Visual **prosody** supports reading aloud expressively.

Ann Bessemans, Maarten Renckens, Kevin Bormans, Erik Nuyts, Kevin Larson

Abstract

Type is not expressive enough. Even the youngest speakers are able to express a full range of emotions with their voice, while young readers read aloud monotonically as if to convey robotic boredom. We augmented type to convey expression similarly to our voices. Specifically, we wanted to convey in text words that are spoken louder, words that drawn out and spoken longer, and words that are spoken at a higher pitch. We then asked children to read sentences with these new kinds of type to see if children would read these with greater expression. We found that children would ignore the augmentation if they weren't explicitly told about it. But when children were told about the augmentation, they were able to read aloud with greater vocal inflection. This innovation holds great promise for helping both children and adults to read aloud with greater expression and fluency.

1. Introduction

Reading is magical. It allows us to communicate over unlimited time and distance. More immediately, when we read, we need to convert the letters into sounds. Successfully making this mapping is key step in learning to read. Once the alphabetical code can be cracked, a child is able to independently decode words (Stanovich, 1986; Rayner & Pollatsek, 1989). Once individual words can be read aloud, it takes another step to read in a way that sounds natural. Beginning readers often struggle to read aloud in a fluent, expressive manner. Reading fluency is defined not only by speed and accuracy but also by proper expression and the naturalness of reading (NAEP, 1995; National Institute of Child Health & Human Development, 2000). Expressive oral reading can be quantified in terms of prosodic variation in pitch, duration, and volume (Patel & Furr, 2011). These features can be of crucial importance in understanding exactly what the speaker or narrator is trying to tell us.

Expressive reading is an increasingly valued component of literacy. The first focus of reading must be on speed and accuracy of decoding. In Belgium and the Netherlands, the reading levels are expressed by the AVI-levels (abbreviation of 'Analyse van Individualiseringsvormen', translated as 'Analysis of forms of individualization'). This kind of standardization makes it possible to judge the reading level of the child, but these tests are measured on reading speed and orthographical errors only. Techniques that are aimed at improving expressive oral reading should be an integral part of reading fluency for ultimately reading success (Hudson et al., 2005; Rasinski, 1990; Samuels, 1988; Schreiber, 1980). There are several reasons why children's prosodic oral reading fluency is important (Duong et al., 2011; Gussenhoven 2004). Prosodic readers are not only easier to understand, but they also have the ability to improve decoding, word recognition, reading accuracy, reading speed and comprehension skills as they are able to segment text into meaningful units (Dowhower, 1991; Miller & Schwanenflugel, 2006; Ashby, 2006; Binder et al., 2013; Young-Suk Grace, 2015). Better prosody correlates to greater reading achievement.

In the Netherlands reading aloud competitions are a tradition for almost 22 years (Stichting Lezen, 2016). Their most important reasons of being are encouraging children to read and awaken their enthusiasm for literature. Reading aloud competitions' main focus is on the pleasantness of listening to reading aloud. Although there are many criteria which judges look for, an important one is the use of the own voice (without using artificial voices). A good reader is able to make use of small changes in tempo, can use a change of pitch and can read louder and softer to convey a mood or emotion (Stichting Lezen, 2016).

The speech of beginning readers appears flat and laborious when reading aloud (Miller & Schwanenflugel, 2006; NIH, 2000). Earlier approaches to aid children's prosodic reading aloud have focused on repetition and imitation of an adult-repeated-reading model (Read Naturally, 2015) and guided oral reading (Kuhn & Stahl, 2003; Playbooks, 2013; PROJECT LISTEN, 2009; Beck & Mostow, 2008). And while beginning readers typically employ prosody in conversational speech, written text does not provide information about the intended prosody. Nonetheless, text could indicate prosodic variations by means of the typeface. We call this visual prosody.

The most well-known ways in which prosody is visualized in typography is in the punctuation of normal typefaces, and in the phonetic transcriptions in comic books (i.e. BAM!). Comic artists take into account a visual form of prosody to liven up the text. However, the text accompanies the image that is rudimentary. It is through the image that the meaning of the visual text can be determined and understood. Some experimental type projects have explored the use of phonetic qualities. The acclaimed poet Paul van Ostaijen made use of visual poetry in his 'sound poems'. His 'ritmiese typografie' was designed by artist Oscar Jespers. The poem Boem Paukeslag (1921) is probably the most well-known.

There also were more artistically inclined type projects which incorporated facets of spoken language, such as the 'New Alphabet' by Tschichold (1930). These typefaces were developed based on an idealism, dogma or philosophy, in this case during the Bauhaus. Conceptual type projects in which aspects of phonetics form the foundations of the typefaces are seen in the work of Kurt Schwitters 'Systemschrift' (1927) and more direct relations to the language itself in the various projects of the French/Italian type designer Pierre di Sciullo.

Researchers have also thought about introducing visual prosody within text. Both van Uden (1973) and Patel and Furr (2011) treat visual prosody by adding a second layer to the text. Van Uden does this in the form of melody bows. Patel and Furr (2011) used two methods to improve visual prosody: manipulated text cues and augmented text cues. In the manipulated text cues they shift letter placement horizontally to indicate duration and vertically to indicate pitch and used grey level to indicate loudness. In the augmented text cues, they add graphs, lines and vertical bars behind the text to indicate the visual prosody. They found that both methods are effective, but that the manipulated version is harder to read, especially by shifting the words vertically. Both van Uden's and Patel & Furr's forms of visual prosody show additional information on top of the text, which reduces the legibility of the text.

It is also worth looking at the intuitive character of visual prosody. It is not unreasonable to assume that children intuitively, without additional explanation, spontaneously interpret certain adjustments as intended by the designers. Evidence for a common sense or intuitive feeling presented within type design can be found in research (Shaik, 2009; Lewis & Walker, 1989). For example, a bold or black typeface is perceived louder against a lighter or greyer one (Shaik, 2009). The intuitive character also gives us information about het learnability of visual prosody.

The goal of this project is to help children read aloud with more expression. Specifically, we want to show with type the three main components that people already use in spoken language: volume, duration or word length and pitch (Sitaram & Mostow, 2012; Peppé, 2009; Schwanenflugel et al., 2006; Cutler et al., 1997; Dowhower, 1991; Bollinger, 1989; Lehiste, 1970). We want to do this without reducing legibility and in a way that will be easy to learn. To explore whether it is possible to make prosody visible in type to guide children's reading aloud, we formulate four research questions:

A. Will children read text aloud with greater expression with text that is designed to show the components of prosody?

B. Will children read the cues as intended: the volume cue read with greater volume, the pitch cue read at a higher pitch, and the duration cue for a longer amount of time?

C. Will the children intuitively understand the visual prosody or is instruction of the visual cues needed?

D. After using the visual prosody, will the children be able to correctly describe what each of the components of visual prosody mean? What does visual prosody tell us about its learnability?

2. Methodology

Participants

118 children participated in the study. No participants were disqualified. The participants in this study were Flemish children aged eight to ten years old and were enrolled in regular elementary school. All children were reading normally for their age (reading level of at least AVI 5). The tests were conducted at the elementary school 'Sint-Rita', located in Sint-Truiden, Belgium. The children's parents were informed about the research by a formal letter. After the parents' insight into the research, a written approval was asked if their child was allowed to participate in the test. The children were randomly divided into two groups, an information group (61 children) and a no-information group (57 children).

Fonts

The typeface Matilda was selected because its legibility has been extensively studied for use with normal and low vision children (Bessemans, 2012). 8 new versions of Matilda were designed for this study to show the volume, duration, and pitch features of prosody. All conditions were shown at a sufficiently large 18 point size.

Volume

The boldness of letters was modified to indicate that a word should be read with increased volume. Figure 1 shows a word with the normal Matilda font, a half bold font, and a full bold font.

De beer is in de tuin.

De **beer** is in de tuin.

De **beer** is in de tuin.

Figure 1: from top to bottom: 'beer' in the normal variation, 'half bold' and 'full bold'. The Dutch sentence translates to "The bear is in the garden."

Duration

The width of letters was modified to indicate that a word should be read slower, or for a longer amount of time. Figure 2 shows the normal Matilda font, a half wide font, and a full wide font.

De arme man bleef alleen achter.

De arme man bleef alleen achter.

De arme man bleef alleen achter.

Figure 2: from top to bottom: ‘alleen’ in the normal variation, ‘half wide’ and ‘full wide’. The Dutch sentence translates to “The poor man was left alone.”

Pitch

Visually describing pitch was the most challenging aspects of prosody. Two attempts were made in order to test if one would work better than the other. In a first version of pitch, letters were raised above the baseline to show that pitch should be raised. Figure 3 shows the normal Matilda font, a half raised font, and a full raised font. In second version of pitch, letters were stretched vertically to show that pitch should be higher. Figure 4 shows the normal Matilda font, a half stretched font, and a full stretched font.

De poes zit op de grote doos.

De poes zit op de grote doos.

De poes zit op de grote doos.

Figure 3: from top to bottom: ‘op’ in the normal variation, ‘half raised’ and ‘full raised’.

De boer had ook een grijze ezel.

De boer had ook een grijze ezel.

De boer had ook een grijze ezel.

Figure 4: from top to bottom: ‘ezel’ in the normal variation, ‘half stretched’ and ‘full stretched’.

In total 9 fonts (variations on one typeface) were used in the study, namely the normal Matilda (n) and its 8 prosodic type design parameters aimed at influencing volume (‘half bold’, ‘full bold’), duration (‘half wide’, ‘full wide’) and pitch (‘half raised’, ‘full raised’, ‘half stretched’, ‘full stretched’).

Sentences

5 unique sentences were examined in this project, each with a key word that would appear in the studied conditions. The reading level for the sentences were slightly below the reading level for the children. This was done in order to assure that the measurements could be focused solely on the children’s reading aloud and not on reading difficulties of words that otherwise might have occurred. The creation of these sentences were done in collaboration with the teachers of the respective classes. Each of the 5 sentences was repeated 9 times,

once in each of the conditions. This made the pronunciation of each of the conditions directly comparable.

45 sentences in total were presented in A5 size booklets with 5 sentences per page on slightly off-white to yellow paper. Figure 5 shows a sample page in one of the booklets.

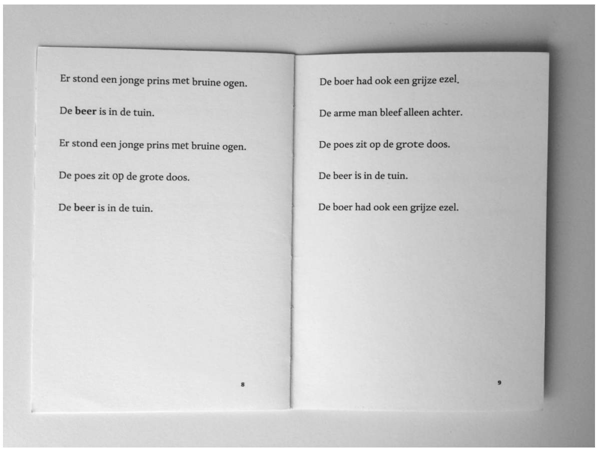


Figure 5: An example of how the sentences are presented to the participants in the booklet.

Procedures

The participants were told that we were investigating ways of making reading easier and more fun. The study was conducted one participant at a time in a quiet, familiar room. The participants were assigned to either the information group or to the no-information group. The no-information group received no introduction to the volume, duration, and pitch conditions they were asked to read, while the information group was shown the different conditions and was given examples of reading the sentences with increased prosody. These instructions were taught in a playful manner, where the child had to effectively look at the testing material and search for prosodic cues. The children discovered the prosodic cues and were taught the envisioned way to read them out aloud.

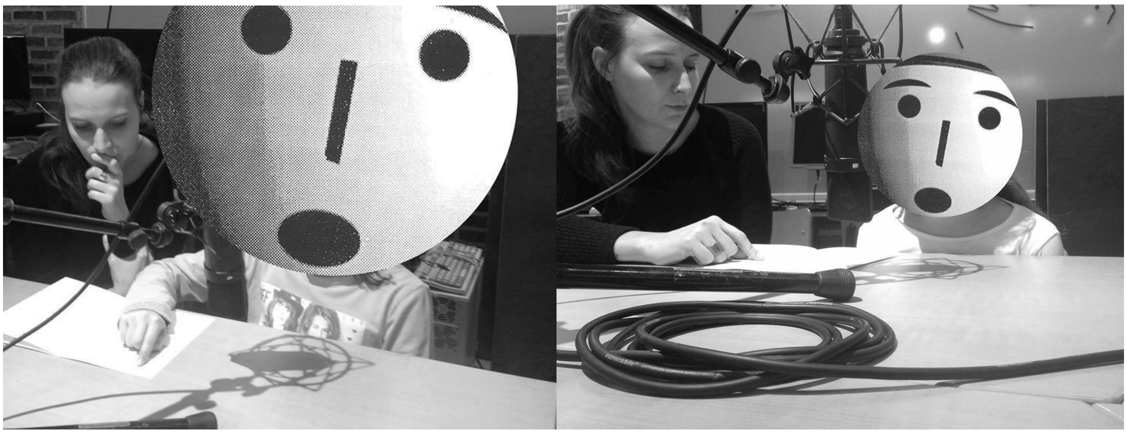


Figure 8A: A child pointing to noticed parameters in the sentence during the talk before or after the test.

Figure 8B: The actual reading test in which the child is reading after getting used to the microphone and the design researcher pointing at the sentence that the child should read.

Each participant was then given a booklet with the 45 sentences presented in a different random order. Their task was to read the sentences aloud the best way they could. Only after a participant showed understanding of the task, the test was started. Some shy children in the information group which were afraid of pronouncing the parameter clearly, were asked to exaggerate a little the pronunciation. If during the reading session, a participant in the no-information group asked about the conditions, he/she was told we were making changes to the text but nothing could be said about it until the end of the experiment. During the test, in order to ensure all sentences were read, the administrator indicated the sentence the child had to read by pointing at it. This was also done to ensure that, during the recordings, there were pauses to indicate clearly the start and ending of every sentence. The participant read all sentences in the book in his/her own pace and was allowed to correct when desired. If deemed necessary, a break in the middle of the book was taken. For the youngest children this break was necessary, as more children than expected lacked the concentration to read 45 sentences consecutively.

At the end of the first day of the experiment, each participant was debriefed. The no-information group received information about the prosodic cues in the same way as the information group got the clues. The purpose of the experiment was explained, and the child was given the chance to ask questions about the study.

One or two days after each participant read the test sentences and was debriefed, the participants were given a questionnaire as a whole class assignment. There were four tasks as part of the questionnaire. The first was to look at prosody marked sentences and identify which words have special prosody. The second task was to write down how they would pronounce the prosody marked words. The third task was to state a preference between the two kinds of pitch conditions. And the fourth task was an open-ended request for feedback.

Measurements

Digital audio recording was done with the Neumann U87ai microphone, designed for voice recording. The digital processing and saving of the audio file was executed via the program Praat, developed at the Department of Phonetic Sciences, University of Amsterdam (Boursma, 2001; Boursma & Weenink, 2016). All data was collected on the most important vowel of the test word. This is in line with Moneta et al. (2008) who focussed only on the vowels to measure voice quality, emotions, in terms of amplitudes and frequencies.

Statistics

With X as the volume, duration or pitch, results are calculated as {average X of one vowel of one specific word} divided by {average X of all the same vowels of the same word of the same child}. E.g. the average pitch of the 'ee' in the word 'beer' written in vt_f, compared with the average of {all the average pitches of all the 'ee' of all the words 'beer' the same child has pronounced}.

The impact of a learning effect was avoided by following procedure. (i) There are 5 sentences and 9 fonts, resulting in 45 sentence-font-combinations. Every booklet has exactly these 45 sentences. (ii) These 45 sentences were randomized. After this randomization, the order was manually adapted (with as minimal changes as possible) in such a way that the same sentence was presented maximally twice immediately after one another. Also the same font was presented maximally twice after one another. (iii) By the combination of this randomization and the limited manually adaptions, sentences nor fonts were too much clustered in the beginning, the middle or the end of the reading task. (iv) This procedure was done 20 times. Hence, there were 20 different booklets with each an unique order of sentence-font-combinations. Every child read one booklet. (v) Since the statistics are performed on these 20 different booklets, that are quite equally spread in the data set, in the overall dataset sentences and fonts are even more equally spread over the beginning, the middle or the end of the reading task. (vi) It is assumed that the learning effect hardly differs between the fonts. By the steps described in the former steps, all fonts were equally spread over the order of the 45 sentences, and thus the learning effect is measured in all fonts in a same way. (vii) By calculating statistics on a ratio: {average of a vowel in a specific font}/{average of this vowel of all fonts} the impact of the learning effect disappears as it is both in nominator and denominator.

The effect of the fonts on the parameters of visual prosody is measured using a Generalized Linear Model with repeated measures in SAS, procedure mixed. This procedure includes adapted Tukey post hoc comparisons that takes the Bonferroni correction into account to test simultaneously the set of all pairwise comparisons $\{\mu i - \mu j\}$. A Generalized Linear Model is an extension of the classical ANOVA, but it has the extra options (repeated measures, Tukey, Bonferroni correction) required for this dataset. For the present paper, only the comparisons of the different fonts with the normal font are used.

3. Results

81 out of 5310 sentences were not processed due to an unknown error from the speech recognition application. The error happened mainly in one sentence, resulting in a highly underrepresentation of this sentence in the sample. The other 44 sentences were equally spread in the sample (average 359 ± standard deviation 12). Not all words recorded

could be used, e.g. if a child stuttered the automatic data recognition program could not recognize the word. In total 14457 words were included in the analyses.

The no-information group

The no-information group showed little difference between the nine fonts (see table 1 and graph 1 till 3).

font	Volume	Duration	Pitch
full bold	100%	101%	100%
full raised	101%	102%	100%
full stretched	100%	101%	101%
full wide	100%	102%	100%
half bold	100%	100%	101%
half raised	100%	101%	97% **
half stretched	100%	98%	99%
half wide	100%	98%	101%
normal	100%	100%	100%

Table 1: Average of one condition divided by the average of the normal condition for volume, duration and pitch when no instructional information was provided to the participants. *'s indicates significant difference from normal font: *=$p<0.05$; **=$p<0.01$; ***=$p<0.001$

There are no statistically significant differences between the fonts for volume, nor for duration. For pitch, only 'half raised' differs statistically significant from the normal font ($p=0.01$).

The information group

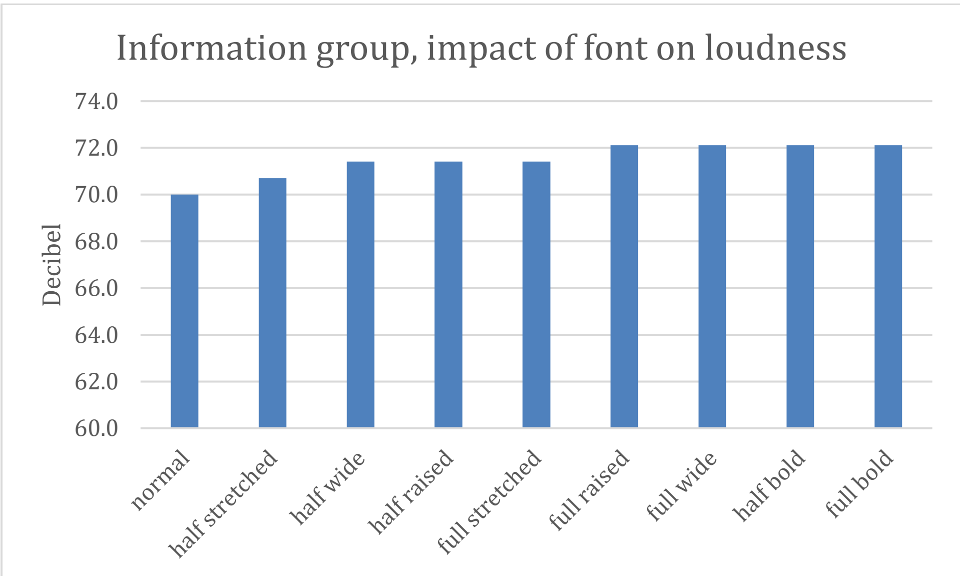
In the information group, all fonts differed significantly for all three measures with the normal font (table 2).

typeface	Volume	Duration	Pitch	Volume example dB	Pitch example Hz	Duration example sec
full bold	103% ***	134% ***	107% ***	72	257	0,19
full raised	103% ***	135% ***	111% ***	72	266	0,19
full stretched	102% ***	133% ***	106% ***	71	254	0,19
full wide	103% ***	149% ***	106% ***	72	254	0,21
half bold	103% ***	129% ***	108% ***	72	259	0,18
half raised	102% ***	116% ***	103% **	71	247	0,16
half stretched	101% *	118% ***	103% **	71	247	0,17
half wide	102% ***	117% ***	103% ***	71	247	0,16
normal	100%	100%	100%	70	240	0,14

Table 2: Average of one condition divided by the average of the normal condition for volume, duration and pitch when instructional information was provided to the participants. *'s in the 3 first columns indicate significant difference from normal font: *=p<0.05; **=p<0.01; ***=p<0.001.

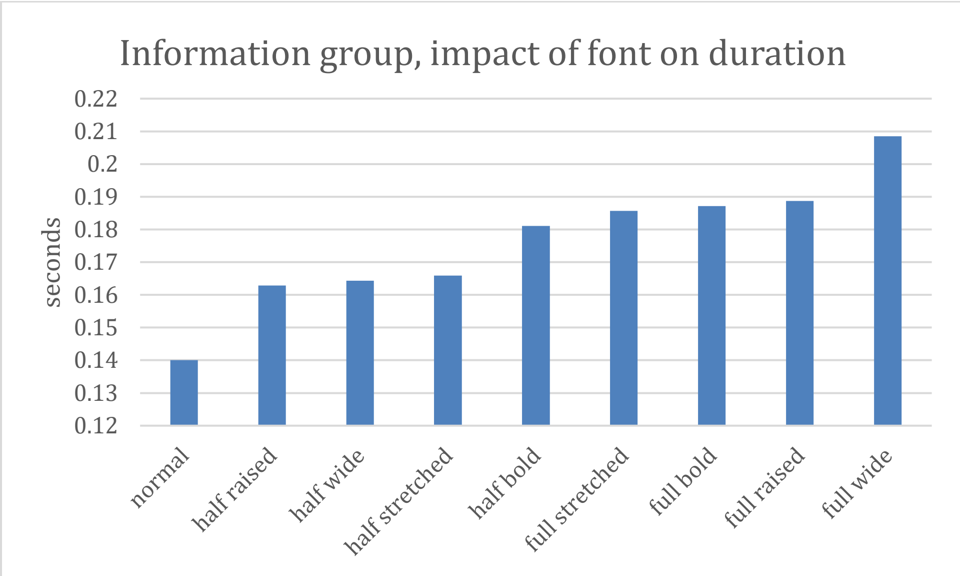
To have a feeling of the impact in a realistic situation, we added for each component an example in the last three columns of table 2. E.g. assume for volume a word which is spoken, when using the normal font, with a volume of 70 dB (which is very near to the average of the volume measured in this experiment). In the condition full bold, this can be on average multiplied with 103%, hence the volume of the pronunciation would be 72dB. For duration a word of 0.14 seconds and for pitch a word of 240 Herz are given as examples. Also these absolute values are very close to the averages found in the dataset when normal font was used.

Children in the information group read words louder when presented in 7 of the fonts compared to the normal font see graph 4). Volumes of all fonts differed statistically significant from the volume of words in the normal font. The largest effect on volume was for the full bold, full wide, half bold, full raised conditions, an increase of 3% over the normal font. The following graph represents the effect of font the volume on a word spoken by a child.



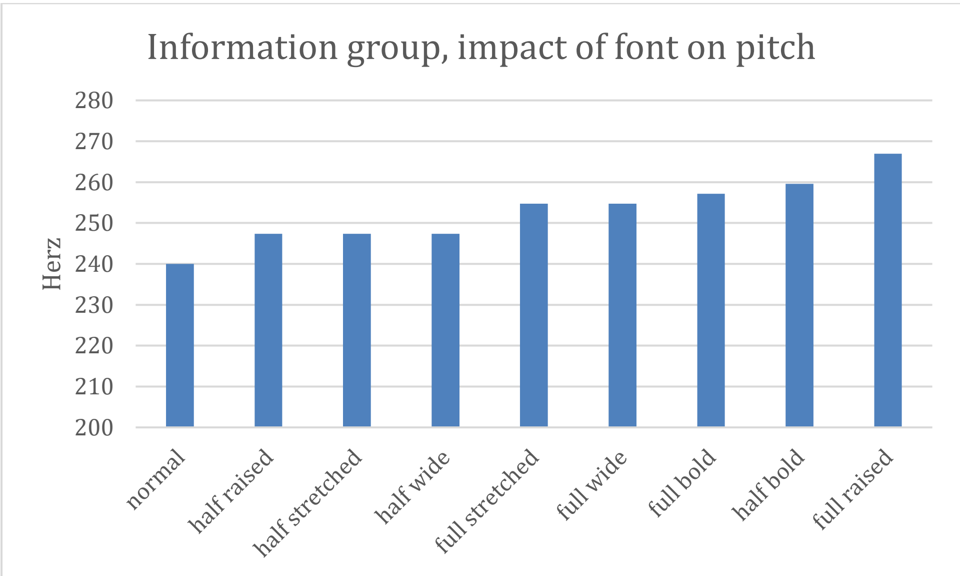
Graph 4: Visualization of the effect of the font on a word that, with a normal font would be expressed with a volume of 70dB

Children in the information group showed a statistically significant increase in duration of reading time for all test fonts compared to the normal font (see graph 5). The largest increase in duration was for the full wide font, which was read 49% longer than the normal font. The full wide font was read statistically significant longer than all other fonts.



Graph 5: Visualization of the effect of the font on a word that, with a normal font would be expressed with a duration of 0.14 seconds

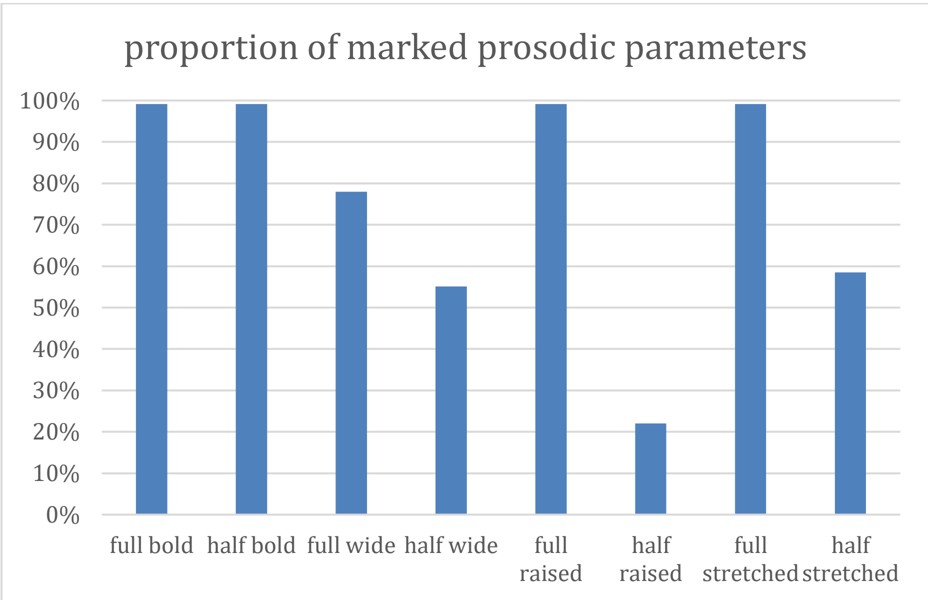
Children in the information group showed a statistically significant increase in pitch for all the fonts compared to the normal font (see graph 6). The largest increase in pitch was for the full raised font, which had vowels spoken at a 11% higher pitch than the normal font. Full raised was read at a statistically significant higher pitch than all other fonts.



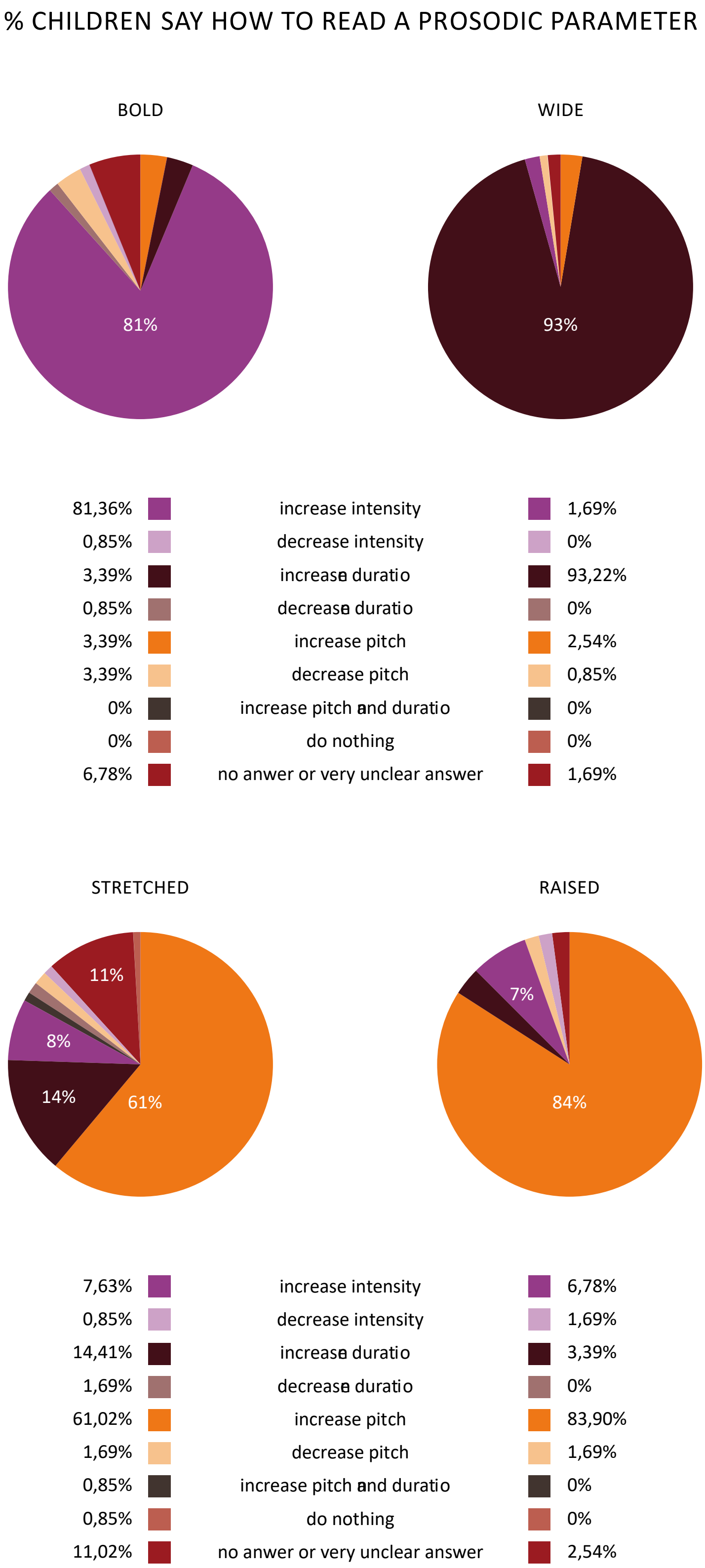
Graph 6: Visualization of the effect of the font on a word that, with a normal font would be expressed with a pitch of 240 Hz.

Questionnaire

Two days after the test, participants were asked to identify words that should be spoken differently and to say how they should be spoken. The half bold and full bold fonts were correctly recognized 99.2% of the time. 81% of the participants said they should be read *loud*, *louder*, or *harder*. Only 7% gave no answer or a very unclear answer. The half wide font was recognized 55% of the time and the full wide was recognized 78% of the time. 93% of the participants correctly said they should be read *long* or *longer*. The half raised font was recognized 22% of the time, while the full raised font was recognized 99% of the time. 83% of the participants correctly said that they should be read *high* or *higher*. 7% of the participants incorrectly said they should be read *louder*. The half stretched font was recognized 49% of the time and the full stretched font was recognized 99% of the time. 61% of the participants correctly said they should be read *high* or *higher*. 15% of the participants incorrectly said they should be read *longer*.



Graph 7: The proportion of words containing a specific parameter that are marked within a sentence, two days after the test.



Pie Chart 1: The answers that the children gave on the question 'how to read the given prosodic parameter?' two days after the test.

At the end of the questionnaire, the participants were given an opportunity to provide their opinions about the fonts. As design researchers, we believe that these rather subjective (when compared to the statistical output) opinions from the participants are of great importance in determining whether or not the test material has an actual chance of being used in real reading material for children, which they find pleasing to use. The reactions were generally positive with the participants enjoying the increased variety in the text and the additional support to help reading aloud. One participant said that it was "easier for reading because you don't have to read on the same tone and then it does not become boring." Another said that reading was "easier because you know if you have to read longer, higher or louder." Yet another one described the experience as if you were communicating in real life instead of reading.

4. Discussion

The goal of this project was to help children read aloud with more expression. We focused on the prosodic components volume (louder pronunciation), duration (slower pronunciation) and pitch (higher pronunciation). While we hoped that the cues would be understood without any explanation, we found that the children who received no explanation choose to focus on reading the sentences quickly instead of with greater expression. Reading for speed is the most common form of reading assessment, so it's not entirely surprising that the kids attempted to read quickly (Mostow & Duong, 2009). It was predicted by some of the teachers that children, due to the intensive testing for measuring their reading level based on speeds and accuracy, would interpret the test in this way and thus read as fast and correct as possible, while ignoring prosody. Consequently, no statistically significant differences in the prosody measures were found.

The children who were given an explanation of the prosody conditions read them aloud quite clearly. Interestingly, prosody marked words tended to be spoken with increased amounts of all kinds of prosody. For example, the words marked with increased volume were read with statistically significant higher volume, but were also read slower and at a higher pitch. The same was true for the other conditions.

All of the conditions caused the children to read the target word louder. The effect was strongest for the full conditions and both full and half bold. Patel and Furr (2011) found no effect of using grey levels to change kids' volume. But in our study, we found a reliable increase of volume for the same conditions that increase pitch reliably (changing the x-height). Pitch and volume are strongly related as usually people speak louder when they raise their pitch. This relationship was not intentional as we hoped to typographically convey each prosody factor independently.

Widening the letters in a word was effective for getting the kids to pronounce a word for a longer amount of time. This is in line with the findings of Patel and Furr (2011). But in our study we found that all conditions increased duration statistically significant, though not as dramatically as wide letters. It is surprising that the other conditions also led to increased durations. This might have happened because the children needed additional cognitive resources to correctly pronounce all of the prosody conditions, causing an increase in duration. We observed this most clearly with the stretched and raised conditions (meant to increase pitch) as the children found raising their pitch effortful.

Both the full raised and full stretched pitch conditions caused a statistically significant increase in spoken pitch. The half raised and half stretched conditions did increase pitch less. The fact that the higher voice is more difficult to understand may be comprehended out of practice. Children do learn the difference between low and high in nursery school, however this is treated as a spatial phenomenon. It was often seen that when children needed to go higher in voice, they moved their body upwards. Because of the way high and low is taught, they often didn't know what to do with their voice. When a child had almost no problems in the pronunciation of a higher voice he was asked afterwards if he had a musical background. Often this was the case.

The half wide, half raised and half stretched font may be too subtle. Two days after the test with the explanation, they are recognized in less than 75% of the situations.

5. Conclusion and further research

Only the data of the information group show differences in recordings of those words that were highlighted with the prosodic cues. We believe that the no-information group experienced this test as a regular reading test for measuring their reading level. This test evaluates the child only on its reading speed and accuracy.

Within the information group, the analysis of the reading tests proves that the prosodic design parameters have the intended effect on the oral reading of children.

Reading aloud of a single prosodic component hardly happened without an interaction with the other prosodic components. However, when isolating each parameter regarding their hypothesized effects, all parameters differed statistically significant from the normal condition for pitch, volume and duration of speech. For each comparison with 'n', the full variation of the intended parameter gave the most significant results for the intended prosodic component. Thickening increased volume the most, widening the duration, raising x-height increased the pitch the most. The effect of 'full raised' on pitch was an average increase of 9%. The effect of 'full bold' on the volume was an average increase of 2%. The effect of ‘full wide’ on the duration was an average increase of 37%. The effects of parameters on the prosodic cues that were not intended were lower, and not always significant.

Based on the findings we recommend type designers to implement a thickened font when they would like children to guide in their speaking aloud with a louder voice by means of a typeface. Both ‘full bold’ and ‘half bold’ are good references for designers when they would like to implement a volume parameter within their typeface. Type designers involved in visual prosody are advised to widen the font, like our parameter ‘full wide’ when children should be guided to read with a slower voice. From a designer’s point of view, we question if it would be possible to design an even more extended type that is still aesthetically justified in terms of letter shapes and text color. When type designers want to implement a design parameter to guide the children in reading aloud with a higher voice, they can raise the x-height, as in ‘full raised’.

The hypothesis that visual prosody in type is able to influence children’s reading aloud with more expression is confirmed by this study. However, based on this research we can not conclude whether visual prosodic cues are sensed in an intuitive manner. Thus, instruction is needed. It is important to note that this research was conducted by Belgian children and that in general, Belgians are known to be rather reluctant when it comes to trying out things differently and rather do it in ways that are familiar to them (Hofstede, 2001). For example, when compared to the Dutch, Belgians are in general more introvert (Laurent, 1973; Portzky et al., 2008; Gerritsen, 2014). This characteristic may attribute to the fact that without instruction, the children may have seen the prosodic cues, but didn’t execute them when reading aloud. There is a chance that, when other nationalities conduct the same test, results may differ regarding the intuitive reading aloud of the parameters.

All in all, these type design parameters have the potential to influence the reading aloud of children, and can therefore assist type and typographic designers to create new typefaces and educational materials that aim to influence expressive reading. Within the new technology of OpenType Font Variations (introduced in 2016) these parameters can be more easily applied by typographers and usable by type designers.

Furthermore, the research to visualizing prosody through text proves to be promising for further research, not only on printed matter but also in digital reading. There is a great deal of enthusiasm for this work by publishing houses as it has the potential of making text more expressive and can teach children more consciously reading aloud skills. Expressive type may reduce the cause of confusion in written communication and might improve reading comprehension. Additionally, prosody also has a diverse range of other uses including making expressive captioning for the deaf community and teaching expression for the autistic community.

Acknowledgements

We would like to thank Wouter Vanmontfort for the adaptation of Praat; Tom Wollaert for optimising the audio recordings and Microsoft Advanced Reading Technologies for the grant.

Bibliography

* ASHBY, J. (2006). “Prosody in skilled silent reading: evidence from eye movements.” *Journal of Research in Reading*. 29, (3), 318-333.
* BECK, J.E., MOSTOW, J. (2008). “How who should practice: Using learning decomposition to evaluate the efficacy of different types of practice for different types of students.” *9th International Conference on Intelligent Tutoring Systems*, Montreal, June 23-27, 353-362.
* BESSEMANS, A. (2012) *Letterontwerp voor kinderen met een visuele functiebeperking*. PhD dissertation, Leiden University & Hasselt University. http://hdl.handle.net/1887/20032
* BOURSMA, P., & WEENINK, D. (2016). *Praat: doing phonetics by computer* (Version 6.0.08), Computer program. http://www.fon.hum.uva.nl/praat
* BOURSMA, P. (2001). "Praat, a system for doing phonetics by computer." *Glot International*. vol. 5, 341-345.
* BINDER, S. K., TIGHE E., JIANG Y., KAFTANSKI K., QI C., & ARDOIN, S. P. (2013). “Reading expressively and understanding thoroughly: An examination of prosody in adults with low literacy skills.” *Read Writ*. 26, 665-680.
* BOLINGER, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford, CA: Stanford University Press.
* CUTLER, A., DAHAN, D., & VAN DONSELAAR, W. (1997). “Prosody in the comprehension of spoken language: A literature review.” *Language and speech*. 40, (2), 141-201.
* DOWHOWER, S. (1991). “Speaking of prosody: Fluency’s unattended bedfellow.” *Theory into Practice*. 30, (3), 165-175.
* DUONG, M., MOSTOW, J., & SITARAM, S. (2011). “Two methods for assessing oral reading prosody ACM.” *Transactions on Speech and Language Processing (Special Issue on Speech and Language Processing of Children’s Speech for Child-machine Interaction Applications) 7,* (4), 11-22. doi 10.1145/1998384.1998388.
* GERRITSEN, M. (2014.). “Vlaanderen en Nederland: één taal, twee culturen?” *Neerlandia/Nederlands van Nu*. (1), 26-29.
* GUSSENHOVEN, C. (2004). *The phonology of tone and intonation*. New York, NY: Cambridge University Press.
* HOFSTEDE, G. (2001). *Culture’s consequences: comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks/London/New Delhi: Sage Publications.
* HUDSON, R.F., LANE, H.B., & PULLEN, P.C. (2005). “Reading fluency assessment and instruction: what, why, and how?” *The Reading Teacher*. 58, (8), 702–714.
* KÖHLER, W. (1929). *Gestalt psychology*. New York, NY: Liveright.
* KUHN, M.R., & STAHL, S.A. (2003). “Fluency: A review of developmental and remedial practices.” *Journal of Educational Psychology*. 95, (1), 3–21.
* LAURENT, P-H. (1973). “The Benelux States and the New Community.” Current History (pre-1986). April, 64, 166-182.
* LEHISTE, I. (1970). *Suprasegmentals*. Cambridge, MA, USA: MIT Press.
* LEWIS, C., & WALKER, P. (1989). “Typographic influences on reading.” *British Journal of Psychology*. 80, (2), 241-257.
* PROJECT LISTEN. (2009). LISTEN’s Reading Tutor. [Website] http://www.cs.cmu.edu/~./listen/ [Assessed: 2017, September, 24th]
* MILLER, J., & SCHWANENFLUGEL, P. J. (2006). “Prosody of syntactically complex sentences in the oral reading of young children.” *Journal of Educational Psychology*. 100, 310-321.
* MONETA, M. E., PENNA, M., LOYOLA, H., BUCHHEIM, A., & KÄCHELE, H. (2008). Measuring emotion in the voice during psychotherapy interventions: A pilot study. *Biological Research*. 41, (4), 389-395. DOI: /S0716-97602008000400004
* MOSTOW, J., & DUONG, M. (2009). “Automated Assessment of Oral Reading Prosody.” Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling, July 20, 189-196. doi 10.3233/978-1-60750-028-5-189
* NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT (NIH). (2000). Report of the National Reading Panel. *Teaching Children to Read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. NIH Publication 00-4769. Washington, D.C., USA: US Government Printing Office.
* NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP). (1995). “Listening to children read aloud: Oral fluency.” *NAEPFacts. 1,* (1), 2-5.
* PATEL, R., & FURR, W. (2011, May, 7-12). “ReadN’Karaoke: Visualizing prosody in children’s books for expressive oral reading.” *CHI Session: Books & Language*, 3203-3206.
* PEPPÉ, S. J. E. (2009). “Why is prosody in speech-language pathology so difficult?” *International Journal of Speech-Language Pathology*. 11 (4), 258-271.
* PLAYBOOKS. (2013). Playbooks; roleplay reader. [Website] http://www.readerstheater.com [Accessed: 2015, Januari, 2nd].
* PORTZKY, G., WILDE DE, E. J. & HEERINGEN VAN, K. (2008.) “Deliberate self-harm in young people: differences in prevalence and risk factors between The Netherlands and Belgium.” *European Child & Adolescent Psychiatry*. 17 (3), 179-186.
* RASINSKI, T. V. (1990). “Effects of repeated reading and listening while reading on reading fluency.” *Journal of Educational Research. 83,* 147-150.
* RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. New Jersey: Prentice Hall.
* READ NATURALLY, INC. (2015). Read naturally. [Website] http://www.readnaturally.com [Accessed: 2015, Januari, 2nd].
*

* SAMUELS, S. J. (1988). “Decoding and automaticity: Helping poor readers become automatic at word recognition.” *The Reading Teacher. 41,* 756–760.
* SCHREIBER, P. A. (1980). “On the acquisition of reading fluency.” *Journal of Reading Behavior. 7,* 177–186.
* SCHWANENFLUGEL, P.J., MEISINGER, E.B., WISENBAKER, J.M., KUHN, M.R., STRAUSS, G.P., & MORRIS, R.D. (2006). “Becoming a fluent and automatic reader in the early elementary school years.” *Reading Research Quarterly. 41,* (4), 496-522.
* SHAIKH, D. (2009). “Know your typefaces! Semantic differential presentation of 40 onscreen typefaces.” *Usability News. 11,* (2).
* SITARAM, S., & MOSTOW, J. (2012). “Mining data from project LISTEN’s reading tutor to analyze development of children's oral reading prosody.” Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference. 478-483.
* STANOVICH, K. E. (1986). “Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy.” *Reading Research Quarterly. 21,* (4), 360-407.
* STICHTING LEZEN. (2016). *The National Reading Aloud Competition.* [Website] https://cdn.denationalevoorleeswedstrijd.nl/wp-content/uploads/2018/05/National-Reading-Aloud-Competition-Rule-Book-for-web.pdf
* VAN UDEN, A. (1973). *Taalverwerving door taalarme kinderen.* Rotterdam: University press Rotterdam.
* YOUNG-SUK GRACE, K. (2015). “Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension.” *Reading Research Quarterly. 50,* (4), 459–481.