpus," http://http://wortschatz.uni-leipzig.de/, accessed: 2015-02-15.

[15] S. Rapp, "Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov models—An aligner for German," in *Proc. of ELSNET Goes East and IMACS Workshop "Integration of Language and Speech in Academia and Industry" (Russia)*, 1995.

[16] G. Möhler and A. Conkie, "Parametric modeling of intonation using vector quantization," in *Proceedings of the Third International Workshop on Speech Synthesis (Jenolan Caves, Australia)*, 1998, pp. 311–316.

[17] H. Baayen, *Analyzing Linguistic Data*. Cambridge University Press, 2008.

[18] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," *Journal of Memory and Language*, vol. 59, no. 4, pp. 390–412, 2008.

[19] B. Winter, "Linear models and linear mixed effects models in r with linguistic applications." arXiv: 1308.5499, 2013. [Online]. Available: http://arxiv.org/pdf/1308.5499.pdf

[20] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014. [Online]. Available: http://www.R-project.org/

[21] D. Bates, M. Maechler, and B. Bolker, *lme4: Linear mixed-effects models using S4 classes*, 2013, r package version 0.999999-2. [Online]. Available: http://CRAN.R-project.org/package=lme4

[22] R. H. Baayen, *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"*., 2011, r package version 1.4. [Online]. Available: http://CRAN.R-project.org/package=languageR

[23] C. Fry and F. Kgler, "Pitch accent scaling on given, new and focused constituents in german," *Journal of Phonetics*, vol. 36, no. 4, pp. 680 – 703, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0095447008000272

[24] C. Gussenhoven, "Intonation and interpretation: Phonetics and phonology," in *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 2002.