Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

INSPECTing read speech – How different typefaces affect speech prosody

Stephanie Berger [a], Carina Marquard [a], Oliver Niebuhr [b]

[a] Dept. of General Linguistics, Kiel University, Germany
[b] Innovation Research Cluster Alsion (IRCA), University of Southern Denmark, Denmark
Stephanie.berger@freenet.de; carina.marquard@web.de; olni@sdu.dk

Abstract

***In**novating **Sp**eech **E**li**c**itation **T**echniques* (INSPECT) is a line of research that aims to describe and quantify how different recording methods and materials affect speech production both inside and outside the laboratory. In addition, it aims to find new elicitation techniques or refine established techniques such that they provide a better control over speaking styles and/or target variables, particularly with respect to informal and expressive speech. Against this background, the present study investigates if and how typefaces of texts affect prosodic patterns in read speech. Analysing recordings of 24 Standard German native speakers showed that typeface has clear effects on the realization of prosodic patterns in read speech tasks. This includes local events like disfluencies, laughter, and breathing, as well as holistic characteristics of prosodic phrases like mean f0 and intensity levels, f0 ranges, f0 slopes, and speaking rates. For example, texts using sans-serif typefaces like Arial were produced very fluently. The opposite was true for typefaces like Forte and Times, which caused more expressive speech, laughter, and L* and H% intonations.

Index Terms: typography, typeface, prosody, laughter, disfluency, breathing, pitch accents, boundary tones, read speech.

1. Introduction

1.1. INSPECT

Phenomena of everyday conversation such as breathing, laughter, emotion, Lombard speech, entrainment, reduction and emphatic accentuation become increasingly important in phonetic research, and so do corpora of monologues and dialogues produced and recorded in real-life situations [1]. However, empirical phonetic research also requires a large number of tokens, control over experimental and confounding variables, and high-quality recordings of acoustic and/or articulatory signals. It is for these reasons that real-life recordings of spontaneous speech provide, in the majority of cases, only valuable insights when they are analysed *in conjunction with* lab-speech corpora, cf. [2]. This is especially true for basic studies on new phonological (segmental or prosodic) questions, as well as for studies examining phonetic detail.

In fact, very little is known about lab-speech recordings, particularly in terms of how social, environmental, and task-related factors shape speech production inside the lab. Knowing more about these factors would help us record and analyse comparable samples across studies, and ultimately contributes to gain control over lab speech and bridge the gap between lab speech and speech in real-life situations. One of the most important questions is how certain speaking styles (especially expressive/informal speech), speaking rates, intonation patterns, or degrees of reduction can be elicited in the lab without stating explicit instructions to the speakers [3].

INSPECT is a research initiative of the Innovation Research Cluster Alsion at the University of Southern Denmark. With a focus on lab speech, INSPECT implements the call of [4] for “a stronger methodological awareness in investigations of speech phenomena” (p.1). For example, earlier INSPECT studies examined the influence of time of day, dialogue partner and elicitation task (e.g., long reading lists and written dialogue templates) on speech production, see [5] for a summary. The present study continues this line of research in dealing with the typefaces used in read speech tasks. The subsequent section will briefly summarise previous research in this field.

1.2. Background and aims

So far, only a few studies have dealt with behavioural effects of typefaces. By far the most studies on behavioural effects of typefaces have their background in marketing and consumer psychology. These studies show that readers instinctively (and quite consistently across studies, languages and cultures) associate typefaces with different attributes, sometimes termed the semantic qualities [6] or connotative dimensions [7] of typefaces. For example, typefaces can be perceived as being juvenile and friendly [8], bookish, traditional, serious, and modern [9], cool and restrained [10], cuddly and assertive [11], male or female [12, 13], and graceful, confident, dramatic, sophisticated, urban, or theatrical [14]. With respect to the systematic and strong links between typefaces and attributes like those above, a lot of recent studies compare typefaces to humans and speak of personas or personality traits of typefaces [15,16,17]. The personas of typefaces are not only connected to the presence or absence of serifs. Although they play an important role, other "anatomical" and aesthetic aspects are also relevant, such as ascender and descender proportions, stroke form, character spacing, harmony, geometry, naturalness, elaborateness, etc., see [16] for a comprehensive analysis.

Based on listener ratings of personality traits and subsequent correlation and factor analyses of these ratings, [15] and [16] found that typefaces form groups for similar personas. On the form part, these groups consist of certain anatomical and aesthetic (i.e. visual) properties, including serifs. Most studies posit three to four of such typeface categories.

Compared to our growing understanding of the personality associations triggered by individual typefaces or their respective typeface categories, we know almost nothing about possible effects of these personality associations on the *phonetic* realization of read texts. However, the literature suggests that there are effects linked to personality associations, especially concerning prosodic parameters. For example, [18] says that "letter-forms have tone, timbre, character, just as words and

 doi: 10.21437/SpeechProsody.2016-105

sentences do" (p.22); [19] adds that a typeface "conveys mood, communicates attitude, and sets tone" (p.60).

The aim of the present study is to fill these impressionistic statements with phonetic substance. We conduct the first detailed and experimental acoustic investigation that goes beyond the cognitive and perceptual aspects of typeface effects. Special emphasis will be on three prosodic domains that are to be controlled in lab speech and represent key carriers of non-lexical communicative meanings and functions: fluency, intonation pattern, and expressiveness. Each domain will be analysed in terms of several parameters.

In order to enhance the generalizability of our results, we used two texts that were both information-oriented but differed greatly at the stylistic level: one of the texts was serious and presents facts, while the other was informal and funny. The prosodic effect of text type and interactions between text type and typeface type will be examined in a later study.

The texts were presented to readers in four different typeface types that – following the perceptual classifications of [15] – can be sorted into four widely different typeface categories: "all-purpose/direct", "happy/creative", "traditional/business", and "assertive bold/display". Moreover, the latter two typefaces have serifs, whereas the other two do not. In addition, two typeface types belong to the group of so-called plain typefaces, whereas the other two do not.

The main question addressed in this study is to what extent (if at all) the typeface of the text that is presented to participants in an experimental reading task influences measures of fluency, intonation pattern and expressiveness. If so,

1. Are texts with sans-serif typefaces generally produced more fluently? Does the "assertive bold/ display" typeface interfere even stronger with fluency, as it is not a plain and popular typeface? Does the "happy/creative" sans-serif typeface cause stronger interference with fluency than other, more popular sans-serif typefaces?
2. Do "happy/creative" and "assertive bold/display" typefaces yield a higher degree of expressiveness and a corresponding change in the make-up of the intonation pattern? Is the "happy/creative" typeface most effective and the "traditional/business" typeface least effective in this respect?
3. Following from (1) and (2), does the "all-purpose/ direct" typeface yield intermediate (non-extreme) results in terms of all three prosodic domains, fluency, intonation pattern, and expressiveness?

2. Method

2.1. Speakers

Twenty-four native speakers of Standard German took part in the production experiment, 13 females and 11 males. They were 25-55 years old and regular readers of digital or print media. None of them reported to have dyslexia. Although it was not an initial requirement, it turned out that most participants had a university education.

2.2. Reading material

We used four different typefaces for our study: Arial, Forte, Snap ITC (Snap), and Times New Roman (Times). Each of them represented a class of typefaces that is associated with specific perceptual and visual attributes [15]. Table 1 shows these attributes and illustrates the corresponding typefaces.

Table 1: *Overview of the investigated typefaces and their corresponding perceptual and visual features.*

Typeface	Perceived attributes	Serifs/Plain
Arial	"all purpose/direct"	no/yes
Forte	"happy/creative"	no/no
Snap ITC	"assertive bold/display"	yes/no
Times New Roman	"traditional/business"	yes/yes

Each typeface was combined with two information-oriented but stylistically very different texts. One text, "Wombats", is a section of a German Wikipedia article on the life and habits of Australian wombats. Being written in a concise, matter-of-fact fashion, it represents facts in the form of a typical encyclopaedia entry. The other text, "Bagger", comes from the famous German satirical magazine "Der Postillion". The "Bagger" text does not present facts but describes an invented plan of the European Union to substitute environmentally harmful disposable plastic bags by young unemployed university graduates, who can be rented for a small fee to carry the shoppers' purchases home. "Bagger" was written in the form of a news article and is hence syntactically similar to the encyclopaedic "Wombats" text, but a role model for humorous writing that is clearly recognized as such from the very beginning.

With 111 and 110 words respectively, both "Wombats" and "Bagger" have about the same length. Moreover, it was inferred from the syntactic structure of the sentences that both texts would be realized as a sequence of 17 prosodic phrases.

All typefaces were set in 14pt with 1.5 line spacing and justified print. Before printing out the texts on separate sheets of paper, we took care that the line breaks in each text always occurred after the same word, independently of typeface.

The eight reading conditions – four typeface types embedded in two different texts – were supplemented by a dummy condition: "The North Wind and the Sun" fable set in Courier New 14pt. It was read twice at the beginning of the production experiment as a warm-up and in order to familiarize speakers with the recording situation.

2.3. Analyses

Our analyses were prepared and conducted on the basis of PRAAT and PRAAT TextGrids [20]. We included 12 dependent variables, selected to cover the three prosodic dimensions of fluency, intonation pattern, and expressiveness.

Fluency was measured in terms of the number of annotated disfluencies (hesitations, repetitions), laughs, and breathing pauses produced in each typeface condition across all speakers. Annotations and frequency counts were made on an auditory basis by MA students of phonetics at Kiel University.

The intonation-pattern analysis was also based on annotations and frequency counts, conducted by the same trained listeners. They used the DIMA system for German intonation annotation [21]. In order to make the annotation task simple and reliable, the trained listeners only had to distinguish L* and H* pitch accents and L% and H% boundary tones. The frequencies of L*, H*, L%, and H% across all speakers in each typeface condition were determined and used as the dependent variables of intonation.

The prosodic dimension of expressiveness was analysed in the form of acoustic parameters. The analysis was conducted automatically by means of the ProsodyPro script [22] for PRAAT. The script was applied to the TextGrid files of the intonation analysis. Thus, all acoustic measurements were made at the level of prosodic phrases. The following measurements were taken: excursion size (i.e. f0 range in st), mean f0 (in Hz), mean intensity (RMS in dB), phrase duration (in s), and maximum velocity of f0 change (i.e. vmax f0 in st/s). Intensity measurements were included as the speakers' distances to the microphone were kept approximately constant. Vmax f0 was determined in the form of absolute values, hence pooling rises and falls, and phrase duration was converted into the established speaking rate measure syllables-per-second (syl/s).

2.4. Experimental procedure

The experiment was conducted in individual sessions. In order to make the speakers feel more comfortable, the recordings took place in a silent room at the speakers' own homes. For the same reason, the readings were not recorded with special recording devices. Rather, the participants spoke into smartphones. Recordings were made in WAV format at a sampling rate of 44.1 kHz and a 16-bit quantization.

Each recording session started with written instructions. The speakers were informed that they would get, on separate sheets of paper, three short texts that they had to read aloud in the given order while speaking into their smartphone, pretending that a friend of theirs would be listening at the other end of the phone. Accordingly, the speakers were asked to make themselves comfortable and read each text in a listener-oriented, committed fashion with an appropriate tone of voice. In case of mistakes, they had to start over again with the corresponding sentence. Speakers were given the opportunity to familiarize themselves with each text prior to being recorded.

Note that our speakers held the smartphone in a comfortable constant position during the production of each text. Thus, the smartphone offered a similar advantage as a head-set (for RMS measurements), and, unlike the latter, additionally placed our speakers in a highly familiar everyday situation. Global RMS differences between text readings that resulted from differences in the way the phone is held (moved back to the ear) by our speakers were normalized prior to analysis.

Table 2: *Order in which familiarization and target texts were read in each of the 4 speaker groups. Half of the speakers in a group additionally read the two target texts in reverse order.*

	Speaker group 1	Speaker group 2	Speaker group 3	Speaker group 4
3 speakers	North Wind	North Wind	North Wind	North Wind
	Wombats (Arial)	Wombats (Forte)	Wombats (Snap)	Wombats (Times)
	Bagger (Arial)	Bagger (Forte)	Bagger (Snap)	Bagger (Times)
3 speakers	North Wind	North Wind	North Wind	North Wind
	Bagger (Arial)	Bagger (Forte)	Bagger (Snap)	Bagger (Times)
	Wombats (Arial)	Wombats (Forte)	Wombats (Snap)	Wombats (Times)

Our four typefaces were distributed across separate speaker groups to which our 24 participants were randomly assigned, see Table 2. By letting each speaker perform the target texts with only one typeface, we effectively hid the actual aim of the reading task from the participants. Within each speaker group, we balanced the reading order of the "Wombats" and "Bagger" texts by swapping this order for half (3) of the speakers in a group. Independent of this balanced reading order, the two target reading conditions in each speaker group were always preceded by two warm-up readings of "The North Wind and the Sun" fable.

3. Results

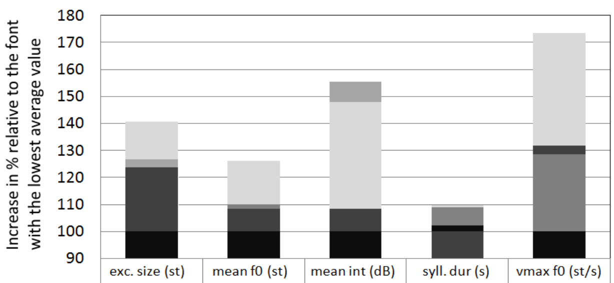
Chi-squared tests and a one-way MANOVA (fixed factor Typeface) were conducted to test for effects of typeface on the frequency counts and acoustic measurements. We found clear effects of typeface on all three prosodic dimensions, i.e. fluency, intonation pattern, and expressiveness.

Results on fluency show a division between the serif typefaces Times and Snap on the one hand and the sans-serif typefaces Arial and Forte on the other, see Table 3. Compared with the sans-serif typefaces, the serif typefaces Times and Snap caused more disfluencies and laughs. They also caused a few more breathing pauses. Overall, these differences between serif and sans-serif typefaces were statistically significant in the corresponding Chi-squared test ($\chi^2[6]=12.9$, $p<0.05$).

Table 3: *Total frequencies of the fluency (top) and intonation (bottom) features produced across all 24 speakers in the four Typeface conditions Arial, Forte, Snap, and Times.*

Total frequencies	Disfluencies	Laughter	Breathing pauses
Arial	13	2	105
Forte	16	7	111
Snap	28	10	125
Times	20	9	122

Total frequencies	H*	L*	H%	L%
Arial	851	500	183	631
Forte	659	784	390	420
Snap	747	535	275	539
Times	621	723	365	447



	exc. size (st)	mean f0 (st)	mean int (dB)	syll. dur (s)	vmax f0 (st/s)
Times	140,6	126,5	148,8	110,8	172,6
Forte	127,2	110,1	155,4	107,4	128,9
Snap	124	108,3	108,5	100	134,7
Arial	100	100	100	102,3	100

Figure 1: *Mean differences between the acoustic parameters in the Typeface conditions Arial, Forte, Snap, and Times. The differences are shown in % relative to the typeface with the lowest average value (circled); each parameter 810 ≤ n ≤ 814.*

Regarding intonation pattern, the typefaces also form pairs in which the results patterns are similar. However, now the dividing line runs *across* the serif and sans-serif typefaces: Arial forms a pair with Snap; and Forte is in a pair together with Times. Compared with the typeface pair Arial/Snap, the typeface pair Forte/Times increased the frequencies of L* pitch accents and H% boundary tones. In contrast, Arial and Snap led to more H* and L% renderings. These diametrically opposed effects of the two typeface pairs Arial/Snap and Forte/Times are again in an overall significance of the corresponding Chi-squared test ($\chi^2[9]=168.2$, $p<0.001$).

The acoustic exponents of expressiveness were also clearly affected by typeface. As can be seen in Figure 1, there is again a dividing line that, like for intonation pattern, separated Arial/Snap from Forte/Times. Differences concern all five prosodic parameters. Compared to the typeface pair Arial/Snap, the typeface pair Forte/Times slowed down the reader's speaking rate, increased his/her intensity and pitch levels, and triggered larger and faster pitch movements. The biggest differences were between Arial and Times. Texts presented in Arial caused four out of the five acoustic parameters (excursion size, mean f0, mean intensity, vmax f0) to be at the lower end of the range of measured values. In contrast, texts presented in Times elicited on average the highest values for all pitch- and duration-related acoustic parameters.

The results of the MANOVA are consistent with this descriptive analysis. Typeface had a highly significant main effect on all acoustic parameters: excursion size ($F[3,812]=22.9$, $p<0.001$, $\eta_p^2=.08$), mean f0 ($F[3,812]=36.4$, $p<0.001$, $\eta_p^2=.12$), mean intensity ($F[3,812]=66.6$, $p<0.001$, $\eta_p^2=.20$), syllable duration ($F[3,812]=3.8$, $p<0.001$, $\eta_p^2=.05$), and vmax f0 ($F[3,812]=22.4$, $p<0.001$, $\eta_p^2=.08$). Multiple pairwise comparisons (with Bonferroni corrections of p-levels) additionally showed that the main effect of Typeface is primarily due to differences between Arial/Snap and Forte/Times: 70 % of the corresponding t-tests reached significance ($p<0.05$). In contrast, only 30 % of the t-tests within typeface pairs reached significance, particularly those comparing Arial and Snap.

4. Conclusions

The results of our speech-production experiment provide strong empirical evidence that typeface has consistent and clear effects on how speakers realize texts in a read-speech task. This includes all three analysed prosodic domains: fluency, intonation pattern, and expressiveness.

Some of these effects were expected. For example, previous studies concluded from eye-tracking or EEG data that sans-serif typefaces are more legible and readable than serif typefaces [31,32,33]. Our results mirror this conclusion at the phonetic level in that the two sans-serif typefaces Arial and Forte caused less disfluencies (hesitations, repetitions) than the two serif typefaces Snap and Times. Effects of typeface on the frequencies of laughter and breathing pauses were smaller but point in the same direction, i.e. serif typefaces made speakers produce more laughter and insert more breathing pauses into the reading flow than sans-serif typefaces. However, disruptions of the reading flow due to laughter and breathing pauses probably not only reflect legibility and readability differences between serif/sans-serif typefaces. First, it is reasonable to assume that laughter and breathing pauses are confounded variables, i.e. laughing is a breath-consuming process and subsequently entails more breathing pauses. Second, the speaker's expressiveness is likely to be another source for the frequency differences in laughter and breathing pauses. An expressive speaking style is associated with a higher trans-glottal airflow [23,24] and moreover fosters laughter.

In the expressiveness domain lies one unexpected finding of our study: Times obviously induced an expressive reading style. Expressiveness was assumed to manifest itself in a larger pitch excursion size, higher mean f0 and intensity levels, a faster speaking rate, and steeper intonation slopes. Notably, Times reached the highest average values of all typefaces in four of these acoustic-prosodic parameters (excursion size, mean f0, speaking rate, and steepness of intonation slopes), as well as the second-highest average value in a fifth parameter (intensity level). Thus, Times even outperformed Forte in terms of expressiveness parameters, and both Forte and Times triggered more expressive readings than Snap.

It fits in with this conclusion that Forte and Times also made speakers realize more low pitch accents (L*) and phrase-final rising (H%) intonations. Unlike H* and L%, which represent the two elements of the typical neutral matter-of-fact intonation in German [25,26], L* and H% are more listener-oriented, and/or signal friendliness, contrastive emphasis, indignation, and surprise [25,26,27,28,29].

Arial turned out to be the "negative image" of Times in that it yielded the lowest average values in almost all expressiveness parameters, the highest frequencies of H* and L% intonations, and the lowest frequencies of disfluencies, laughter, and breathing. Arial was therefore no "all purpose" typeface in the sense that it yielded intermediate reading performances in all three prosodic domains. Rather, it was an extreme typeface yielding a conservative, "sober" reading style.

In fact, besides the serif-related findings on fluency, we found no evidence in the present study that visual features (e.g., plain vs. not plain) or perceived personality traits of typefaces (see Tab.1) surfaced in the speakers' readings. Times, associated with tradition and business, induced much more expressive readings than, for example, the positive and self-confident typeface Snap. Snap, in turn, caused more laughter than Forte whose visual characteristics were found to embody happiness and creativeness. Thus, our study opens a large field of new questions concerning the relations between perceptual attributes of typefaces and their actual phonetic effects in reading tasks. It seems that personality attributes associated with typefaces do not allow straightforward predictions about how they influence the prosody of read speech. However, the fact that Arial and Times yielded, in almost all respects, the most extreme reading performances in our study could mean that our results are biased insofar as all our speakers certainly know Arial and Times from different aspects of their everyday life, whereas they probably never saw Forte and Snap before. This familiarity difference could have led to more extreme readings in combination with the familiar typefaces, cf. the results of [30].

Based on our results, we can give the following advice for the use of typefaces in read speech: Arial should be used for eliciting the prosodic patterns of a typical reading style. Snap is as good as Arial in triggering H* and L% intonations and moreover very suitable for eliciting disfluencies, laughter, and breathing. Forte fosters laughter without causing a lot of other disfluencies or pauses. Times yields almost as many L* and H% intonations as Forte and is even more suitable than Forte for supporting expressive readings inside the laboratory. Future studies will have to refine this picture by including further typefaces and searching for individual or text-specific differences as well as interactions with – amongst others – typeface size, typeface and paper colour, paper type, and culture.

5. Acknowledgements

We are greatly indebted to all students of the course on "Phonetic databases and speech processing" (winter 2014/15 at Kiel University) for their help in selecting texts and typefaces, recruiting and recording speakers, and analysing the data. Special thanks go to Alumni Kiel e.V. for generously supporting the first two authors.

6. References

[1] N. Campbell and P. Mokhtari, “Voice quality: the 4th prosodic dimension,” *Proceedings of the XVth Int. Congress of Phonetic Sciences, Barcelona, Spain*, vol. 3, pp. 2417–2420, 2003.

[2] Y. Xu, “In defense of lab speech,” *Journal of Phonetics*, vol. 38, pp. 329–336, 2010.

[3] O. Niebuhr, B. Peters, R. Landgraf, and G. Schmidt, “The Kiel Corpora of "Speech & Emotion" – A Summary,” *Proceedings of the 41st Conference of the German Acoustical Society, Nuremberg, Germany*, pp. 1-4, 2015.

[4] P. Wagner, J. Trouvain, and F. Zimmerer, "In defense of stylistic diversity in speech research," *Journal of Phonetics*, vol. 48, pp. 1-12, 2015.

[5] O. Niebuhr and A. Michaud, "Speech data acquisition – The underestimated challenge," *Kieler Arbeiten zur Linguistik und Phonetik (KALIPHO)*, vol. 3, pp. 1-42, 2015.

[6] E. R. Brumberger, "The Rhetoric of typography: The persona of typeface and text," *Technical Communication*, vol. 50, no. 2, pp. 206-223, 2003.

[7] C. L. Rowe, “The connotative dimensions of selected display typefaces,” *Information Design Journal*, vol. 3, pp. 38-50, 1982.

[8] R. Sassoon, “Through the eyes of a child – Perception and type design,” In R. Sassoon (ed.), *Computers and typography*. Oxford, UK: Intellect Books, pp. 178–201, 1993.

[9] Ch. Kostelnick and D. D. Roberts, *Designing visual language: Strategies for professional communicators*. Needham Heights: Allyn and Bacon, 1998.

[10] E. Spikermann and E. M. Ginger, *Stop counting sheep and find out how type works*. Berkeley: Peachpit, 2003.

[11] J. D. Berry, *Now read this: The Microsoft ClearType typeface collection*. Seattle: Microsoft Corporation, 2004.

[12] A. D. Shaikh, D. Fox, and B. S. Chaparro, "The Effect of Typeface on the Perception of Email," *Usability News*, vol. 9.1, 2007.

[13] A.D. Shaikh, "Know Your Typefaces! Semantic Differential Presentation of 40 Onscreen Typefaces," *Usability News*, vol. 11.2, 2009.

[14] R. Shushan and D. Wright, *Desktop publishing by design*. Redmond: Microsoft Press, 1994.

[15] A. D. Shaikh, B. S. Chaparro, and D. Fox, “Perception of typefaces: Perceived personality traits and uses,” *Usability News*, vol. 8.1, 2006.

[16] Y. Li, *Typeface personality traits and their design characteristics,* MA thesis, Montréal, Canada: Concordia University, 2009.

[17] Y. Li and C. Y. Suen, “Typeface personality traits and their design characteristics,” *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, New York, USA*, pp. 231-238, 2010.

[18] R. Bringhurst, 1996. *The elements of typographic style (2nd edition)*, Point Roberts: Hartley and Marks, 1996.

[19] R. C. Parker, *Looking good in print*. Research Triangle Park: Ventana Communications Group Inc., 1997.

[20] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, pp. 341-345, 2001.

[21] F. Kügler, B. Smolibocki, D. Arnold, B. Braun, S. Baumann, M. Grice, S. Jannedy, J. Michalsky, O. Niebuhr, J. Peters, S.B. Ritter, Ch.T. Röhr, A. Schweitzer, K. Schweitzer, P. Wagner, "DIMA – Annotation guidelines for German intonation," *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, UK*, pp. 1-5, 2015.

[22] Y. Xu, "ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis," *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France*, pp. 7-10, 2013.

[23] K. R. Scherer, T. Johnstone, and G. Klasmeyer, “Vocal expression of emotion,” In R. J. Davidson, K. R. Scherer, and H. Goldsmith (eds), *Handbook of the Affective Sciences*, pp.433–456. Oxford: Oxford University Press, 2003.

[24] E.B. Holmberg , R.E. Hillman, and J.S. Perkell, "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *Journal of the Acoustical Society of America*, vol. 84, pp. 511-529, 1988.

[25] M. Grice and S. Baumann, (2002). “Deutsche Intonation und GToBI,” *Linguistische Berichte*, vol. 191, pp. 267-298, 2002.

[26] O. Niebuhr, “The signalling of German rising-falling intonation categories – The interplay of synchronization, shape, and height,” *Phonetica*, vol. 64, pp. 174-193, 2007.

[27] K. J. Kohler, “Prosody in speech synthesis: the interplay between basic research and TTS application,” *Journal of Phonetics*, vol. 19, pp. 121–138, 1991.

[28] K. J. Kohler, “Timing and communicative functions of pitch contours,” *Phonetica*, vol. 62, pp. 88–105, 2005.

[29] J. Peters, *Intonation – Kurze Einführung in die germanistische Linguistik*. Heidelberg: Winter. 2014.

[30] S. Juni and J. S. Gross, “Emotional and persuasive perception of typefaces,” *Perceptual and Motor Skills*, vol. 106, pp. 35-42, 2008.

[31] Robinson, D.O, M. Abbamonte, and S.H. Evans, "Why serifs are important: the perception of small print," *Visible Language*, vol. 4, pp. 353-359, 1971.

[32] Bernard, M.L., B.S. Chaparro, M.M. Mills, and C.G. Halcomb, "Comparing the Effects of Text Size and Format on the Readability of Computer-Displayed Times New Roman and Arial Text," *International Journal of Human-Computer Studies*, vol. 59, pp. 823-835, 2003.

[33] Beymer, D., D. Russell, and P. Orton, "An eye tracking study of how font size and type influence online reading," *Proceedings of People and computers XXII : culture, creativity, interaction: 22nd British HCI Group annual conference, Liverpool John Moores University, UK,* pp. 15-18, 2008.