Visible
Language

55 . I

I

2021

april

Advisory Board

Contents

a lower V

Visual prosody supports reading aloud expressively for deaf readers

Maarten Renckens[1]
Leo De Raeve[2, 3]
Erik Nuyts[1]
María Pérez Mena[1]
Ann Bessemans[1]

voice

Type is a wonderful tool to represent speech visually. Therefore, it can provide deaf individuals the information that they miss auditorily. Still, type does not represent all the information available in speech: it misses an exact indication of prosody. Prosody is the motor of expressive speech through speech variations in loudness, duration, and pitch. The speech of deaf readers is often less expressive because deafness impedes the perception and production of prosody. Support can be provided by visual cues that provide information about prosody—visual prosody—supporting both the training of speech variations and expressive reading.

We will describe the influence of visual prosody on the reading expressiveness of deaf readers between age 7 and 18 (in this study, 'deaf readers' means persons with any kind of hearing loss, with or without hearing devices, who still developed legible speech). A total of seven cues visualize speech variations: a thicker/thinner font corresponds with a louder/quieter voice; a wider/narrower font relates to a lower/faster speed; a font raised above/lowered below the baseline suggests a higher/lower pitch; wider spaces between words suggest longer pauses.

We evaluated the seven cues with questionnaires and a reading aloud test. Deaf readers relate most cues to the intended speech variation and read most of them aloud correctly. Only the raised cue is difficult to connect to the intended speech variation at first, and a faster speed and lower pitch prove challenging to vocalize. Despite those two difficulties, this approach to visual prosody is effective in supporting speech prosody. The applied materials can form an example for typographers, type designers, graphic designers, teachers, speech therapists, and researchers developing expressive reading materials.

1 Hasselt University in collaboration with PXL University College, Belgium;
2 Independent Information Centre about Cochlear Implantation (ONICI) Zonhoven, Belgium;
3 KIDS (Royal Institute for the Deaf) Hasselt, Belgium.

Keywords
type design,
visual prosody,
prosody,
deaf readers,
expressive reading,
reading comprehension

1. Introduction

1.1. About prosody

Type visually provides information to deaf individuals that they otherwise do not hear. Therefore, type can be a useful tool for deaf individuals who are able to read. However, it remains an incomplete representation of speech: type often only focuses on "which words are said," but less on "how words are said."

How words are said is referred to as "speech expression" (Veenendaal, Groen & Verhoeven, 2014), and its motor is speech prosody: variations in the speech features loudness, duration, and pitch (Bessemans et al., 2019; Chan, 2018; Soman, 2017; Belyk & Brown, 2014; Karpiński, 2012; Sitaram & Mostow, 2012; Nakata, Trehub & Kanda, 2011; Patel & Furr, 2011; Patel & McNab, 2010; Wagner & Watson, 2010). Prosody plays an important role in language comprehension: it distinguishes homographs such as PREsent versus preSENT. It can add information to what is said, such as statements, questions, sarcasm, surprise, and it can influence the meaning of a sentence. For example, the sentence "That old man cannot hear you very well" has a different meaning if "cannot" or "you" is emphasized (Seidenberg, 2017; Wagner & Watson, 2010; Carlson, 2009; Verstraete, 1999; Guberina & Asp, 1981 for similar examples).

Prosody is required for all applications of language, even in fluent reading where a proper expression (of which prosody is the motor) is as important as a proper rate (speed), and accuracy (correct decoding of the letters) (Groen, Veenendaal & Verhoeven, 2019; Reading Rockets, 2019; International Literacy Association (ILA), 2018; Hasbrouck and Glaser, 2012; Paige, Rasinski, Magpuri-Lavell, 2012; Sitaram & Mostow, 2012; National Institute of Child Health and Human Development (NIH), 2000; National Assessment of Educational Progress (NAEP), 1995). During prosodic reading, a total of six prosodic characteristics are defined: pausal intrusions, length of phrases, appropriateness of phrases, final phrase lengthening, terminal intonation contours (e.g., lowering the voice after a group of words), and stress (Kuhn & Stahl, 2003 citing Dowhower, 1991). Correctly applying prosody in reading can improve word recognition, reading accuracy, reading speed, and comprehension skills, because expressive readers can segment text into meaningful units (ILA, 2018; Young-Suk Grace 2015; Veenendaal, Groen, Verhoeven, 2014; Binder et al., 2013, Carlson, K. 2009; Miller & Schwanenflugel, 2006; Ashby, 2006). A highly fluent reading leads to better reading motivation and comprehension (Hasbrouck & Glaser, 2012), while a less developed prosody is related to poorer comprehension (Groen, Veenendaal & Verhoeven, 2019; Gross et al., 2013 citing National Research Council, 1999). Kuhn and Stahl (2003, citing Schreiber, 1987) suggest that speech is easier to understand than reading because of its prosody. Even during silent reading, prosody plays an active role in fluent reading and reading comprehension (Breen et al., 2016; Leinenger, 2015; Young-Suk Grace, 2015; Gross et al., 2013; Ashby, 2006; Fodor, 1998). For example, prosody, indicated by periods, commas, question marks, exclamation marks, or other prosody indicators, influences how we read and clarifies the intention of the sentence.

1.2. Speech prosody influenced by hearing loss

Despite its importance, prosody is challenging to master by almost all individuals with hearing loss (Hutter, 2015; Marx et al., 2014; Stiles & Nadler, 2013; See et al., 2013; Wang et al., 2013; Nakata, Trehub & Kanda, 2012; Vander Beken et al., 2010; Lyxell et al., 2009; Peng, Tomblin & Turner, 2008; Markides, 1983). It is important to look at their use of prosody within the broader context of their hearing problems.

The terms *deaf*, *hard of hearing*, *hearing loss*, or *hearing impaired* refer to a suboptimal perception of sounds from the environment. The cause may be a deficit within the outer ear, the inner ear, a damaged nerve, and/or brain damage. Hearing loss can occur before or after mastering basic understanding of oral communication (pre-and post-lingual deafness).

All kinds of hearing loss limit the information perceived from the environment, which results in fewer stimuli to develop cognition. So, individuals with hearing loss often experience a disadvantage in their general learning process. If not countered by the use of sign language, hearing devices, or very good support, a hearing loss can cause a delay or suboptimal development in:

- cognition and language development (see, for example, De Raeve, 2014; Boons et al, 2012, 2013a, 2013b, 2013c; Fagan & Pisoni, 2010);
- speech (see, for example, Baudonck et al., 2015; Limb & Roy, 2014) and use of prosody (see, for example, De Clerck et al., 2018; Øydis, 2013);
- reading fluency (see, for example, Mayer et al., 2016; Luckner & Urlbach, 2011). Not all individuals with hearing loss can read well because they are not able to relate letters with the correlating sounds.

Most of these steps are non-chronologically interwoven, and mastering each of them is a process taking several years.

The group of individuals with hearing loss is a very diverse group with much differentiation. Individuals with the same amount of hearing loss and who receive the same support can still have a different cognitive development. Due to the hindrances in connecting to the 'hearing society,' individuals with hearing loss often rely on sign language, and they developed their own Deaf culture (doof.nl, 2017; Fevlado, 2013).

The sooner intervention takes place, the lesser the delays in (speech) development. Nowadays, hearing devices such as cochlear implants (CI's) mostly restore the provision of sound stimuli, enabling the development of spoken language at an age-appropriate level (Hearing Team first, 2017). Still, the sound output of the devices does not perfectly resemble the sound perceived by a hearing individual (Scarbel et al., 2012), and each device needs to be calibrated for the individual wearing it, the "fitting." For a hearing individual, the sound received by the cochlear implant could be described as a low-quality sound (listen to the Daily Mail Online (2014) at https://www.dailymail.co.uk/sciencetech/article-2636415/What-deaf-hear-Audio-file-reveals-s-like-listen-world-using-cochlear-implant.html for an auditory example made by Michael Dorman, an Arizona State University professor of speech and hearing science). After a while, the brain will adjust to the input and process the sound information in the best possible way. The improved hearing status improves phonological awareness, resulting in an increase in general literacy (Mayer et al., 2016; Harris, 2015; Dillon, Cowan & Ching, 2013).

So, while speech has become more accessible than ever before for a part of this group, learning to speak fluently remains an adventure that not all of them bring to a positive end. The impact on the topic at hand, namely prosody, is that neither prosodic perception nor prosodic production by individuals with hearing loss is similar to that of their hearing peers. While some children with implants even produce speech containing minimal to no differences compared to typical hearing children (Boons et al., 2013b, See et al., 2013; Boons, 2013; Vanherck & Vuegen, 2009), the perception of prosody is hindered by limitations of the hearing devices. Production of prosody is in general flatter than that of their hearing peers. Their speech still can be reliably distinguished from their hearing peers (Boonen et al., 2017). The achieved speech quality depends on the hearing threshold, the hearing devices, the applied therapy, and more. One of the aspects that differ compared with their hearing peers is their production of prosody, even for the younger generation of deaf individuals who are wearing hearing devices from an early age (De Clerck et al., 2018; Øydis, 2014; Wang et al., 2013; Chin, Bergeson & Phan, 2013). Compared to typical hearing children, children with cochlear implants demonstrate a smaller pitch range in their utterances (De Clerck et al., 2018; Øydis, 2014); a lower pitch modulation (Wang et al., 2013); a divergent nasal resonance (Baudonck et al., 2015) or a lesser application of prosody in general (Chin, Bergeson & Phan, 2013). Several of those imperfections relate to the three speech variations loudness, duration, and pitch, which form the base of prosody. To optimize the speech and prosody of children with hearing loss, training sessions that include singing, vocal exercises, movements with the body, and more are default practice in their education (KIDS, n.d.; Advanced bionics, unpublished; Vander Beken et al., 2010; Asp, 2006; Guberina & Asp, 1981). On the one hand, these sessions ensure that the vocal cords receive the training they need, while on the other hand, the children become aware of the vocal variations, namely "how words are said."

1.3. Evaluating visual prosody for deaf readers

While prosody is essential during the reading process, type lacks exact representations for several prosodic functions. If type would implement more indicators to the intended prosody, suggesting "how words are said," deaf readers would gain more access to the prosody they partly miss.

Visualizations of prosody in type already exist and are mostly intended to encourage speech variations (Bessemans et al., 2019; Patel & McNab, 2010; Staum, 1987; van Uden, 1973). These typographic visualizations of speech prosody are referred to as *visual prosody*. Visual prosody adds visual cues to type, each cue hinting to a particular aspect of prosody, such as one of the three individual speech variations loudness, duration, or pitch. Several centers of expertise in deaf education apply visual prosody in teaching materials on a pragmatical basis to exercise vocal variations with deaf readers (KIDS, n.d.; Advanced bionics, unpublished; Staum, 1987; van Uden, 1973). For some examples, see Figure 1. When visual prosody is applied, most existing cues are relatively intuitive, meaning that their intention can be spontaneously interpreted (Shaikh, 2009; Lewis and Walker, 1989). For example, a bold typeface relates to louder sounds (Shaikh, 2009). The intuitive character provides information about "how easily a reader would apply the prosodic cues as intended if no explanation is provided."

Figure 1.a, b & c.

Some schools in (Dutch) deaf education apply visual prosody as a teaching method. 1. Horizontally stretched words (lllllaaaaannnngggg, lllloooonnnnngggg) correlate with a longer duration. 2. Bolder and larger text correlates with more loudness, and 3. A higher position, vertically stretched text, a rising line, or music notes correlate with a higher pitch in the voice (left to right: image from KIDS, n.d.; image from Advanced Bionics, unpublished; image based on van Uden, 1973).

Sadly enough, empirical information about how deaf readers handle prosodic cues and/or read them aloud is not yet available. We, as design researchers specialized in typography and type design, are interested in how changes in typography influence reading and how designers can optimize those changes. Therefore, this research aims to optimize visual prosody for deaf readers between 7 and 18 years old in the Dutch language (Flanders region in Belgium and the Netherlands) by testing several cues representing all speech variations.

Visual prosody is positioned as a visual manner to encourage expression, and where necessary, to teach how speech prosody should sound. In this study, we will only focus on the reading expressiveness (the application of prosody while reading) and not yet on reading fluency. While we're very much aware that visual prosody is not a magical solution to solve all problems that readers with hearing loss encounter (cognition, language development, speech, prosody, and fluent reading), we believe visual prosody could support part of this audience when developing reading and speech skills. If successful, these cues could be applied in later studies evaluating speech training over a longer time period, or in studies aiming to improve their reading fluency and thus overall literacy.

In this study, visual prosody is approached with the hypothesis, *"Visual prosody leads to more vocal prosody while reading aloud, and influences reading comprehension of deaf readers between 7 and 18."* In this study, 'deaf readers' refers to 'deaf students who have developed spoken language that is distinct enough to be understood, in a signed bilingual educational setting.' Three research objectives, with multiple sub-questions, evaluate the hypothesis:

1. " How intuitive is visual prosody for deaf readers between 7 and 18?"
    a. "What is the reader's perceived intention of visual prosody?"
    b. "How noticeable are the prosodic cues?"
    c. "How well does each deaf reader relate the prosodic cues to the intended speech variation?"
2. "Does visual prosody increase the speech variations of deaf readers between 8 and 18?"
3. "Does visual prosody influence the understood meaning of a sentence for deaf readers between 8 and 18?"

While evaluating these sub-questions, we collected as much relevant information as possible about each participant: the amount of hearing loss, pre/post-lingual deaf, the native language, hearing device, etc., to evaluate their influence on how visual prosody is handled.

2. Methodology

Participants

Because of the great diversity in this audience (some describe it as the most diverse audience in terms of perceived problems, provided support, and thus personal development), we set out three principles which the participants had to meet.

Firstly, because this is the first known study to evaluate the prosodic cues for individuals with hearing loss empirically, we are interested in the cues' effect on their reading aloud. Thus, all participants in this study should be able to read aloud. In these times (and in the Western world), this is not a big issue. This research took place primarily in Flanders, the Dutch-speaking region of Belgium, where the care for individuals with hearing loss is well established. This region offers universal neonatal hearing screening to all newborns since 1998, within the first three weeks of life (Vander Beken, 2010). Children who were referred by the universal hearing screening test (the Maico test) are redirected to a referral center for audiological diagnostics and early intervention. The broad application of neonatal hearing screening ensures that most of the individuals with hearing loss receive early supervision, and fast implantation is recommended and re-funded by the Flemish government. Most of the younger individuals with hearing loss now wear cochlear implants. At the secondary education level in Belgium, more than 70 percent of students with hearing loss attend regular schools (De Raeve et al., 2012). They receive additional support from schools for the deaf for several hours per week in the form of speech therapy or extra exercises, a sign or speech-to-text interpreter, etc. This way, most individuals with hearing loss in Flanders can develop spoken language to a certain degree (this is independent from spoken or signed language being their native and/or most used language).

Secondly, we determined that deaf readers aged 7 to 18 could benefit most from visual prosody during their educational career. Therefore, it was a prerequisite in this study that participants mastered technical reading, in the form of the automatized decoding of letters, which is needed before fluent reading can be established (Groen, Veenendaal & Verhoeven, 2019; Miller & Schwanenflugel, 2006). Decoding text is the act of recognizing letter sequences as a word (technical reading). They first start to relate letters to a sound, to "crack the code," learning which letter belongs to which speech sound. When the technical reading skills are fully acquired, learners can spend more attention on prosody. The age of 7 became the youngest age to participate because, at this age, readers possess the technical reading skills to read sentences as a whole unit instead of separate words. The target age was limited up to the age of 18: at this age, compulsory schooling in the Dutch-speaking regions ends.

Thirdly, cognitive disorders of any kind that heavily impede their learning development were excluded. Participants' characteristics were carefully checked, and participants who did not meet the requirements were excluded from the research.

Within these boundaries, a very heterogeneous group of 38 deaf readers participated in this study. Their characteristics are described in table 1.

- They were on average 12.21 years old; the youngest one was 7.2y and the oldest one 19.4y (on June 30th, 2018. This was not the date on which all participants were tested).
- One participant had two deaf parents; two participants had a deaf mother; the other 35 participants had hearing parents.

- Thirty-three participants were prelingual deaf. Two were post-lingual deaf (after the age of 3), and from two participants, this information was not available.
- Nineteen participants had a bilateral hearing loss > 90 dB on one or more sides; 15 participants had a hearing loss between 89 and 27 dB. For 4 participants, this information was not available.
- Thirty-six participants were wearing CI's, hearing aids, or a combination. Two participants did not wear a hearing device because those participants only had mild hearing loss.
- A total of 30 participants were educated in a regular school. Only 8 were educated in a special school for the deaf. In general, that is a fair reflection of the target audience, of which 70% attend secondary school within regular education (De Raeve et al., 2012). The smaller number of participants in special secondary education correlates with the trend that deaf participants move to regular education after primary school.

Table 1.

Information about the diverse group of participants in this study.

Information about the pupils							Information about the language		Information about the ears				
Code number	Primary or secondary?	Special education (SE) or support network (SN)	class	Gender (M/F)	Age at June 30th, 2018	Deaf parents?	Knows Sign Language?	Native language (NL, FR, VGT, NGT,...)	Months when becoming deaf/hearing impaired (0 is born deaf)	Left: threshold level	Left: hearing aid? None/Hearing aid/CI	Right: threshold level	Right: hearing aid? None/Hearing aid/CI
HI002	Pri	SN	6	F	11y9m	N	N	NL	48	120	CI	120	CI
HI003	Pri	SN	5	F	11y2m	N	N	NL	36		HA		HA
HI004	Pri	SN	5	F	11y5m	Y	Y	NL	0	91	HA	91	HA
HI005	Sec	SN	2	F	14y10m	N	N	NL	0	120	CI	120	CI
HI006	Sec	SN	6	F	18y3m	N	Y	NL	0	120	CI	120	CI
HI007	Sec	SN	6	M	17y11m	N	N	NL	0	50	HA	50	HA
HI008	Sec	SN	6	M	17y7m	N	N	NL	0	average	HA	average	HA
HI009	Sec	SN	3	M	16y3m	N	N	NL	0	47	HA	45	HA
HI010	Sec	SN	3	M	15y0m	N	N	NL	0		N	light loss	Y
HI011	Sec	SN	2	F	14y0m	N	N	NL	0		N		N
HI012	Sec	SN	1	M	14y4m	N	Y	Maroccan (learned NL)	0	27	CI	120	CI
HI013	sec	SN	4	F	15y7m	N	Y	NL	30	50	HA	55	HA
HI014	Pri	SE		M	9y9m	N	Y	NL	0	110	CI	115	CI
HI015	Pri	SN	6	F	16y2m	N	little	NL	0	90	HA (but not always wearing)	120	CI
HI016	Sec	SE	1	M	13y9m	N	N/little according to him	NL	0	113	CI	113	CI
HI017	Pri	SE		F	11y7m	N	Y	NL	18	95	CI	115	CI
HI018	Pri	SN	5	M	12y		Y						
HI019	Pri	SE	3	F	19y4m	N	Y		24	67	HA	75	HA
HI020	Pri	SE	5	M	13y2m	N	Dutch with gestures	Turkish (learned NL)	0	72	HA	75	HA
HI021	Pri	SE	5	F	10y6m		Dutch with gestures	NL	6	77	HA	78	HA

Information about the pupils							Information about the language		Information about the ears				
Code number	Primary or secondary?	Special education (SE) or support network (SN)	class	Gender (M/F)	Age at June 30th, 2018	Deaf parents?	Knows Sign Language?	Native language (NL, FR, VGT, NGT,…)	Months when becoming deaf/hearing impaired (0 is born deaf)	Left: threshold level	Left: hearing aid? None/Hearing aid/CI	Right: threshold level	Right: hearing aid? None/Hearing aid/CI
HI024	Pri	SN	4	M	9y9m	N	N	NL	0	83	HA	72	HA
HI027	Sec	SN	5	F	16y8m	N	N	NL	0	100	HA	88	HA
HI028	Sec	SN	2	F	14y0m	N	N	NL	0	110	CI	91,65	HA
HI029	Pri	None	2	M	8y0m	N	Dutch with gestures	NL	32		HA		HA
HI030	Sec	SN	5	M	17y7m	N	N	Turkish & NL mixed	0	average	HA	average	HA
HI031	Sec	SN	2	F	14y4m	N	N	NL	0	120	CI	120	CI
HI032	Pri	SN	5	M	11y0m	Mother	Y	NL	0	118	CI	118	CI
HI033	Pri	SN	3	M	9y4m	N	Y	NL	0	120	CI	120	CI
HI034	Pri	SN	2	F	8y1m	N	N	NL	0	63	Y	62	Y
HI035	Pri	SN	6	F	11y9m	N	little	NL	0	120	CI	120	CI
HI036	Sec	SN	1	F	12y		Little	NL		120	CI	120	CI
HI037	Pri	SN	6	F	12y2m	N	Y	NL	0	100	CI	100	CI
HI038	Pri	SN	6	F	12y2m	N	Y	NL	0	100	CI	100	CI
HI041	Pri	SN	2	M	7y9m	N		NL	0	100	CI	100	CI
HI042	Pri	SN	4	M	7y6m	N		NL	0	71	HA	71	HA
HI045	Sec	SN	4	F	16y5m	Mother	N	NL	0	33	HA (but not always wearing)	58,8	HA (but not always wearing)
HI046	Pri	SE	6	M	10y7m	N	N	NL	4	65	HA	55	HA
HI047	Pri	SE	5	F	7y2m	N	N	NL	1	30		30	

Visual prosody applied in this study

The prosodic cues as applied in Bessemans et al. (2019) formed the basis for this study. These cues were adjusted to represent both directions of the speech variations: the thickness of the letters correlates with a louder/softer voice; the width of the letters correlates with the duration of what is said; the vertical height of the letters correlates with the height of the pitch. Additionally, a larger space connects to the duration, correlating with a longer pause. All applied fonts are shown in Figure 2.

Note that not all cues are symmetrical. For example, where the raised cue was moved up 250 units, the lower cue was only moved 125 units. Design experiments showed that moving letters down below the straight baseline was more noticeable than moving letters above the often curved x-height. The advantage of less vertical displacement is avoiding collisions between lines of text.

Figure 2.A, B, C & D.

Words set differently within a sentence form prosodic cues that indicate a specific speech variation. For some cues, gradations were implemented. The Dutch sentence translates to "The poor man stayed behind, alone."

A.	De arme man bleef alleen achter.	'Full thinner' for a quieter vocalization
	De arme man bleef alleen achter.	'Half thinner' for a quieter vocalization
	De arme man bleef alleen achter.	'Normal' for a normal vocalization
	De arme man **bleef** alleen achter.	'Thicker' for a louder vocalization
B.	De arme man *bleef* alleen achter.	'Full oblique' for a faster vocalization
	De arme man *bleef* alleen achter.	'Half oblique' for a faster vocalization
	De arme man bleef alleen achter.	'Full narrower' for a faster vocalization
	De arme man bleef alleen achter.	'Half narrower' for a faster vocalization
	De arme man bleef alleen achter.	'Normal' for a normal vocalization
	De arme man bleef alleen achter.	'Wider' for a slower vocalization
C.	De arme man bleef alleen achter.	'Lower' for a lower pitch
	De arme man bleef alleen achter.	'Normal' for a normal vocalization
	De arme man bleef alleen achter.	'Higher' for a higher pitch
D.	De arme man bleef alleen achter.	'Normal' for a normal vocalization
	De arme man bleef alleen achter.	'Double space' for a longer pauze
	De arme man bleef alleen achter.	'Tripple space' for a longer pauze

The test materials

Test material for Objective 1: "How intuitive is visual prosody for deaf readers between 8 and 18?"

The sub-question *"What is the reader's perceived intention of visual prosody?"* was evaluated by means of a short video fragment including subtitles showing prosodic cues [Figure 3]. In the booklet, the participants were asked, *"Why do some words look different in the sentence, according to you?"* The intended answer required a link to speech variations or a reference to speech expression.

Figure 3.

Presenting the prosodic cues together with a short video fragment evaluated if participants related visual prosody to speech prosody.

The sub-question *"How noticeable are the prosodic cues?"* was evaluated by presenting all cues from Figure 1 in mixed order within a list of sentences. Participants were invited to mark the prosodic cue within each sentence if one was present. Marked more often indicates a higher noticeability. One additional sentence was added to the list to check if the noticeability of the lowered pitch cue would be influenced by the word's context: when this cue is followed by letters with descenders it might become less noticeable.

The sub-question *"How well can each deaf reader manage to relate the prosodic cues to the intended speech variation?"* was evaluated by sentences wherein a prosodic cue was applied on one word, followed by the question of how they would pronounce that one word. Participants could mark the answer in a list containing all possible speech variations: louder, quieter, higher, lower, faster, slower. The enlarged space was treated differently: in a multiple-choice, participants could choose from breath in, divide the sentence into parts, breath out, wait longer/take a pause, something else.

Test material for Objective 2: "Does visual prosody increase the speech variations of deaf readers between 8 and 18?"

To answer the second objective, booklets with sentences intended to read aloud were created. Those booklets carefully incorporated the results from objective 1. To optimize the representation of all cues, the prosodic cues which were marked most often in the test for sub-question *"How noticeable are the prosodic cues?"* were applied: *'thicker,' 'full thinner,' 'full narrower,' 'wider applied on a longer word,' 'higher,' 'lower'* and the *'triple space.'* At the same time, the *'oblique cue'* was not implemented as this cue was not often related to its intended speech component.

Before the actual reading test, the participants received a small exercise-booklet containing all cues. This information allowed them to memorize the intended voice variation for each prosodic cue and to exercise those voice variations for a short while. Providing a separate exercise booklet prevented the participants from seeing the final test sentences in advance while still acquainting them with the usage of prosodic cues.

For the actual reading tests, the design of the booklets differed per age group. Participants were grouped according to third and fourth grade (approximately 7 till 10 years old) and fifth and sixth grade (approximately 10 till 12 years old) of the primary school plus the secondary school (approximately 12 till 18 years old). Each age group received five different sentences adjusted to their reading level, and each of those five sentences was presented nine times: twice in a regular condition, and seven times alternating the word that contained one of the prosodic cues. To avoid the impact of learning effects (by the repetition of the same sentence) on the outcome of the experiment, five different booklets were made for each age group. Those five booklets all contained a different random order of the sentences.

To create an optimal reading experience, the sentences were presented in a way similar to reading materials familiar to each age group. All sentences were presented in a booklet with slightly off-white to yellow paper. For the age group 7-10, there were 5 sentences per page in a corps of 16 pt. For the age group 10-12, there were 5 sentences per page in a corps of 14 pt. For the older ones, aging 12-18, there were 8 sentences per page in a corps of 12 pt, the size almost reflecting that of reading books for adults.

To increase the reading pleasure for the two youngest age groups, the encouraging sentence "Halfway! Well done." was expressed in the middle of the booklet. This allowed the participants to have a break, which was found necessary to keep the youngest participants focused till the end of the test (as in Bessemans et al., 2019).

Test material for Objective 3: "Does visual prosody influence the understood meaning of a sentence for deaf readers between 7 and 18?"

This latest objective evaluated if visually emphasized words within a sentence influence the understood meaning of the whole sentence. Ten sentences were created. To compensate for the divergent reading levels of the participants, two different sentences were developed per age group 7–9, 9–10, 10–11, 11–12, and 12–18 years old. The sentences were reviewed by speech therapists on feasibility, and each of those sentences was presented three times to the participant, each time with a different emphasized word.

One such sentence was, "That old man cannot hear you very well." The prosodic cue "thicker" was used to emphasize one of the words. Participants were then asked to mark the perceived meaning of the sentence in a list. The possible meanings in that list referred to a specific word, such as "you." If the word "you" was emphasized, participants were then expected to mark the corresponding meaning "do something about your speech." The possible meanings in the list relied as little as possible on a literal definition of one of the words.

The research procedure

In the first stage, the schools for the deaf were contacted. To comply with the privacy regulations, the supervisors (teachers/therapists) selected the children who met the participation requirements in this research. After that, each participant was visited twice in their school: an initial visit to test how intuitive visual prosody is and a second follow-up visit for testing the reading aloud and the influence of visual prosody on the meaning of a sentence.

During the initial visit for the first test of the study, they received the first booklet about how intuitive visual prosody is. No information was provided beforehand. Participants whose first answer did not relate to speech were encouraged to guess a second time what the emphasized words could mean. Independent of the second answer's correctness, the test continued with the next question. During the test, participants gradually received the required information for each exercise. At the end of this first visit, each participant knew that visual prosody serves to enhance expressive reading.

During the second visit, the focus was first on reading aloud. Participants received an exercise-booklet first. They were allowed to briefly repeat the visual cues to get used to reading visual prosody aloud. This short repetition helped to refresh and memorize the intended speech variations; to briefly train the vocalization of speech variations; to grow comfortable to the test and to speak into the microphone [Figure 4]. It was emphasized to the participants that they were allowed to read at their own pace to avoid acting as if this was a reading test for speed. As soon as participants were at ease with the procedure, a second booklet that matched their reading level was provided to them. Participants were asked to read the sentences aloud the best way they could with attention for the expressiveness. During the test, the researcher pointed with a finger to the sentence that was to be read aloud, making sure that all the sentences were read.

Figure 4.

Each participant was free to set up the microphone and booklet as desired. The participant in this photo was one of the few participants who preferred the booklet next to the microphone.

After the reading test, participants received the questionnaire about how visual prosody influences the understanding of a sentence. Participants were asked to mark the answer which corresponded the most with the sentence. Only when a participant could not understand the intention of the test, specific questions were asked to draw attention to the emphasized word within the sentence and what the location of the emphasis would involve for the meaning of the sentence. If really needed, participants were encouraged to read the sentence aloud. No hint was given about how a relation could be made.

At the end of the second visit, each participant was able to write down feedback about visual prosody and to provide comments in open questions focusing on the appreciation of the cues. The written feedback allowed each participant to express their opinions, ideas, concerns, or suggestions about this approach to visual prosody. It also provided the possibility for the researcher to ask additional questions, for example, about difficulties experienced during the test.

The data collection, conversion, and analysis

The reading aloud of each participant was recorded with an XML 990 microphone and processed with the application Praat (Boersma & Weenink, 2014). The application was extended and given the ability to split the recordings between sentences and automatically name and number the files (Renckens & Vanmontfort, 2015a). The research group ESAT (Catholic University of Leuven) performed the speech recognition to determine the place of the most important vowels within all recordings. A newly developed plugin for Praat extracted the required data of each sound recording (Renckens & Vanmontfort, 2015b). The analysis of prosody (loudness, duration, and pitch) was based on the values of the most important vowels in the words marked with prosodic cues. The decision to use vowels was based on:

- Within a single word, prosody can vary fast and several times. Peaks are often situated on the vowel. Analyzing longer speech fragments (such as whole words) would make it more difficult to compare the effect of the cues, a problem that Patel & McNab (2011) probably encountered in their first analysis that did not deliver the expected results.
- Smaller fragments within the speech allow a more precise analysis of the intended effect. We aim at correct vocalizations, such as "bEEEEEEr" for the Dutch word beer (bear) instead of the incorrect pronunciation "beeRRRRRRR." That latest would sound wrong in the Dutch language. An analysis on the vowel omitted unintended effects of visual prosody.

With X as the loudness, 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 the thicker cue, compared with the average of {all the average pitches of all the "ee" of all the words "beer" the same child has pronounced}.

The effect of the fonts on the parameters of visual prosody is measured by a one-way ANOVA with repeated measures. ANOVA compares the averages between the different prosodic cues. Tukey's method is used to test the set of all pairwise comparisons {$\mu i - \mu j$} simultaneously.

Pauses are not recognized by speech recognition software. When speaking, most sounds are connected to each other. For this reason, the analysis is based on {the point in time of the latest millisecond of the latest vowel of the word before the space} till {the point in time of the first millisecond of the first vowel of the word after the pause}. Measuring pause this way enables comparisons, even when there is no real pause detected with the speech recognition. It is a useful technique as long as comparisons are made within the data of the same participant.

For tests where the children had to link a presented object with one of n items, proportion tests were used to test if the percentage of how often a given item was selected differed from chance (being 1/n). E.g., if a cue was presented and they had to choose if it indicated a louder, quieter, higher, lower, faster or slower voice, n equals 6.. Proportion tests were performed to test if the percentage of the intended vocalization was significantly larger than 1/6 = 17%.

3. Results

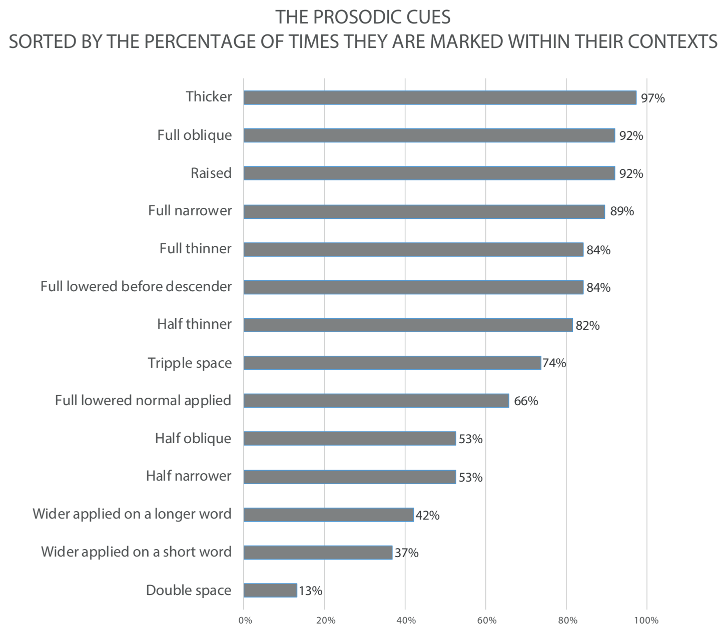
Results for Objective 1: "How intuitive is visual prosody for deaf readers between 8 and 18?"

While evaluating the first sub-question, *"What is the reader's perceived intention of visual prosody?"*, 23 of the 38 participants did not provide an answer related to expressive speech when no information about visual prosody was provided. Only 3 participants provided an answer related to speech expression on a first try, while 12 participants provided an answer with this relationship when they were asked to make a second attempt. In total, 15 out of 38 (39%) of the participants related visual prosody to one or more aspects of expressive speech in a first encounter. Answers not related to expressive speech stated that visual prosody might have the intention to "make things easier," "indicate the verbs," or "lead to better knowledge." Ambiguous answers were evaluated by the researcher during the test to determine if the participant meant an expressive speech. Answers that were deemed correct included "some of those words were read louder," "it influences the intonation," or "that it serves the pronunciation."

While evaluating the second sub-question, *"How perceptible are the prosodic cues?"*, a notable result from this question is that all full versions of prosodic cues were marked more often than subtle versions [Chart 1].

Chart 1.

The prosodic cues on top are marked in a statistically significant number of occurrences.



The answers to the third sub-question, *"How well does each deaf reader relate the prosodic cues to the intended speech variation?"* clarified that most participants would change their voice as intended on five out of seven prosodic cues [indicated by a full border in Chart 2.]. Two cues were more often correlated to unintended speech variations than to their intended speech variations:

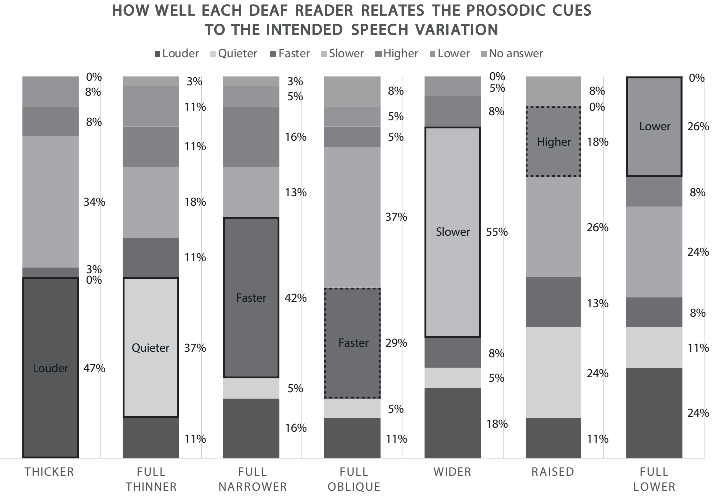
— Raised type, which was mostly connected to slower (26%) or quieter (24%) while higher (18%, proportion test p=0.29 was intended;

— Oblique, which was related to slower (37%, proportion test p=0.02) while faster (29%) was intended.

They were also related to the intended speech variations for a statistically non-significant number of times [indicated by a dashed border in Chart 2.].

Chart 2.

How well each deaf reader relates the prosodic cues to the intended speech variation. A border indicates the intended speech variation. A full border points to a statistically significant result, a dashed border points to a statistically non-significant result.



For the enlarged space, a total of 58 answers were provided, and 30 out of 58 answers (51%, proportion test $p \leq 0.001$) related an enlarged space to waiting longer between words.

A closed question assessed if participants acknowledge the benefit of reading text containing visual prosody. A vast majority of the participants (86%) believed that visual prosody would help them to some degree to read with more expression. Only a minority (14%) expressed that they did not deem visual prosody helpful.

The open feedback pointed to a ratio of positive:negative comments of 2.3:1. Some participants wrote positive as well as negative feedback at the same time. Participants made 53 positive comments in total, including *"Indicating important words makes me pay more attention and clarifies the text," "I would have liked to learn to read with these kinds of booklets," "the voice sounded nicer than normal," "it helped me as deaf person and I think it will help others to read and speak. It probably will help normal hearing individuals as well," "because of the emphasis, the sentence receives much more meaning," "I found reading the sentences aloud a good instruction because I could use and train the voice better with this. It supports you very well if you just learn how to read. I would have preferred learning to read like this"* and *"Yes, it supported me. E.g., faster, slower, louder, quieter—it seemed interesting."* A total of 23 negative comments were provided, of which 13 only indicated that visual prosody was difficult. Two participants related this statement with the parameter that the participant deemed the most difficult: once the quieter voice; once the higher/lower voice. Other comments were *"I sometimes forgot about it"* or *"it is difficult to read with other voices."*

Results for Objective 2: "Does visual prosody increase the variations in the speech features of deaf readers between 8 and 18?"

A total of 38 individuals participated in the test, but due to an unknown microphone error, the vocalization of one participant was not saved. So, 37 participants remained.

A total of 4,995 vowels were expected in the recordings (37 individuals read aloud 45 sentences, and within each sentence, 3 words were selected to compare the effect of the prosodic cues.). The speech recognition software recognized 3,994 words, which is an accuracy of 80%.

A total of 135 words were expected for each participant during analysis. The minimum number of words recognized in the recordings of one participant was 44%, and the maximum number of words detected in the recordings of another participant was 95%. Four out of six cues resulted in statistically significant speech variations as intended when compared with the normal font [indicated by a full border in Table 2].

Table 2.

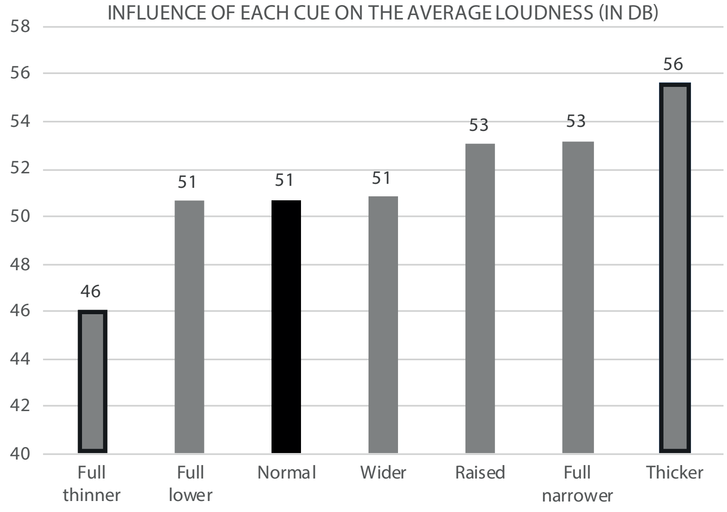
Four out of six cues resulted in a statistically significant intended effect (indicated by the full borders), and two cues resulted in the intended but non-statistically significant effect (indicated by the dashed borders). All other cells contain non-intended effects, which are all smaller than the intended effects. The average of one condition is divided by the average of the normal condition for volume, duration, and pitch. Asterisks (*) indicate a significant difference from the normal font: *=p<0.05; **=p<0.01; ***=p<0.001. The examples are based on an example "neutral condition" of respectively 240Hz, 51 dB, and 0.13sec.

Prosodic cue	Effect on intensity of a vowel	Loudness example (on 51dB)	Effect on duration of a vowel	Duration example (on 0.13sec)	Effect on pitch of a vowel	Pitch example (on 240Hz)
Raised	105% ***	53	145% ***	0.19	127% ***	305
Full lower	100%	51	143% ***	0.19	99%	237
Wider	100%	51	166% ***	0.22	106% ***	254
Full narrower	105% ***	53	98%	0.13	109% ***	262
Thicker	109% ***	56	150% ***	0.19	114% ***	273
Full thinner	90% ***	46	107%	0.14	101%	242
Normal	100%	51	100%	0.13	100%	240

Participants read the prosodic cue intended to read louder ('thicker') with a statistically significant 9% increase in intensity compared with the normal voice, and on average statistically significant louder than all other prosodic cues. The prosodic cue intended to read quieter ('full thinner') is performed with a statistically significant 10% decrease of intensity when compared with the normal voice, and on average quieter than all other prosodic cues [Chart 3.]. Effects of the other cues on the loudness were always smaller than the effect of 'Full thinner' of 'Thicker' and were not always significant [Table 2].

Chart 3.

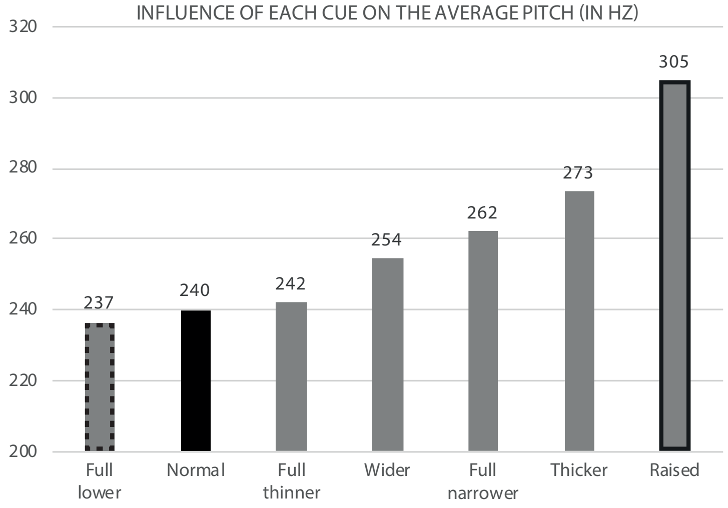
Comparisons of the effects of the different cues on the average loudness of the voice illustrated on an example of 51 dB. The columns represent the average for each cue. The thick borders indicate that the cues intended to influence loudness are in their intended place and had a statistically significant result.



Participants read a prosodic cue intended to read slower ('wider') with a statistically significant 66% increase of the duration of the voice compared with the normal voice, and on average, statistically significantly slower than all other prosodic cues. The prosodic cue intended to read faster ('full narrower') is performed with a 2% decrease in duration when compared with the normal voice, and on average faster than all other prosodic cues. But this prosodic cue does not differ significantly from the normal condition [Chart 4.]. Effects of other cues on the duration were always an increase of the duration, smaller than the effect of 'wider' and not always significant [Table 2].

Chart 4.

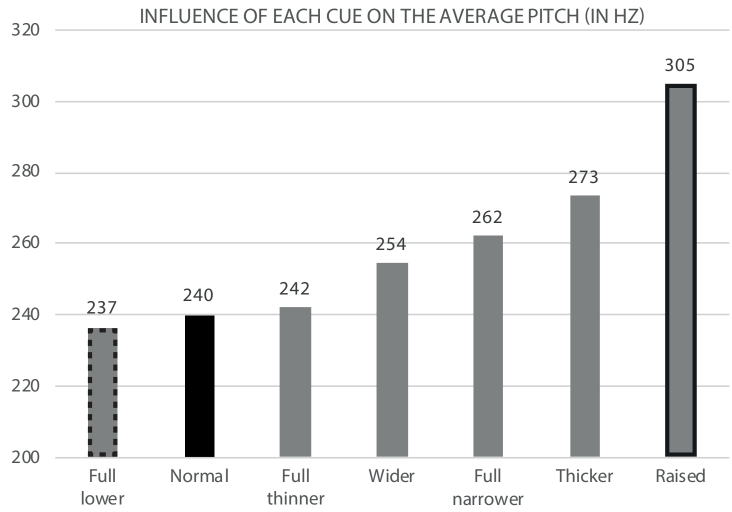
Comparisons of the effects of the different cues on the average duration of the voice illustrated on an example of 0,13sec. The columns represent the average for each cue. The thick border indicates that the cue intended to influence the duration is in its intended place and had a statistically significant result. The dashed border indicates that the cue intended to influence the duration is in its intended place, but the result was not statistically significant.



Participants read a prosodic cue intended to read with a higher voice ('full raised') with a statistically significant 27% higher pitch compared with the normal voice, and on average higher than all other prosodic cues. The prosodic cue intended to read with a lower voice ('full lower') is performed with a 1% lower pitch when compared with the normal voice, and on average lower than all other prosodic cues. But this prosodic cue does not differ significantly from the normal condition [Chart 5.]. Effects of other cues on the duration were always an increase of the pitch, smaller than the effect of 'full raised' and not always significant [Table 2].

Chart 5.

Comparisons of the effects of the different cues on the average pitch of the voice illustrated on an example of 240Hz. The columns represent the average for each cue. The thick border indicates that the cue intended to influence the pitch is in its intended place and had a statistically significant result. The dashed border indicates that the cue intended to influence the pitch is in its intended place, but the results were not statistically significant.



The prosodic cue indicating pause resulted in a pause on average 4.26 times longer pause. This time span is measured independently of how the pause was created: creating a pause by briefly waiting for the next word; creating a pause by breathing in/out between words.

Several prosodic cues had a statistically significant effect on unintended speech variations. Those effects were always smaller than the influence on the intended speech variation. The possible relations between loudness, pitch, and duration are expressed with the Pearson correlation coefficients. All Pearson correlation coefficients remain below 0.28 (intensity-pitch: 0.28 with $p<.0001$; duration-pitch: 0.12 with $p<.0001$ and intensity-duration: 0.16 with $p<.0001$). While a correlation coefficient has an exact mathematical meaning, the interpretation of the magnitude of a correlation coefficient is ambiguous (Kotrlik et al., 2011). However, the various interpretations by different experts describe a correlation coefficient lower than 0.3 as "low," "small," "little if any" (Kotrlik et al., 2011). Therefore, it can be stated that while a cue can have an effect on several speech variations at the same time, the effects on intended and unintended speech variations hardly relate to each other.

Participant's characteristics such as age, type of education, amount of hearing loss, type of hearing device all can influence the vocalization of visual prosody. They were collected for statistical analysis, but most of those analyses delivered no consistent insights. Therefore, it is suggested that the same prosodic cues can be used for all deaf readers between 7 and 18, without differentiation. The only analysis which delivered a certain pattern came from participants in regular education versus those in special education. Participants in special education applied less intensity and duration but more pitch to vocalize the prosodic cues. This can point to a less controlled vocalization and the need for more support by a supervisor.

Results for Objective 3: "Does visual prosody influence the understood meaning of a sentence for deaf readers between 7 and 18?"

A total of 38 booklets were filled in, each containing six sentences, accounting for 228 sentences. In 148 out of 228 sentences (66%), participants marked the intended meaning of the sentences correct. This outcome is statistically significantly higher than the 25% chance level on the correct answer when guessing out of four possible answers (proportion test, $p<0.0001$).

4. Discussion

This research aims to optimize visual prosody for deaf readers between 7 and 18 years old as a support tool to encourage speech variations while reading. The focus was on visualizations for loudness, duration (including pauses), and pitch in both directions: increase and decrease. Therefore, six cues for speech variations and one separate cue for the pause were developed. While this approach to visual prosody led to a successful influence on the reading expressiveness, the cues in their current form cannot be applied without some additional guidance from a supervisor.

The intuitive use of the cues

Readers with hearing loss experience difficulties starting to use those cues. Only 39% of the participants did create the link between visual prosody and a form of prosody/ expression/ speech variations automatically. That is in line with research pointing out that their access to prosody (both perception and production) is lower than for their hearing peers (De Clerck et al., 2018; Øydis, 2014; Wang et al., 2013; Chin, Bergeson & Phan, 2013). Two possible reasons why deaf readers do not automatically create the link between the cues and prosody were found.

The first reason is that at school, children already receive different types of augmented texts to learn about grammar and the structure of the sentence. For example, different colors are used to indicate verbs, nouns, subjects, articles, etc. This explains why some participants related visual prosody to grammar: *"to comprehend text better"* or *"to indicate verbs"* after seeing visual prosody for the first time.

The second reason is that during reading evaluations at school, children are mostly evaluated for speed and accuracy of decoding (Bessemans et al., 2019; ILA, 2018; Mostow & Duong, 2009). Thus, it could be possible that participants focus on the words but not on how those words were stated. This is more valid for deaf children following regular education as they more often participate in reading evaluations. Children within special education do not always participate in reading evaluations. This second reason is also supported by the results of Bessemans et al. (2019), where a group of children receiving no information about speech variations in advance did not read prosodic cues with more expression.

Because not all readers relate visual prosody to expressive reading, it is important that a supervisor is present during the first use of those cues. The supervisor needs to explain the purpose of the prosodic cues before any exercise on speech prosody starts.

The noticeability of the cues

When visual prosody was presented in a text, the wider cue was the only full cue that was not noticeable enough, only marked in 42% of the occurrences. When this cue was applied on a short two-letter word, the number of times marked was even lower, at 37%. This difference supports, but does not yet prove, the hypothesis that the visibility of the wider font might depend on the word on which it is applied. Even more surprisingly, the cue was identical to the research of Bessemans et al. (2019), wherein the same cue was marked in 78% of all occurrences. The findings seem to contradict each other but can be explained by the order of the tests. Bessemans et al. evaluated the noticeability after the reading aloud test, while in this research, the noticeability was evaluated before the reading test. The participants in Bessemans et al. thus already had a training session and thus were more accustomed to all cues.

For future use, additional widening is recommended for the wider cue. That would make noticing this cue easier, for example, when there is no training session beforehand or when there is no supervisor to point to this cue. The current design for the wider font in this study was based on the full-wide version of duration in the study of Bessemans et al. (2019), which was judged the most aesthetically justified variation in letter shapes without disturbing the text color too much in relation to the other full versions and the non-adapted Matilda regular. No adaptations were made to the wider font during this research. However, the typeface Matilda applied for this prosodic parameter a custom spacing system: the white space on the left and right of a letter is of similar width in each font. While it is more common in type design to provide each font a different spacing, this choice was made for research purposes: the design parameter 'letter' was taken into account and not the 'letter-spacing' parameter (in that sense, they were kept constant) to precisely point out the effect to wider letters (and not possible interaction effects with wider letter spacing). To ensure that future visual prosody users will notice the wider cue more readily, it is suggested to follow the standard spacing system and to insert more letter spacing. More letter spacing can increase the noticeability as the letters have more distance between each other.

The low noticeability seems to contradict the reading aloud test wherein this cue resulted in a 66% longer duration of the vocalization. The large influence on the reading aloud is explained by the repetition of the cues just before the reading aloud test. During that training session, participants were made aware of the cues within the text until they noticed them. Therefore, the participants did not yet know what to look for while we tested the noticeability, but participants gradually got more accustomed to the prosodic cues throughout the whole research.

The intuitive relation between the cues and speech variations

Once readers with hearing loss understand the relation between the prosodic cues and the speech variations, they intuitively relate five cues correctly to the intended speech variation. And once the relation between each cue and its intended speech variation is explained, participants found that relationships easy to remember, for example, a raised font with a higher pitch. The obvious relationships are in line with Niebuhr et al. (2017), who state that iconic visualizations (visual representations of speech variations) are more intuitive than symbolic visualizations (prosody indicated by symbols added to the text).

Based on the literature that reviews a lesser pitch perception by individuals with hearing loss (Svirsky, 2017; Marx et al., 2014; Perreau, Tyler & Witt, 2010), it was expected that the cues for a higher and lower pitch would be difficult for deaf readers to relate to their intended speech variation. They were indeed the two cues related the least often to the intended speech variation. The explanation could be that pitch perception is more difficult for deaf individuals (Limb & Roy, 2014), and they have a more limited understanding of pitch.

Unexpected was that the bold cue was related to a louder vocalization in only 47% of all occurrences. This cue was expected to be the best related to the intended speech variation in this test because thicker fonts are often related to volume (Lewis & Walker, 1989) and are widely applied in comics and reading books to express volume. It was therefore expected that this cue would be the easiest to interpret. The current research cannot explain why the percentage was relatively low.

Rather unexpected was that the experimental oblique font was more related to a slower vocalization than to a faster vocalization. Because the oblique font was not related to its intended speech variation, it was not studied in the reading aloud tests of the current research and was put aside for possible future studies. The correlation of an oblique cue with slower reading could originate in the common application in reading materials, where italics are often applied to highlight important parts. A later follow-up study can evaluate if deaf readers perceive oblique/italic text as more important and to be read with more attention (thus slower). Based on such future studies, new guidelines for reading italic/oblique fonts can be formulated.

An explanation is needed to correct the readers who confuse the cues and their intended speech variation. Because deaf readers can have less knowledge about prosody (and speech variations), the presence of a supervisor is recommended. This supervisor can correct the reader and fill in possible gaps in his knowledge about speech variations.

The influence of visual prosody on the reading aloud

Deaf readers read aloud the cues for a louder, quieter, slower, and higher vocalization, plus the pause, as intended. Therefore, these cues can support their expressive reading and can be used in speech therapy to train their speech expressiveness, as several organizations already do on an experimental and pragmatic basis with their custom cues (KIDS, n.d.; Advanced bionics, unpublished; Staum, 1987; van Uden, 1973). The two cues which did not result in a statistically significant and intended speech variation were the narrower font and the lower font. The correlating speech variations to these cues (faster speed and lower pitch) are deemed much more challenging to produce. The difficulties performing a faster or lower vocalization and the comments on the narrower font do not diminish the relevance of those two cues: although not all children will (be able to) perform the related speech variations, their presence is useful to start a discussion about vocal speed and pitch.

In general, pitch is a difficult speech component to attain. At the beginning of the reading test, several deaf readers needed some extra exercises on this speech component. More than once, readers moved their whole body upwards when producing a higher voice; the notions "higher" and "lower" are more often used to indicate objects within a person's spatial environment. A bodily motion was literally mentioned by one participant in this study and also noticed in Bessemans et al. (2019).

A lower-pitched vocalization could be difficult to execute because the regular speaking voice already sounds low and is close to the lowest limit in a voice's pitch range [Table 3.] (Meijer, 2015; De Bodt et al., 2015). Further, it is known that technological limitations constrain the pitch perception through cochlear implants (Limb & Roy, 2014). That hinders to perceive prosodic speech information accurately (Kalathottukaren, Purdy & Ballard, 2017). Both reasons could have contributed to the fact that the results of the prosodic cue intended to read with a lower voice is statistically non-significant.

Table 3.

The average pitch during speech is already close to the lowest pitch that a voice can produce. The values can differ slightly for individual measurements. See, for example, Anderson (1977) or Benninger and Murry (2008). Note that the maximal pitch mentioned here is taken from the singing voice, which can reach a higher limit than the speaking voice, but illustrates the voice's full pitch range. The table is a simplified version of De Bodt et al. (2015, citing Mathieson, 2001).

	Voice type	Vocal range	Average pitch in speech
Man	Bass	82-333 Hz	98 Hz
	Baritone	98-392 Hz	124 Hz
	Tenor	131-523 Hz	165 Hz
Woman	Alto	147-587 Hz	175 Hz
	Mezzo-soprano	165-880 Hz	196 Hz
	Soprano	196-1174 Hz	247 Hz

The speech variation on some cues was more than once exaggerated. For the higher pitch, a change in vocal register often occurred. A switch to a higher vocal register allows the voice to reach higher pitch values but is not common in daily speech: it can cause the vocalization to sound forced. When visual prosody is applied in speech training, a supervisor will need to indicate when the pitch is going too high.

Each prosodic cue had an effect on unintended speech components. For example, the prosodic cue intended to result in a higher pitch not only caused a higher pitch, but also significantly increased both loudness and duration. Bessemans et al. (2019) and Patel, Kember & Natale (2014) noticed the same side effect. The cause needs to be sought in the human anatomy: all human anatomical motion tends to cooperate in order to create speech. An extra effort in one body part influences the achievements of the other parts as well. It is known that pitch rises exponentially if intensity increases (Buekers & Kinsma, 2005). Although some cues gave an unintended increase in loudness/duration/pitch, the intended effect of the cue was always significantly larger. The unintended effects, therefore, do not diminish the positive outcome of this research.

Visual prosody influencing the understood meaning of a sentence

The influence of visual prosody on understanding the meaning of the sentence is strong: without any explanation, 66% of the participants marked the intended answer that correlated with the emphasized word. The result is expected to improve even more when participants get more acquainted with visual prosody or receive more explanation in advance.

Why (visual) emphasis in written sentences correlates with emphasis in speech (while reading silent) is not to be determined within this research. Gross et al. (2013) express the same caution. We propose two possible reasons: an auditory one and a visual one. The first possible reason might be that an inner voice is active during reading and that the emphasized word triggers speech prosody. The second possible reason might be that participants isolate the visually stressed word and base their answer on the meaning of this one word separately. Whatever the reason may be, because the influence is the same in spoken and written sentences, visual prosody can be applied to achieve a correct interpretation of a sentence and thus to discuss why and where speech prosody should be applied to create emphasis.

The perception of visual prosody

Visual prosody is perceived as useful by the readers, and where some participants expressed difficulties in handling the cues, more exposure and more extended training will automatically result in habituation. That should encourage developers of reading/learning materials to adopt visual prosody. The positive reaction of parents and speech therapists lies in line with the perception of the deaf readers. Speech therapists found prosody an aspect of speech that deserved more attention. One parent of a deaf child having more than average difficulties with developing speech responded: *"In his way [my son] was so enthusiastic and so proud of his certificate [that he received after participating in this research], immediately after arriving home he was overjoyed and wanted to display the cues. I think I could sense he had understood the reading program. He explained it to me completely, which is quite remarkable for him. It clearly made a good impression on* *him."* (literal translation of communication with a parent, 2018). The positive comments are in line with five other studies that mention a positive attitude of the participants towards visual prosody (Bessemans et al., 2019; Patel, Kember & Natale, 2014; Patel & McNab, 2010; Argyropoulos et al., 2009).

In this study, there could have been an under-representation of low achievers in reading. In this research, only a couple of participants had signs of what could be severe to problematic speech development delays. This study is not the first with this conclusion: Mayer et al. (2016) also concluded that their study did not completely represent the heterogeneous group of deaf. As mentioned in Sininger, Grimes & Christensen (2010) and Holly (1997), low achievers do not always participate in research. While it does not diminish the research results, supervisors need to evaluate first where possible problems with speech will occur before commencing speech training.

We close the discussion with a prospect on possible future research. Prosody, as the motor of expressive speech, is part of fluent reading (according to the definition of the National Reading Panel in NIH, 2000). Fluent reading is important because those readers *"processed the text smoothly, identify and understand words easily, efficiently and rapidly, discern syntax, and focus on the meaning"* (Luckner & Urbach, 2011). Fluent reading is a part of the reading process wherein deaf readers generally develop slower than their hearing peers (Mayer et al., 2016; Luckner & Urlbach, 2011). Even with CI's, they still do not reach the same level as their hearing peers (Boons et al., 2013; Mayberry, 2002; Vermeulen et al., 2007). Future research needs to determine the full effects of visual prosody; whether it supports fluent reading in general, and if so, how much visual prosody is able to support reading comprehension.

5. Conclusion

This study confirms the hypothesis, *"Visual prosody leads to more vocal prosody while reading aloud, and influences reading comprehension of deaf readers between 7 and 18."* 'Deaf readers' in this study refers to readers who have developed spoken language that is distinct enough to be understood. For this audience, the approach to visual prosody used in this study is successful in creating more speech variations, and thus a more expressive voice. Therefore, visual prosody can be used not only in reading materials aimed at expressive reading but also in speech therapy to learn about speech variations or to train the prosody of deaf readers.

Typographers, type designers, graphic designers, teachers, speech therapists, and researchers who are developing reading materials intended to support expressive speech could use this study as a base when developing materials supporting expressive reading by relating loudness to the thickness (blackness) of a font; duration to the width of a font; pitch to the vertical position of a font; and a pause to a wider space. The example cues in this article illustrate a good starting point for further development. During further developments, some of the suggested improvements for the noticeability can be implemented, such as an even wider cue to read slower.

It is important to remember that expressive reading cannot be achieved only by visual prosody. Prosodic cues are not intuitive enough to be handled by a reader alone, and some readers lack the necessary knowledge about speech variations. If a reader would start using those cues without supervision, some cues will be read without a change in vocalization, or some other errors will be made in the vocalization. Therefore, a supervisor is needed to guide the reader through the process and to provide corrections where needed. But once the intention of visual prosody is clear, readers seem to handle the cues well.

6. Acknowledgements

Thanks to all schools in the Dutch-speaking regions that supported this research and facilitated connecting with deaf students. In alphabetical order: Antwerp Plus (Antwerp), BUSO Zonnebos (Schoten), Cor Emous (The Hague), Kasterlinden (Kasterlinden), KIDS (Hasselt), Koninklijk Instituut Woluwe (Woluwe), Sint Gregorius (Ghent-Bruges), Sint-Lievenspoort (Ghent), Spermalie (Bruges). Without their help, this research would not have been possible.

7. References

ADVANCED BIONICS. (unpublished version). *A musical journey through the rainforest.*

ANDERSON, VA (1977). *Training the Speaking Voice.* New York: Oxford University Press.

ARGYROPOULOS, V.; SIDERIDIS, G.; KOUROUPETROGLOU, G.; & XYDAS, G. (2009). "Auditory discriminations of typographic attributes of documents by students with blindness." *The British Journal of visual impairment.* 27(3): 183–203. http://dx.doi.org/10.1177/0264619609106360

ASHBY, J. (2006). "Prosody in skilled silent reading: evidence from eye movements." *Journal of Research in Reading.* 29 (3) 318–333. http://dx.doi.org/10.1111/j.1467-9817.2006.00311.x

ASP, CW (2006). Verbotonal Speech Treatment. San Diego: Plural Publishing.

BAUDONCK, N.; VAN LIERDE, K.; D'HAESELEER, E.; & DHOOGE, I. (2015). "Nasalance and nasality in children with cochlear implants and children with hearing aids." *International Journal of Pediatric Otorhinolaryngology.* 79 (2015): 541–545.

BENNINGER, MS; & MURRY, T. (2008). *The Singer's Voice.* Abington: Plural Publishing.

BESSEMANS, A.; RENCKENS, M.; BORMANS, K.; NUYTS, E; LARSON, K. (2019). "How to visualize prosody in order to help children read aloud with more expression." *Visible Language.* 70(3): 28-49.

BELYK, M. & BROWN, S. (2014). "Perception of affective and linguistic prosody: an ALE meta-analysis of neuroimaging studies." *Social Cognitive and Affective Neuroscience.* 9 (9): 1395-1403.

BINDER, KS; TIGHE, E.; JIANG, Y., KAFTANSKI, K.; QIC.; ARDOIN, SP (2012). "Reading expressively and understanding thoroughly - An examination of prosody in adults with low literacy skills." *Reading and Writing.* 26 (5): 665–680. http://dx.doi.org/10.1007/s11145-012-9382-7

BOERSMA, P., WEENINK, D. (2014). *Praat: doing phonetics by computer* [Online] http://www.fon.hum.uva.nl/praat [January 7th, 2015].

BOONEN, N.; KLOOTS, H.; VERHOEVEN, J.; & GILLIS, S. (2017). "Can listeners hear the difference between children with normal hearing and children with a hearing impairment?" *Clinical Linguistics & Phonetics.* 33 (4): 316-333. https://doi.org/10.1080/02699206.2018.1513564

BOONS, T.; BROKX, J.; FRIJS, J.; PHILIPS, B. VERMEULEN, A.; WOUTERS, J.; VAN WIERINGEN, A. (2013c) "Newborn hearing screening and cochlear implantation: Impact on spoken language development." *B-ENT.* 9 (Suppl. 21): 91-98.

BOONS, T.; DE RAEVE, L.; LANGEREIS, M.; PEERAER, L.; WOUTERS, J.; VAN WIERINGEN, A. (2013b). "Narrative spoken language skills in severely hearing impaired school-aged children with cochlear implants." *Research in Developmental Disabilities*. 34: 3833–3846. http://dx.doi.org/10.1016/j.ridd.2013.07.033

BOONS, T.; DE RAEVE, L.; LANGEREIS, M.; PEERAER, L.; WOUTERS, J.; VAN WIERINGEN, A. (2013a). Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation. *Research in Developmental Disabilities.* 34: 2008–2022.

BREEN, M.; KASWER, L.; VAN DYKE, JA; KRIVOKAPIĆ, J.; LANDI, N. (2016). "Imitated prosodic fluency predicts reading comprehension ability in good and poor high school readers." *Frontiers in Psychology.* 7 (1026). http://dx.doi.org/10.3389/fpsyg.2016.01026

BULKERS, R.; & KINGSMA, H. (2009). "Impact of phonation intensity upon pitch during speaking: A quantitative study in normal subjects." *Logopedics Phoniatrics Vocology*. 22 (2): 71-77. http://dx.doi.org/10.3109/14015439709075317

CARLSON, K. (2009). "How Prosody Influences Sentence Comprehension." *Language and Linguistics Compass.* 3 (5): 1188–1200.

CHAN, H. (2018). "A method of prosodic assessment: Insights from a singing workshop." *Cogent education*. 5 https://doi.org/10.1080/2331186X.2018.1461047

CHIN, SB; BERGESON, TR; PHAN, J. (2012). "Speech Intelligibility and Prosody Production in Children with Cochlear Implants." *Journal of Communication Disorders.* 45 (5). 355–366. http://dx.doi.org/10.1016/j.jcomdis.2012.05.003.

DAILY MAIL ONLINE. (May 23rd, 2014). The world through a deaf persons ears Video reveals what its like to listen to sound using a cochlear implant [Online] https://www.dailymail.co.uk/sciencetech/article-2636415/What-deaf-hear-Audio-file-reveals-s-like-listen-world-using-cochlear-implant.html [March 19th, 2019]

DE BODT, M.; HEYLEN, L.; MERTENS, F.; VANDERWEGEN, J.; VAN DE HEYNING, P. (2015). Stemstoornissen. Handboek voor de klinische praktijk (Speaking disorders. Manual for the clinical practice). Antwerpen: Garant.

DE CLERK, I; PETTINATO, M.; GILLIS, S.; VERHOEVEN, J.; & GILLIS, S. (2018). "Prosodic modulation in the babble of cochlear implanted and normally hearing infants: a perceptual study using a visual analogue scale." *First Language.* 38 (5): 481–502. https://doi.org/10.1177/0142723718773957

DE RAEVE, L. (2014). Paediatric Cochlear Implantation: outcomes and current trends in education and rehabilitation [Dissertation] Radboud University.

DE RAEVE, L.; BAERTS, J.; COLLEYE, E.; & CROUX, E. (2012). "Changing Schools for the Deaf - Updating the Educational Setting for Our Deaf Children in the 21st Century, a Big Challenge." *Deafness & education international.* 14 (1): 48–59. https://doi.org/10.1179/1557069X11Y.0000000012

DILLON, H.; COWAN, R.; & CHING, TY (2013). "Longitudinal outcomes of children with hearing impairment (LOCHI)." *International Journal of Audiology.* 52, (Suppl 2): S2-3. https://doi.org/10.3109/14992027.2013.866448

DOOF.NL. (2017). Update erkenning Nederlandse Gebarentaal (Update on the acknowledgment of the Dutch Sign Language). [Online] https://www.doof.nl/samenleving-maatschappij/update-erkenning-nederlandse-gebarentaal [1 November 2018].

FAGAN, MK, PISONI, DB (2010). Hearing experience and receptive vocabulary development in deaf children with cochlear implants. Journal of Deaf Studies and Deaf Education. 15(2): 149-61. https://doi.org/10.1093/deafed/enq001

FEVLADO (2013). Visietekst dovenonderwijs (vision tekst on deaf education). [Online document] http://fevlado.be [21 March 2018].

FODOR, J. D. (1998). "Learning to parse?" *Journal of Psycholinguistic Research*. 27 (2): 285–319.

GROEN, MA; VEENENDAAL, NJ; VERHOEVEN, L. (2019). "The role of prosody in reading comprehension: evidence from poor comprehenders." *Journal of Research in Reading.* 42 (1): 37–57. ISSN 0141-0423 https://doi.org/10.1111/1467-9817.12133

GROSS, J.; MILLETT, AL; BARTEK, B.; BREDELL, KH; WINEGARD, B. (2013). "Evidence for prosody in silent reading." *Reading Research Quarterly.* 49 (2): 189–208. https://doi.org/10.1002/rrq.67

GUBERINA, P. & ASP, C. W. (1981). The Verbo-tonal Method for rehabilitation people with communication problems. [Online] http://www.suvag.com/ang/histoire/autrestextes.html [28 February 2018]

HARRIS, M. (2015). The Impact of Cochlear Implants on Deaf Children's Literacy. In: MARSCHARK, M. & SPENCER, E. (2015). *The Oxford Handbook of Deaf Studies in Language.* New York: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780190241414.013.27

HASBROUCK, J.; & GLASER, DR (2012). *Reading fluency: Understanding and teaching this complex skill.* Austin, TX: Gibson Hasbrouck & Associates.

HEARING TEAM FIRST. (2017). Start with the brain and connect the dots - supporting Children who are deaf or hard of hearing to develop literacy through listening and spoken language. [Online pdf] https://hearingfirst.org/blog/2017/08/17/Connecting-the-Dots [10 April 2019].

HOLLY, F-B. (1997). "Cochlear Implant Use by Prelingually Deafened Children: The Influences of Age at Implant and Length of Device Use." *Journal of Speech, Language, and Hearing Research.* 40(1): 183-199. ISSN: 1092-4388.

HUTTER, E.; ARGSTATTER, H.; GRAPP, M.; Plinkert, P.K. (2015). "Music therapy as specific and complementary training for adults after cochlear implantation: A pilot study." *Cochlear Implants International.* 16 (S3): S13.

INTERNATIONAL LITERACY ASSOCIATION (ILA). (2018). *Reading fluently does not mean reading fast.* [Online pdf] https://literacyworldwide.org/statements [13 August 2019].

KALATHOTTUKAREN, RT; PURDY, SC & BALLARD, E. (2017). "Prosody perception and musical pitch discrimination in adults using cochlear implants." *International Journal of Audiology.* 54 (7): 444-452.

KARPIŃSKI, M. (2012). "The boundaries of language: dealing with paralinguistic features." *Lingua Posnaniensis.* 54 (2): 37-54. ISSN 0079-4740, ISBN 978-83-7654-252-2.

KIDS (Unpublished). Teaching materials for children with language disorders. [Internal materials]

KOTRLIK, J.; WILLIAMS, H.; & JABOR, M. (2011). "Reporting and Interpreting Effect Size in Quantitative Agricultural Education Research." *Journal of Agricultural Education.* 52(1): 132–142.

KUHN, MR; & STAHL, SA (2003). "Fluency - A review of developmental and remedial practices." *Journal of Educational Psychology*. 95 (1): 3–21. https://doi.org/10.1037/0022-0663.95.1.3.

LEINENGER, M. (2015). "Phonological coding during reading." *Psychological Bulletin*. 140 (6): 1534-1555.

LEWIS, C., & WALKER, P. (1989). "Typographic influences on reading." *British Journal of Psychology.* 80, (2), 241-257.

LIMB, C. J., & ROY, A. T. (2014). "Technological, biological, and acoustical constraints to music perception in cochlear implant users." *Hearing Research.* 308: 13–26. https://doi.org/10.1016/j.heares.2013.04.009.

LUCKNER, JL; & URBACH, J. (2011). Reading Fluency and Students Who Are Deaf or Hard of Hearing: Synthesis of the Research. Communication Disorders Quarterly: 2011. https://doi.org/10.1177/1525740111412582.

LYXELL, B.; WASS, M.; SAHLÉN, B.; SAMUELSSON, C.; ASKER-ÁRNASON, L.; IBERTSSON, T.; MÄKI-TORKKO, E.; LARSBY, B.; & HÄLLGREN, M. (2009). "Cognitive development, reading and prosodic skills in children with cochlear implants." *Scandinavian Journal of Psychology.* 50: 463–474.

MARKIDES, A. (1983). *The speech of hearing-impaired children.* Manchester University Press, New Hampshire, USA.

MARX, M.; JAMES, C.; FOXTON, J.; CAPBER, A. FRAYSSE, B.; BARONE, P. DEGUINE, O. (2014). "Speech prosody perception in cochlear implant users with and without residual hearing." *Ear & Hearing.* 36 (2): 239–248. http://doi.org/10.1097/AUD.0000000000000105

MAYBERRY, RI (2002). "Cognitive development in deaf children - the interface of language and perception in neuropsychology." *Handbook of Neuropsychology, 2nd Edition*. Vol. 8, Part II.

MAYER, C.; WATSON, L.; ARCHBOLD, S.; YEN, Z.NG. MULLA, I. (2016). "Reading and Writing Skills of Deaf Pupils with Cochlear Implants." *Deafness & Education International.* 18 (2): 71-86. http://dx.doi.org/10.1080/14643154.2016.1155346

MEIJER, A. (2015) Kopstem en Spreekstem. [Online] https://muziekschool.nl/kopstem-en-spreekstem [16 April 2019].

MILLER, J.; & SCHWANENFLUGEL, PJ (2006). "Prosody of syntactically complex sentences in the oral reading of young children." *Journal of Educational Psychology.* 98 (4): 839-843.

NAKATA, T.; TREHUB, SE; & KANDA, Y. (2011). Effect of cochlear implants on children's perception and production of speech prosody. *J. Acoust. Soc. Am.* 131: 2.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP). (1995, August). Listening to children read aloud: Oral fluency. NAEPFacts, 1, (1): 2-5.

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, DC, USA: US Government Printing Office.

NIEBUHR, O.; ALM, MH; SCHÜMCHEN, N.; & FISCHER, K. (2017). "Comparing visualization techniques for learning second language prosody: first results." *International Journal of Learner Corpus Research.* 3(2): 250-277. https://doi.org/10.1075/ijlcr.3.2.07nie

ØYDIS, H. (2013). Acoustic Features of Speech by Young Cochlear Implant Users. [Dissertation] Universiteit Antwerpen.

PAIGE, DD; RASINSKI, TV; MAGPURI-LAVELL, T. (2012). "Is fluent, expressive reading important for high school readers?" *Journal of Adolescent & Adult Literacy*. 56 (1): 67–76. https://doi.org/10.1002/JAAL.00103

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.

PATEL, R. & MCNAB, C. (2010). "Feasibility of augmenting text with visual prosodic cues to enhance oral reading." *Speech Communication.* 53 (3): 431-441.

PATEL, R., KEMBER, H. & NATALE, E. (2014). "Feasibility of Augmenting Text with Visual Prosodic Cues to Enhance Oral Reading." *Speech Communication.* 65. https://doi.org/10.1016/j.specom.2014.07.002

PENG, Y.; HSU, M-W.; TAELE, P.; LIN, T-Y.; LAI, P-E., HSU, L.; CHEN, T-C.; WU, T-Y.; CHEN, Y-A.; TANG, H-H.; CHEN, MY (2018). SpeechBubbles: Enhancing Captioning Experiences for Deaf and Hard-of-Hearing People in Group Conversations. [Conference paper]. *CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems Paper No. 293.* CHI 2018, Montréal, QC, Canada. https://doi.org/10.1145/3173574.3173867

PENG, SC; TOMBLIN, JB; & TURNER, CW (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. Ear Hear. 29 (3): 336-51. http://doi.org/10.1097/AUD.0b013e318168d94d

PERREAU, A.; TYLER, RS; WITT, SA (2010). "The Effect of Reducing the Number of Electrodes on Spatial Hearing Tasks for Bilateral Cochlear Implant Recipients." *J Am Acad Audiol.* 21 (2): 110–120.

READING ROCKETS. (2019). Fluency. [Online] http://www.readingrockets.org/teaching/reading-basics/fluency [13 January 2019].

RENCKENS, M.; & VANMONTFORT, W. (2015a). *Adjustments to Praat.* [Application] https://github.com/READSEARCH/praat [21 July 2020].

RENCKENS, M.; & VANMONTFORT, W. (2015b). *Praat plugin: export time-intensity-pitch.* [Script]. https://github.com/READSEARCH/Praat_plugin_export_time-intensity-pitch [21 July 2020].

SCARBEL, L.; VILAIN, A.; LOEVENBRUCK, H.; SCHMERBER, S. (2012). "An acoustic study of speech production by French children wearing cochlear implants." *3rd Early Language Acquisition Conference.* Dec 2012, Lyon, France.

SHAIKH, D. (2009). "Know your typefaces! Semantic differential presentation of 40 onscreen typefaces." *Usability News. 11,* (2).

SEE, RL; DRISCOLL, VD; GFELLER, K.; KLIETHERMES, S. & OLESON, J. (2013). "Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing." *Otology and Neurotology.* 34 (3): 490–8. https://doi.org/10.1097/MAO.0b013e318287c985.

SEIDENBERG, M. (2017). *Language at the speed of sight.* New York: Basic Books.

SININGER, YS; GRIMES, A.; & CHRISTENSEN, E. (2010). "Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss." *Ear & Hearing.* 31 (2): 166-185.

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.

SOMAN, UG (2017). *Characterizing Perception of Prosody in Children with Hearing Loss.* [Dissertation]. Tennessee: Graduate School of Vanderbilt University.

STAUM, MJ (1987). "Music Notation to Improve the Speech Prosody of Hearing Impaired Children." *Journal of Music Therapy.* 24 (3): 146-159.

STILES, DJ; & NADLER, LJ (2013). "Sarcasm Recognition in Children with Hearing Loss - The Role of Context and Intonation." *Journal of Educational Audiology.* 19: 3-11.

SVIRSKY, M. (2017). "Cochlear implants and electronic hearing." *Physics Today.* 70 (8): 52-58. https://doi.org/10.1063/PT.3.3661

VANDER BEKEN, K.; DEVRIENDT, V.; VAN DEN WEYNGAERD, R.; DE RAEVE, L.; LIPPENS, K.; BOGAERTS, J.; MOERMAN, D. (2010). *Personen met een auditieve handicap (Persons with an auditory limitation).* In: BROEKAERT et al. *Handboek bijzondere orthopedagogiek (Manual for extraordinary remedial education).* (2016: fourteenth press) Garant. Antwerpen-Apeldoorn. p 131-210.

VANHERCK, C.; VUEGEN, D. (2009). Prosodie bij kinderen met een cochleair implant - verkenning van de vaardigheden en objectivering met behulp van het prosogram. [Online] https://exporl.med.kuleuven.be/web/index.php/Public:Masterproefverdedigingen/2009/Prosodie_bij_kinderen_met_een_cochleair_inplant:_verkenning_van_de_vaardigheden_en_objectivering_met_behulp_van_het_prosogram [9 April 2017].

VAN UDEN, A. (1973). Taalverwerving door taalarme kinderen (Language acquisition by language-poor children). Universitaire Pers Rotterdam, Rotterdam.

VEENENDAAL, NJ; GROEN, MA; VERHOEVEN, L. (2014). "The role of speech prosody and text reading prosody in children's reading comprehension." *British Journal of Educational Psychology.* 2014 (84): 521–536.

VERMEULEN, A.M.; VAN BON, W.; & SCHREUDER, R. (2007). Reading Comprehension of Deaf Children With Cochlear Implants. Journal of Deaf Studies and Deaf Education 12(3): 283-302.

VERSTRAETE, E. (1999). De stilte verbroken – Diagnostiek en revalidatie van personen met een auditieve handicap: Deel 1: Theoretische basis (The silence broken – diagnosis and revalidation of persons with an auditory limitation: part 1: theoretical basics). Vormingsdienst SIG. Destelbergen, Belgium.

WANG, DJ; TREHUB, SE; VOLKOVA, A.; VAN LIESHOUT, P. (2013). "Child implant users' imitation of happy- and sad-sounding speech." *Frontiers in Psychology.* 4 (351). http://dx.doi.org/10.3389/fpsyg.2013.00351.

WAGNER, M. & WATSON, D. (2010). "Experimental and theoretical advances in prosody: A review." *Language and Cognitive Processes.* 25 (7-9): 905–945. http://dx.doi.org/10.1080/01690961003589492.

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. ISSN: ISSN00340553.

Authors

Maarten Renckens

Maarten Renckens is a teacher and design researcher with a love for letters and a heart for people. Dealing with a reading difficulty himself, he is very interested in the reading process. His projects include the typeface 'Schrijfmethode Bosch' (Writing Method Bosch) that learns children how to write and typefaces to encourage beginner readers and readers with hearing loss to read more expressively. With a background in architectural engineering, he is used to approach concepts technically and mathematically. He applies this technical knowledge to unravel letterforms, in order to determine the effects of different letterforms on the reading process.

Leo De Raeve

Leo De Raeve PhD has 3 professions: he is a Doctor in Medical Sciences, psychologist and teacher of the deaf. He is founding director of ONICI, the Independent Information and Research Center on Cochlear Implants, is lecturer at University College Leuven-Limburg and scientific advisor of the European Users Association of Cochlear Implant (EURO-CIU).

Erik Nuyts

Prof. Dr. Erik Nuyts is researcher and lecturer at the University College PXL and associate professor at the University of Hasselt. He got a master degree in mathematics, and afterwards a PhD in biology.

Since his specialty is research methodology and analysis, his working area is not limited to one specific field. His experiences in research, therefore, vary from mathematics to biology, traffic engineering and credit risks, health, physical education, (interior) architecture and typography.

His responsibilities both at the University College PXL as at the University of Hasselt involve preparation of research methodology, data collection and statistical analyses in many different projects. He is responsible for courses in research design, statistics, and mathematics.

María Pérez Mena

Dr. María Pérez Mena is an award-winning graphic and type designer. She is postdoctoral researcher at the legibility research group READSEARCH at PXL-MAD School of Arts and Hasselt University. María teaches typography and type design in the BA in Graphic Design at PXL-MAD and is lecturer in the International Master program 'Reading Type & Typography' and the Master program 'Graphic Design' at the same institution. She received her PhD "with the highest distinction" from University of Basque Country and is a member of the Data Science Institute UHasselt.

Ann Bessemans

Prof. Dr. Ann Bessemans is a legibility expert and award-winning graphic and type designer. She founded the READSEARCH legibility research group at the PXL-MAD School of Arts and Hasselt University where she teaches typography and type design. Ann is the program director of the international Master program 'Reading Type & Typography'. Ann received her PhD from Leiden University and Hasselt University under the supervision of Prof. Dr. Gerard Unger. She is a member of the Data Science Institute UHasselt, the Young Academy of Belgium and lecturer at the Plantin Institute of Typography.

a higher voice

A preliminary study exploring the relation between visual prosody and the prosodic components in sign language

Maarten Renckens[1]
Leo De Raeve[2, 3]
Erik Nuyts[1]
María Pérez Mena[1]
Ann Bessemans[1]

1 Hasselt University in collaboration with PXL University College, Belgium;
2 Independent Information Centre about Cochlear Implantation (ONICI) Zonhoven, Belgium;
3 KIDS (Royal Institute for the Deaf) Hasselt, Belgium.

Type enriched with visual prosody is a powerful tool to encourage expressive reading. Visual prosody adds cues to text to guide vocal variations in loudness, duration, and pitch. More vocal variations result in a less monotonous voice and thus more expression. A positive effect of visual prosody is known on the voice of normal hearing readers and of signed bilingual deaf readers who developed signed language and spoken language. These deaf readers rely on speech as well as sign language and both modalities can be used interchangeably to compensate each other.

This preliminary study explores visual prosody in text in relation to Flemish Sign Language to see if sign language can be used to explain prosody. We asked deaf readers between 7 and 18 to relate prosodic cues to videos presenting prosodic components of Flemish Sign Language. We found that those readers connect the prosodic cues with the components in Flemish Sign Language as intended. Larger word-spacing correlates with a pause between signs, a wider font with a sign with 'longer duration', a thicker font with more 'displacement' in the sign, a raised font with a 'faster velocity' in the sign. However, some confusion occurred as participants seemed to extract only two prosodic components in the sign language: both the 'faster velocity' and 'longer duration' were referred to in terms of 'speed' and were not perceived as separate prosodic components. Participants were confused about why there were three cues in the text. Therefore, it is advised to re-evaluate and to re-design visual prosody for sign language with only 'displacement' and 'speed' in mind.

Keywords –
type design,
visual prosody,
prosody,
deaf readers,
sign language,
expressiveness

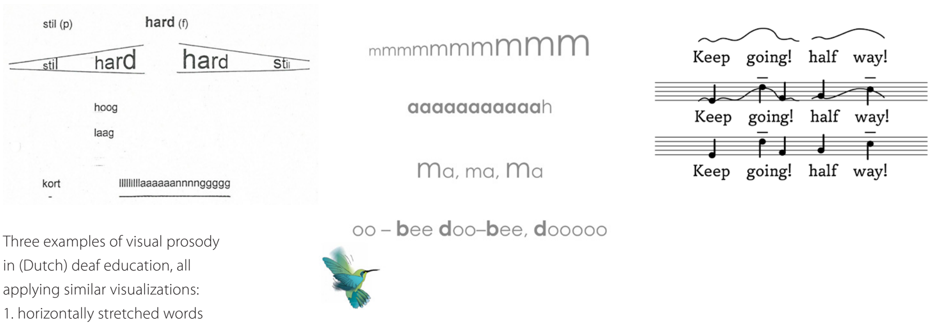
Introduction

Visual prosody visualizes information that is otherwise mostly absent in text: information about 'how words are said.' That is the prosody, variations in the loudness, duration, and pitch of the voice. Prosody is the motor of expressive speech and plays an important role in understanding language. It distinguishes words such as PREsent versus preSENT, and it adds additional information: emphasis, statements, questions, sarcasm, emotion, and more. Additionally, it can influence the meaning of a sentence: for example, "That old man CANNOT hear you very well" or "That old man cannot hear YOU very well." Visual prosody adds prosody to text in a visual way using prosodic cues. There exist several different approaches to visual prosody, both for reading a text with more speech variations and for reading comprehension (Renckens et al., 2021; Bessemans et al., 2019; Rude, 2016; Patel & McNab, 2010).

The perception and production of prosody of deaf readers are not equal to that of their hearing peers (De Clerck et al., 2018; Øydis, 2014; Boons et al., 2013; See et al., 2013; Chin, Bergeson & Phan, 2013; Vander Beken et al., 2010; Markides, 1983). In our study, 'deaf readers' refers to 'deaf students who have developed spoken language (legible enough to be understood) and Flemish Sign Language in 1) regular education with additional support from a school of the deaf or 2) a signed bilingual educational setting'.

Even high technological digital hearing devices do not provide full access to speech since the perceived sound quality is limited (Dorman et al., 2017; Scarbel, Vilain, Loevenbruck, Schmerber, 2012). Therefore, deaf readers could benefit from visual prosody, which is already applied in teaching materials to exercise vocal variations with deaf readers by several institutions (KIDS, n.d.; Advanced bionics, unpublished; Staum, 1987; van Uden, 1973). For some examples, see Figure 1. Prosodic cues in reading materials help deaf readers aged 7 to 18 to read with more expression and with understanding of the intended meaning of a sentence (Renckens et al., 2021).

Figure 1.a, b & c.



Three examples of visual prosody in (Dutch) deaf education, all applying similar visualizations: 1. horizontally stretched words (lllllaaaaannnnggggg, translated as llllooooonnnnngggg) correlate with a longer duration. 2. bolder and larger text correlates with more loudness, and 3. a higher position, vertically stretched text, a rising line, or music notes correlate with a higher pitch in the voice (left to right: image from KIDS, n.d.; image from Advanced Bionics, unpublished; image based on van Uden, 1973).

Until now, visual prosody was developed with speech properties in mind (loudness, duration, pitch) but not sign language. Because deaf readers often rely not only on spoken language but also on sign language, this study aims to evaluate if visual prosody in text has a consistent relation with (Flemish) sign language. If it has, then sign language could possibly be used to explain/support (speech) prosody for deaf readers in bilingual education.

There are similarities and differences between oral and sign languages. Similar to words in oral languages, signs follow one after another. But unlike in spoken languages, sign language can engage several information sources simultaneously through different body parts and signs require free space surrounding a person to perform a sign (Koenen et al., 2005). Prosody in sign language can be found in two ways: in non-manual markers and in variations on the signs. First, signs are accompanied by non-manual markers (NMMs), which are all components of sign language that are not formed by the hands (Brentari, Falk & Wolford, 2015; Elliott & Jacobs, 2013; Van Herreweghe, 1995). Examples of prosody not made by the hands are the non-manual markers such as mimicry, body posture, shoulder raising, head position, facial expressions, eye blinks, eye shifts, gazes, the position of the lips, and position of the brows. Second, the literature review by Brentari, Falk & Wolford (2015) pointed to language-independent prosodic components in sign language, made by variations in duration, velocity, and displacement. Those prosodic components are always present, independent of how a language applies them, similar as to how pitch, loudness, and duration are always in the voice, regardless of language. It is this second kind of prosody that can be useful for this study: the progress of time is the same for word duration and sign duration in both modalities. The velocity in movement is supposed to resemble frequency in speech, and displacement of movement would correlate with intensity in speech (Brentari, Falk & Wolford, 2015).

But relating visual prosody to sign language is more difficult than relating it to the voice. Current prosodic cues are designed to be intuitive with the speech in mind. 'Intuitive' means that their intention can be interpreted without much explanation (see also Shaikh, 2009; Lewis and Walker, 1989). For example, 'boldness' correlates to 'loudness' (Shaikh, 2009). But in sign language, the line between syntax and prosody is not yet defined (Sandler, 2010). The prosodic components in sign language are more interwoven with the modality of the language itself: for example, while a signer consciously makes use of hands, face, and body to represent a concept in a sign (van den Bogaerde, 2012), signing the sign 'walking' slowly can mean that the person is walking slow. This interwovenness with the content could influence the intuitive understanding of prosodic cues for sign language.

To evaluate the aim of this research, the hypothesis of this research became *"Deaf readers relate typographic prosodic cues designed with the speech properties in mind consistently to prosodic components in Flemish Sign Language."*

Methodology

Prosodic cues

The existing prosodic cues within the typeface Matilda were used for this study. The relation of these prosodic cues with prosody in speech (according to Bessemans et al., 2019) and the components of sign language (according to the suggestions of Brentari, Falk & Wolford, 2015) is described in Table 1.

Figure 2.

A summary of the four prosodic cues used in this study. a thicker font, a wider space, a font raised above the baseline, and a wider font. The sentence is translated as "The big bird flies high."

- De grote vogel vliegt **hoog.**
- De grote vogel vliegt   hoog.
- De grote vogel vliegt hoog.
- De grote vogel vliegt hoog

Table 1.

The suggested relations/similarities between the prosodic cues, the speech variations, and the prosodic components in sign language (based on 1. Bessemans et al., 2019; which is based on earlier studies, and 2. Brentari, Falk & Wolford, 2015).

Visual cue in text for this research	Correlates with prosody in speech (Bessemans et al, 2019)	Correlates with prosody in sign language (Brentari, Falk & Wolford, 2015)
Thicker font	Louder voice	More displacement
Larger space	Longer pause	Longer pause
Font raised above the baseline	Higher pitch	Faster velocity
Wider font	Longer duration	Longer duration

This table is not a full summary of prosody: the opposite cues for the opposite direction of each speech variation exist: for example, a softer voice instead of a louder voice. However, a study does not have to include all possible cues to prove the effectiveness of a subset of cues (as in, for example, Patel & McNab, 2010, Patel, Kember & Natale, 2014, or Bessemans et al., 2019). To reduce complexity of this test for young readers, these additional cues were not applied in this study.

Video fragments

Four video fragments were created. In those videos, a 'neutral' sentence in Flemish Sign Language was shown first [Figure 1.A.], followed by the same sentence in Flemish Sign Language but with a modulated prosody on the last sign: a "longer pause," "longer duration," "faster velocity" or "more displacement" [Figure 1.B].

Please note that prosody in sign language is much more interwoven with the content of the message than the prosody in speech (e.g., the sign for walking, executed slowly). To avoid such connections, prosody in this research was treated as an independent factor, not connected to the meaning of the message. The effect of prosodic cues on a sign's meaning could become the subject of a follow-up study.

Figure 3.

A and B. In videos, a signer performs one recurring sentence with varying prosodic variations. A: The neutral sign is shown first. B: Each neutral sign was followed by a sentence containing one specific variation (e.g., in image B more displacement than in A).

Participants

The cues were tested by 38 deaf readers. In this study, this term refers to readers with hearing remnants who are able to speak. Their native language and preferred language could be either spoken or signed, but all participants were educated 1) regular schools and received additional support from schools of the deaf (often after starting their first education in a school of the deaf), or 2) in a bi-lingual educational environment. They thus had a high chance to come in contact with both spoken and signed languages. All readers were between 7 and 18 years old. The participants' age was evenly spread: for all ages between 7 and 18, each age was represented by at least two participants. These readers followed primary and secondary education in regular schools or schools for the deaf, and most wore one or two hearing devices. Some participants preferred Flemish Sign Language as primary language, and others preferred speech.

Because this research was executed at the same time as the study evaluating visual prosody's influence on the reading aloud (Renckens et al., 2021), the participants enrolled in both research studies at the same time. Therefore, the participants were aware that there was a relationship between the cues and speech prosody but did not yet know which one. Participants were not provided any information about prosody in sign language when this test started.

Procedure

The participants were asked to look carefully to the video [Figure 3]. Then, they were asked to choose from a list of sentences [Figure 2] the sentence that, according to him/her, corresponded with the last video shown. Participants did not have to know Flemish Sign Language fluently to follow the video fragments. All participants were encouraged to mark an answer. If they were not sure, they could see the video one more time before marking an answer.

This procedure with the videos was repeated: after connecting all movie fragments to written sentences a first time, the movie fragments were shown a second time without the participants knowing that the same videos were shown again. This way, participants watched all prosodic components in Flemish Sign Language twice. The order in which all participants saw the videos was: pause, thicker, extended, higher, pause, extended, higher, thicker.

Participants could provide feedback while answering the question, and the researcher asked questions when a participant was stuck.

Statistical Analysis

While relating the cues to the movies, a binomial test with one chance on four to guess correctly, results of 38 persons differ significantly from pure chance ($p<0.05$) if 14 or more (i.e., 36% or more) give the same answer.

To see if there was a learning effect and to test if there was a difference between primary and secondary school, proportion tests for two proportions were used.

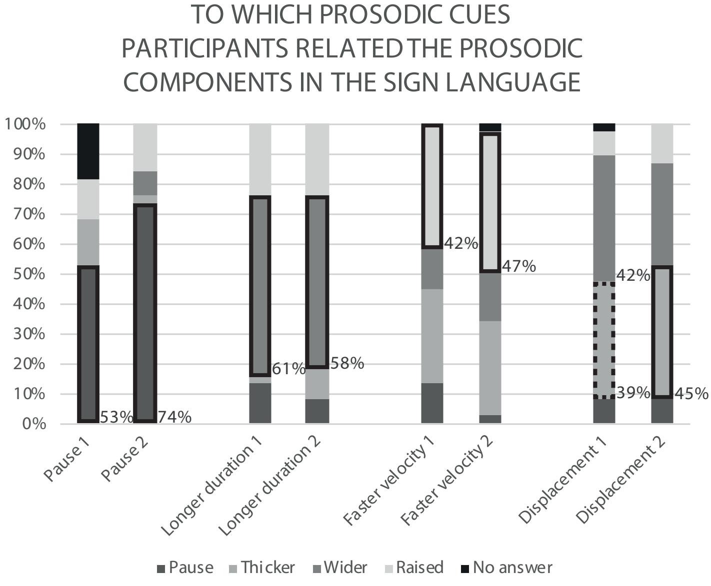
Results

Relating cues to the movies

Except for the first time that 'more displacement' was shown, all cues were related to their intended component in Flemish Sign Language [Chart 1.]. The first time that 'faster displacement' was presented, participants marked the non-intended wider font more often than the intended thicker font [as shown in Figure 1.C] (42% versus 39%). All cues were significantly more often related to their intended component in Flemish Sign Language than if participants would have been guessing. No statistically significant learning effects were found between the first and second time a video fragment was shown.

Chart 1.

Except for the first time that 'more displacement' was shown (dashed border), prosodic components in Flemish Sign Language were related most of their occurrences to their intended cue (solid border).



No difference between primary and secondary school

In total, 22 of the readers took classes in primary school, and 16 of the readers took classes in secondary school. How the participants related the videos with Flemish Sign Language to the intended prosodic cues is stated in Table 2.

Table 2.

When participants had to relate videos with prosodic cues, the results varied between 38% and 77% correctness.

Number of correct answers when participants relate the videos to prosodic cues						
	all (#)	all (%)	Primary (#)	Primary (%)	Secondary (#)	Secondary (%)
Pause 1	20	53%	11	50%	9	56%
Displacement 1	15	39%	9	41%	6	38%
Longer duration 1	23	61%	13	59%	10	63%
Faster velocity 1	16	42%	10	45%	6	38%
Pause 2	28	74%	17	77%	11	69%
Displacement 2	17	45%	11	50%	6	38%
Longer duration 2	22	58%	12	55%	10	63%
Faster velocity 2	18	47%	9	41%	9	56%

There was no significant difference between the percentage of correct answers of primary school readers versus secondary school readers for any of the cues presented. It can be argued that this lack of significance is due to the relatively small number of samples. However, notice that in four situations, the students at the secondary school score better, and that in the other four situations, the children of the primary school score better. This favors the idea that there is no age effect.

Some comments of the children during the sessions

The children's comments were not recorded explicitly, as the children were free to comment on anything that they felt was important. However, two comments kept returning and were deemed important enough to be mentioned here.

First, we noticed that not a single participant used the terms "faster velocity" or "slower duration"; both were described in terms of faster/slower "speed."

Second, participants deemed the faster velocity more 'difficult' to relate to a font. Four participants explicitly said that they were searching for a prosodic cue that was not in this test but was present in the other tests evaluating speech (such as a narrower font resembling a faster speech, see Renckens et al., 2021). At least six more participants were stuck on this cue, even to that level that the researcher had to intervene and ask what the problem was. The researcher had to tell that, even when they did not see the typeface they were looking for, an answer had to be chosen. Thus, one participant on four was clearly struggling with the cue for a faster velocity.

Discussion

The cues designed with speech prosody in mind do relate to the intended prosodic components in Flemish Sign Language. A longer pause in sign language correlates with a larger word-space; a sign with a longer duration with a wider font; a sign with more displacement with a thicker font; a sign with a faster velocity correlates with a raised font. These findings are in line with Brentari, Falk & Wolford (2015)'s review of the similarities between the prosody of speech and sign language (table 1).

We did not find an effect of age. This implies that participants in primary school similarly associated the textual prosodic cues with prosody in Flemish Sign Language as the children in secondary school. This has the advantage that, once the children can read, prosodic cues could be introduced and afterward be used during the whole period they go to school.

But the participants' oral feedback exposed confusion when these cues are applied to sign language. The prosodic cues used in this study were designed with three very distinct speech variations in mind: loudness (in decibel), duration (in milliseconds), and pitch (in Hertz). The prosodic components in sign language are very different: displacement (in distance), duration (in milliseconds) and velocity (distance per millisecond). Here, velocity closely resembles duration: both are expressed in 'time'-units (often seconds or milliseconds). This close interwovenness caused confusion for the participants. It seemed that they intuitively only extracted two prosodic components in sign language: displacement and speed. This constitutes the main drawback of this study. We did not anticipate the difficulty the children had with the three prosodic components. We only learned during the data collection that the interwovenness of duration and velocity was exceptionally strong. Because we did not anticipate this issue, we had no alternative prosodic cues to address it.

Despite this confusion, the intended cue for a faster velocity (the raised text) was marked most of the time (42% and 47%).

That could be explained by the fact that the cue that some participants expected, a narrower font, was not presented (participants knew that cue from the research about the voice that ran at the same time). Participants thus choose a raised font, probably because the bold and wider fonts were already related to another prosodic component in sign language. If more cues would have been provided to choose from, it is uncertain that the task of relating cues to the movies would have been as consequently answered as it was now.

While this study focused on Flemish Sign Language, we can assume that the findings are valid for multiple sign languages, in line with Brentari, Falk, and Wolford (2015), who discussed several sign languages. This should be evaluated with different sign languages and different cultures as a reference, which both could have another perception of prosody.

At this moment, the practical usage of this preliminary study is limited. We evaluated the three prosodic cues most commonly used in visual prosody for speech. The participants' oral feedback pointed out that this approach cannot be transferred to sign language easily. We advise performing a new study that evaluates if two prosodic cues are sufficient to represent the intrinsic characteristics of (Flemish) Sign Language: on one side 'displacement,' on the other side velocity and duration grouped together as 'speed.'

Conclusion

This research confirms the hypothesis that *"Deaf readers relate typographic prosodic cues designed with the speech properties in mind consistently to prosodic components in Flemish Sign Language."* The test was done with participants relating three prosodic cues in type to movies showing prosodic components in sign language. The prosodic cues were correctly related to their intended prosodic component in the sign language.

However, the three prosodic components in sign language are more closely interwoven with each other than the three prosodic speech variations on which the prosodic cues are based. That caused confusion in the terminology used by the readers: "velocity" and "duration" were both referred to as "speed." Thus, readers intuitively extract only two prosodic components in sign language and encounter problems with three prosodic cues in text.

Based on the results of this study, we recommend that visual prosody has to be adapted for sign language. This requires further research.

Acknowledgments

We would like to thank all schools in the Dutch-speaking regions that supported this research and facilitated connecting with deaf students. In alphabetical order: Antwerp Plus (Antwerp), BUSO Zonnebos (Schoten), Cor Emous (The Hague), Kasterlinden (Kasterlinden), KIDS (Hasselt), Koninklijk instituut Woluwe (Woluwe), Sint Gregorius (Ghent-Bruges), Sint-Lievenspoort (Ghent), Spermalie (Bruges). Without their help, this research would not have been possible.

References

ADVANCED BIONICS. (unpublished version). *A musical journey through the rainforest.* [Teaching material]

BESSEMANS, A.; RENCKENS, M.; BORMANS, K.; NUYTS, E.; & LARSON, K. (2019) "Visual prosody supports reading aloud expressively." *Visible Language.* 53 (3): 28-49.

BOONS, T.; DE RAEVE, L.; LANGEREIS, M.; PEERAER, L.; WOUTERS, J.; VAN WIERINGEN, A. (2013). "Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation." *Research in Developmental Disabilities*. 34: 2008–2022.

BRENTARI, D.; FALK, J.; & WOLFORD, G. (2015). "The acquisition of prosody in American Sign Language." *Linguistic Society of America.* 91 (3): e144-e168. https://doi.org/10.1353/lan.2015.0042.

CHIN, S.B.; BERGESON, T.R.; PHAN, J. (2012). "Speech Intelligibility and Prosody Production in Children with Cochlear Implants." *Journal of Communication Disorders.* 45(5). 355–366. http://dx.doi.org/10.1016/j.jcomdis.2012.05.003.

DE CLERCK, I; PETTINATO, M.; GILLIS, S.; VERHOEVEN, J.; & GILLIS, S. (2018). "Prosodic modulation in the babble of cochlear implanted and normally hearing infants: a perceptual study using a visual analogue scale." *First Language.* 38 (5): 481–502. https://doi.org/10.1177/0142723718773957.

DORMAN, M.F.; COOK NATALE, S.; BUTTS, A.M.; ZEITLER, D.M.; & CARLSON, M.L. (2017). "The sound quality of cochlear implants - Studies with single-sided deaf patients." *Otol Neurotol.* 38 (8): e268–e273. https://doi:10.1097/MAO.0000000000001449.

ELLIOTT, E.A. & JACOBS, A.M. (2013). "Facial Expressions, Emotions, and Sign Languages." *Frontiers in Psychology*. 4 (115) https://doi.org/10.3389/fpsyg.2013.00115.

KIDS (Unpublished). *Teaching materials for children with language disorders.* [Internal teaching materials].

KOENEN, L.; BLOEM, T.; JANSSEN, R.; & VAN DE VEN, A. (2005). *"Gebarentaal. De taal van doven in Nederland (Sign Language. The language of deaf in the Netherlands)."* Vi-Taal: The Hague.

LEWIS, C., & WALKER, P. (1989). "Typographic influences on reading." *British Journal of Psychology.* 80, (2), 241-257.

MARKIDES, A. (1983). *The speech of hearing-impaired children.* Manchester University Press, New Hampshire, USA.

ØYDIS, H. (2013). *Acoustic Features of Speech by Young Cochlear Implant Users.* [Dissertation] Universiteit Antwerpen.

PATEL, R., KEMBER, H. & NATALE, E. (2014). "Feasibility of Augmenting Text with Visual Prosodic Cues to Enhance Oral Reading." Speech Communication. 65. https://doi.org/10.1016/j.specom.2014.07.002

PATEL, R. & MCNAB, C. (2010). "Feasibility of augmenting text with visual prosodic cues to enhance oral reading." *Speech Communication*. 53 (3): 431-441.

RENCKENS, M.; DE RAEVE, L.; NUYTS, E.; PEREZ MENA, M.; BESSEMANS, A. (2021). Visual prosody supports reading aloud expressively for deaf readers. *Visible Language:* 55 (1).

RUDE, M. (2016). *Prosodic writing shows L2 learners intonation by 3d letter shapes - state, results, and attempts to increase 3d perception.* [Intern document]. University of Nagoya.

SANDLER, W. (2010). "Prosody and syntax in sign languages." *Trans Philol Soc*: 108(3): 298–328. https://doi.org/10.1111/j.1467-968X.2010.01242.x.

SCARBEL, L.; VILAIN, A.; LOEVENBRUCK, H.; SCHMERBER, S. (2012). "An acoustic study of speech production by French children wearing cochlear implants." *3rd Early Language Acquisition Conference, Dec 2012,* Lyon, France.

SEE, R.L.; DRISCOLL, V.D.; GFELLER, K.; KLIETHERMES, S. & OLESON, J. (2013). "Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing." *Otology and Neurotology*. 34 (3): 490–8. https://doi.org/10.1097/MAO.0b013e318287c985.

SHAIKH, D. (2009). "Know your typefaces! Semantic differential presentation of 40 onscreen typefaces." *Usability News. 11,* (2).

STAUM, M.J. (1987). "Music Notation to Improve the Speech Prosody of Hearing Impaired Children." *Journal of Music Therapy.* 24 (3): 146-159.

VAN DEN BOGAERDE, B. (2012). "Kun je alles zeggen in gebarentaal? Over taal van doven en slechthorenden (Can you say anything in sign language? About language of deaf and hard of hearing)." [Online] http://www.taalcanon.nl/vragen/kun-je-alles-zeggen-in-gebarentaal. [1 March 2017].

VANDER BEKEN, K.; DEVRIENDT, V.; VAN DEN WEYNGAERD, R.; DE RAEVE, L.; LIPPENS, K.; BOGAERTS, J.; MOERMAN, D. (2010). "Personen met een auditieve handicap (Persons with an auditory handicap)." In: BROEKAERT et al. (2016: fourteenth press) *Handboek bijzondere orthopedagogiek (Manual for extraordinary remedial education).* Garant. Antwerpen-Apeldoorn. p 131-210.

VAN HERREWEGHE, M. (1995). *De Vlaams-Belgische gebarentaal: een eerste verkenning (The Flemish-Belgian Sign Language: a first exploration).* Ghent: Academia Press.

VAN UDEN, A. (1973). *Taalverwerving door taalarme kinderen (Language acquisition by language-poor children).* Universitaire Pers Rotterdam, Rotterdam.

Authors

Maarten Renckens

Maarten Renckens is a teacher and design researcher with a love for letters and a heart for people. Dealing with a reading difficulty himself, he is very interested in the reading process. His projects include the typeface 'Schrijfmethode Bosch' (Writing Method Bosch) that learns children how to write and typefaces to encourage beginner readers and readers with hearing loss to read more expressively. With a background in architectural engineering, he is used to approach concepts technically and mathematically. He applies this technical knowledge to unravel letterforms, in order to determine the effects of different letterforms on the reading process.

Leo De Raeve

Leo De Raeve PhD has 3 professions: he is a Doctor in Medical Sciences, psychologist and teacher of the deaf. He is founding director of ONICI, the Independent Information and Research Center on Cochlear Implants, is lecturer at University College Leuven-Limburg and scientific advisor of the European Users Association of Cochlear Implant (EURO-CIU).

Erik Nuyts

Prof. Dr. Erik Nuyts is researcher and lecturer at the University College PXL and associate professor at the University of Hasselt. He got a master degree in mathematics, and afterwards a PhD in biology.

Since his specialty is research methodology and analysis, his working area is not limited to one specific field. His experiences in research, therefore, vary from mathematics to biology, traffic engineering and credit risks, health, physical education, (interior) architecture and typography.

His responsibilities both at the University College PXL as at the University of Hasselt involve preparation of research methodology, data collection and statistical analyses in many different projects. He is responsible for courses in research design, statistics, and mathematics.

María Pérez Mena

Dr. María Pérez Mena is an award-winning graphic and type designer. She is postdoctoral researcher at the legibility research group READSEARCH at PXL-MAD School of Arts and Hasselt University. María teaches typography and type design in the BA in Graphic Design at PXL-MAD and is lecturer in the International Master program 'Reading Type & Typography' and the Master program 'Graphic Design' at the same institution. She received her PhD "with the highest distinction" from University of Basque Country and is a member of the Data Science Institute UHasselt.

Ann Bessemans

Prof. Dr. Ann Bessemans is a legibility expert and award-winning graphic and type designer. She founded the READSEARCH legibility research group at the PXL-MAD School of Arts and Hasselt University where she teaches typography and type design. Ann is the program director of the international Master program 'Reading Type & Typography'. Ann received her PhD from Leiden University and Hasselt University under the supervision of Prof. Dr. Gerard Unger. She is a member of the Data Science Institute UHasselt, the Young Academy of Belgium and lecturer at the Plantin Institute of Typography.

Co-designing to Improve Practice in Treating Urinary Tract Infections: a case study of reducing inappropriate antibiotic treatment

Guillermina Noël
Darren Pasay
Denise Campbell-Scherer
Lynora Saxinger

Although co-design has made important contributions to practice in many fields, healthcare has only recently employed it for shaping best practices. This paper explores an aspect of medical practice that challenges many hospitals: the decision-making process for ordering urine testing and the use of antibiotics, specifically, in treating Urinary Tract Infections (UTI) and Asymptomatic Bacteriuria (ASB). The case investigates how physicians and nurses make decisions about testing urine for infection and the use of antibiotics. To explore the issue, the researchers conducted three co-design workshops to (1) uncover the medical decision-making process in ordering urine testing and treating UTI and ASB, (2) determine the needs of clinicians as they make such decisions, and (3) collectively design a decision aid that would fit users' cognitive, emotional and physical needs. The case shows how human-centered design approach led to an evidence-based decision-making tool – guiding clinicians to improve their practices by reconsidering when and if to order urine testing and prescribe antibiotics.

Keywords –
human-centred design,
co-design,
collaboration,
healthcare,
clinicians,
urinary tract infections,
asymptomatic bacteriuria

1. Introduction

This case study applies a human-centered approach to improve medical practice. Human-centered design is a framework to improve people's way of living, working, and doing other essential daily activities that focuses attention on people's capacities and needs, seeks to define the problem and its underlying causes collectively, and uses iterative cycles to design and evaluate prototypes (based on Meyer & Norman, in press). While human-centered design approaches have flourished in healthcare practice improvement, few studies describe the processes followed and provide a detailed account about how to engage stakeholders in the design process. In healthcare, medical expertise has tended to drive the creation of communication tools – such as patient education materials, guidelines, and decision aids – written and designed, for the most part, by healthcare professionals.

Recently, increased collaboration between experts in design and healthcare is emerging (Breslin et al., 2008; Hargraves, 2018; Noël et al., 2019; Roberts et al., 2016; Witteman et al., 2015). This article demonstrates how human-centered design can help reduce unnecessary antibiotic prescription. This case illustrates the process and methods used to support medical practice change.

This study begins with an overview of both the medical problem and the communication design problem faced. The study elaborates how we engaged clinicians with heavy demands in co-design and together co-created a decision aid. Co-design is a human-centered design method, where "the person who will… be served through the design process is given the position of 'experts of their experience,' and plays a large role in knowledge development, idea generation and concept development" (Sanders & Stappers, 2012, pp. 23-24).

The medical problem: Confronting and changing lore

Urinary tract infections (UTIs) are a common bacterial infection that doctors diagnose during Emergency Department visits. According to Abbo and Hooton (2014), "10.8 million patients in the United States visited an Emergency Department (ED) for the treatment of a UTI between 2006 and 2009" (p. 175). There is great variation in practices regarding when to order a urine test, how to interpret symptoms of UTI, and when to initiate antibiotic treatment (Beckford-Ball, 2006). It is common practice to order urine testing (both urinalysis and urine cultures) in the absence of typical UTI symptoms, such as dysuria (discomfort) and urinary frequency (Flokas et al., 2017; Nicolle et al., 2019). This practice occurs partly because diagnosis of UTI is difficult in older adults with non-specific symptoms, as they are more likely to have bacterial colonization of the bladder (Beckford-Ball, 2006, p. 1). Criteria for the diagnosis of urinary tract infection varies, depending on the patient and the context: symptoms / no symptoms, men / women, pregnant women / non-pregnant women, catheter / no catheter, adult / children, and so on (Scottish Intercollegiate Guidelines Network (SIGN), 2012).

ASB occurs with a positive urine culture but without local or systemic symptoms of a UTI. Except in certain cases, such as pregnancy and prior to invasive urological procedures, treating ASB with antibiotics is not beneficial and is potentially harmful. Clinicians often find it difficult not to treat positive urine culture results, and patients are exposed to unwarranted antibiotic therapy (American College of Radiology, 2012; Barlam et al., 2016; Fleming-Dutra et al., 2016).

Antimicrobial exposure can, for example, contribute to the development of bacterial resistance, result in adverse and allergic reactions in patients, and have unintended consequences, such as *Clostridium difficile* infections. These bacteria can cause problems for patients – ranging from diarrhea to life-threatening inflammation of the colon (File et al., 2014). Antimicrobial stewardship endeavours to limit antimicrobial exposure to cases where it is absolutely necessary.

Problematically, established practices, professional lore, and public beliefs show unawareness of the diagnostic criteria of UTI and that treating ASB with antibiotics is not beneficial.

Motivation for study

This case study was initiated by the "Appropriateness of Care: Asymptomatic Bacteriuria (ASAB)" initiative of Alberta Health Services (Canada), which sought to reduce the inappropriate urine culture testing and use of antibiotics in hospitals and long-term care facilities. In particular, the health service explored the following beliefs and practices:

- Beliefs that non-specific clinical status changes or cloudy/smelly urine (particularly in older adults or catheterized patients) are the result of UTI (Rowe et al., 2014).
- Beliefs that a positive urinalysis or a test of urine cultures is always indicative of UTI (Juthani-Mehta, 2015).
- 'Routine' and 'opportunistic' collection of urinalysis and urine cultures (as part of a panel of investigations) as accepted practice (Rowe et al. 2014).

The ASAB group partnered with the Physician Learning Program (PLP) at the University of Alberta to better understand the problem and to collaborate on the design of an antimicrobial and diagnostic stewardship decision-making tool that would ultimately improve patient care.

Reframing the design problem: Achieving safer outcomes

In 2017, the PLP applied a human-centered design approach to advance health practice. Human-centered design has shown to increase the effectiveness of communication tools in healthcare, particularly decision aids and patient information aimed at translating new evidence into practice (Erwin & Krishnan, 2016; Garvelink et al., 2016; Thornhill et al., 2017; Noël et al., 2018; Ragouzeos et al., 2019; Adam et al., 2019; Zeballos-Palacios et al., 2019).

The ASAB team provided a draft of an educational toolkit for patients and clinicians. In this article, we focus on the co-design of one key component of this toolkit: An evidence-based decision aid designed to improve clinicians' decision-making about when to test urine for infection. To see the complete final toolkit, visit https://www.albertahealthservices.ca/info/Page15718.aspx.

The design problem we faced was not only to optimize the decision aid through more effective visual and verbal language. The designer worked with the partner collaborators to reframe the problem and its scope–from improving a visual tool to supporting decision making to achieve safer health outcomes.

Situating co-design and collaboration

We tried to create the optimal conditions for sharing, reflection, collective reasoning, and exploration through co-designing. If the decision aid was to be successful, it would need to encourage clinicians to make decisions based on recent evidence. Co-designing helps participants feel ownership of the project. Each participant has a role in creating a better reality.

Co-design promotes the collective understanding of what is occurring (Britton, 2017, p. 40). Healthcare change practices benefit from co-design because it gives the community a major say in the process (Meyer & Norman, 2020, p. 17). Understanding the diverse needs and perceptions of a situation people have is essential to change practices. Others refer to this as the human dimensions of change (NHS Institute for Innovation and Improvement, 2005). However, engaging very busy healthcare professionals (physicians, nurses, and pharmacists) in co-design is not straightforward. People do not always have in mind what could help (Flores, 2012, p. 4). Part of the role of the human-centered designer was to develop a strategy to engage clinicians in reflecting about a problem that some were not aware of, with participants sensing and becoming increasingly aware of divergent perspectives, and ultimately, connecting, understanding, exploring, and solving issues collaboratively.

The designer creates activities and facilitation tools to foster engagement, and the sharing of information, perspectives, and knowledge to re-construct and adapt new problem representations. If the designer facilitating the process is unable to create suitable activities, problem exploration and reframing will not happen.

Some healthcare leaders tend to refer to the practice of producing knowledge with stakeholders as co-creation (Greenhalgh et al., 2016). They view co-creation as a collaborative process in which academics and stakeholders produce knowledge, rather than only translating it into action. Elwyn et al. (2019) consider that co-production fosters learning healthcare systems. "Coproduction also makes a connection between practice improvement and organizational design by leveraging the power of learning health systems towards the increasing focus on value-based care" (p. 715). This learning requires iterative processes of questioning and re-thinking (Senge, 2006, p. 324).

We see co-design as a process of mutual learning, exploring possibilities, and making them visible through visual and verbal language – a process in which designers and non-designers work collaboratively (Sanders & Stappers, 2008). Increasingly, collaboration and facilitation skills are becoming more relevant to designers' work (Voûte et al., 2020, p. 63).

During co-design, participants typically reveal aspects of the problem that the leading team does not have access to and is unaware of. Multiple perspectives help the team both modify their representation of the problem and restructure it to account for the team's collective expertise. As participants re-represent their task, they engage in reframing the problem.

2. Methods

Rarely, publications reporting on antibiotics stewardship discuss the team's design process and methods, making it difficult to see the contribution of design to the project (Betsch et al., 2018; Formoso et al., 2013). This section outlines how we engaged clinicians and how we gained knowledge.

We strove to understand the people using the tool – their knowledge, assumptions, and misconceptions. Human-centered design means learning about people, their worlds, and their goals (Frascara, 2017).

To build understanding from the UTI and ASB problem from the clinicians' perspectives, we created three co-design workshops to collectively:

1. Uncover the medical decision-making process in assessing UTI and ASB.
2. Identify ways to guide clinicians to make more evidence-based decisions about diagnosing and treating UTI and ASB.

3. Design a tool that would meet the cognitive, emotional, and physical needs of healthcare providers.

Participants

We recruited participants through the University of Alberta Hospital. Our protocol was approved by the Health Research Ethics Board of the Faculty of Medicine and Dentistry at the University of Alberta.

Forty-four people volunteered to participate in three co-design sessions. Participants were nurse practitioners, nurse educators, nursing students, pharmacists, and physicians from the Emergency, Gastroenterology, and Orthopedic Departments of the University of Alberta Hospital. Volunteers were excluded if their daily routine did not include urine testing.

The role of the participants is to bring different voices, their experiences, and deep knowledge of the caring process; to share their ways of doing; their learnings. The role is to engage in dialogue and collaborate in designing more appropriate care practices shaping their future ways of practicing.

Timeframe

Two workshops were conducted in September of 2017, with six participants and eight participants, respectively. A third workshop took place in March of 2018, with thirty participants: unit managers, clinical nurse educators, registered nurses, licensed practical nurses, and patient care managers. Participants co-created in groups. Each workshop was conducted for 90 minutes.

After the first two co-design sessions, the ASAB steering committee met to discuss the learnings, and changes proposed, and approved them. The decision aid was revised in December 2017, piloted in January 2018, and revised again in February 2018. At this point, the team decided to hold a third co-design workshop. Between March 2018 and January 2019, the decision aid was revised, piloted, and approved by the ASAB steering committee.

Tools and materials

To examine the problems associated with assessing and treating UTI and ASB, we created:

1. **Poster-sized quotes**: To encourage participants to think about the problem, we extracted 15 quotes from 'Urinary Tract

Infection—Requiem for a Heavyweight' (Finucane, 2017). The A3 quotes were hung on a wall.

2. **Scenarios**: To situate participants, we explored the following: *"Suppose I* (facilitator) *am a patient…?"* and *"If this is the process…"*

3. **Dialogue prompters**: To help participants engage in collective thinking, we designed tools for fostering dialogue, reflection, and problem representation.

4. **Provotypes and prototypes**: Provotypes help provoke initial exploration. Prototypes help team members think through a design problem by testing solutions.

Procedures and activities

Participants were invited to immerse themselves in reflection to think about how to change the situation. They were recognized as having valuable experience and knowledge – without which, the problem could not be effectively understood. We asked participants for their help in thinking through how to help clinicians change their routine practices.

Together, the team introduced the problem from medical and design perspectives. From a medical perspective, the problem definition was framed using research evidence (Blakiston & Zaman, 2014; Flokas et al., 2017), local data, and knowledge about the local clinical setting and circumstances. The basic medical problem was as follows:

> *There is no clinical benefit obtained by treating individuals with ASB, particularly older adults. However, the belief that cloudy or smelly urine is indicative of UTI leads to overdiagnosis and unnecessary antibiotic use. This promotes antimicrobial resistance. In Alberta, there are more than 610,000 urine cultures processed annually. These tests result in costs exceeding $9,000,000 (CAD). How can we improve this?*

This setup allowed the group to embrace a common interest and mission, to become a "pledge group" (Von Krogh et al., 2000, p.15). The design problem built on the medical problem and explored:

> *What is the current decision-making process when assessing UTI and ASB, and how can we better help clinicians to quickly decide when it is appropriate to order urinalysis and urine cultures to test urine for infection?*

The two perspectives helped us build a common goal to pursue. Each co-design activity began with a set of instructions that were read aloud and discussed. The instructions were a series of steps, showing how to explore the problem by breaking it into smaller and simpler problems (Rumelhart, 1992).

The three co-design sessions were facilitated by the physician lead, who specializes in infectious diseases, and the designer. The role of the physician lead was to bring her depth of knowledge about UTI and ASB, the current problem we were facing, and clinical practice. The role of the designer was to orchestrate activities to help clinicians immerse in a problem space, reflect on the problem situation, help see the problem from diverse perspectives, foster engagement reasoning, and enable creativity and imagination. The designer observed interaction dynamics, posed questions, noted how the diverse collective activities (listening, reasoning, imagination) were going, and adapted or proposed changes to keep momentum and achieve the goals.

Activity 1: Process mapping

We invited participants to begin by reading the 15 poster-sized quotes. In the first 10 minutes, participants co-created a process map. We engaged in a simulation to identify the decision-making process that healthcare professionals typically follow when assessing UTI and ASB. The goal was to help participants situate themselves in daily practice—to think, raise awareness, and reflect about what they do. We presented the following scenario: *Suppose I* (facilitator) *am a patient waiting in the Emergency Room; what steps would you* (clinician) *take to diagnose and treat UTI?* Then the groups discussed the process map, each group describing the procedures they would follow.

To help clinicians engage in process mapping, we use dialogue prompters to pre-structure activities and encourage the team to think collectively about the problem (see Figure 1).

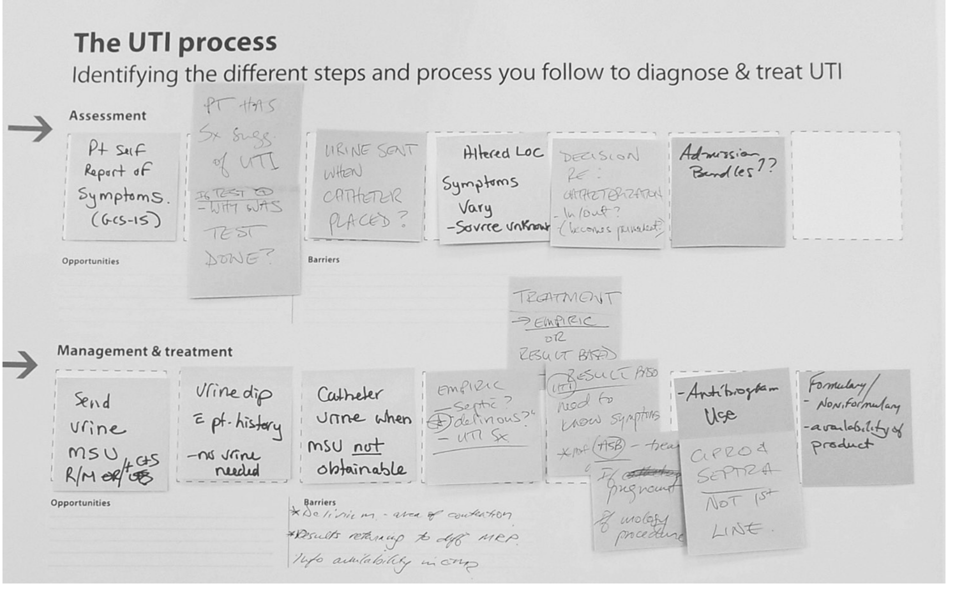


Figure 1.

The dialogue prompter used to facilitate reflection about the process.

These tools create a shared communication space (Frascara, 2016, p. xix-xx). The dialogue prompters were visible to the participants throughout the process.

Activities 2 & 3:

Co-analysis & co-design of the decision aid

For a second co-design activity, we designed another dialogue prompter (see Figure 2) with the following scenario*: If this is the process you follow, in what way does this decision aid assist in decision making*? The dialogue prompter helped participants reflect on the sequence of information, the relevance and clarity of the content, the appropriateness of the vocabulary, the tone of voice, and other aspects. The groups had 20 minutes to co-design the *decision aid* and 10 minutes to share their reflections. This activity was repeated for each workshop.

The dialogue prompter (Figure 2) included a provotype of the decision aid called "UTI Algorithm." Our goal was to provoke discussion about the suitability of the content and its visual presentation. Previous work suggests that provotypes can help participants explore possible solutions at early stages of the design process (Erwin & Krishnan, 2017; Boer & Donovan, 2012). We wanted to learn about how clinicians would interact with the decision aid and use that understanding to generate ideas for improvement. This activity helped arrive at initial performance requirements (Frascara & Noël, 2012).

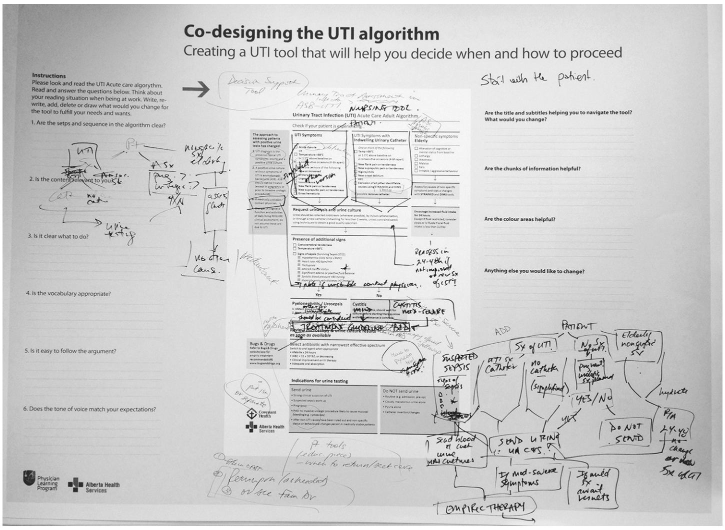


Figure 2.

Dialogue prompter with the provotype and the changes suggested by the participants. The diagrammatic sketch at the bottom-right of the image reveals the thinking process participants were following to adapt the algorithm to their practice processes.

The diagrammatic sketch at the bottom-right in Figure 2 provided the basis for the first prototype used in Workshop 2 (Figure 3). The group made the decision points more apparent.

Figure 3.

Dialogue prompter with the emergent prototype created by the co-design participants in workshop 2.

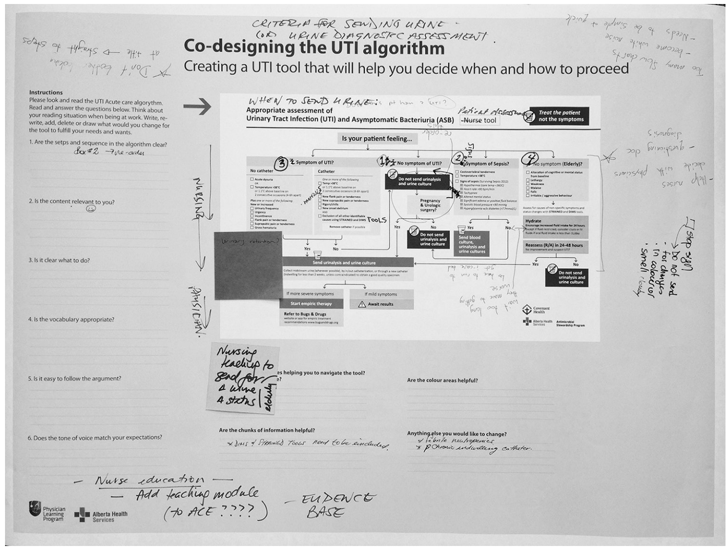


Figure 4.

Participants during the third workshop engage in co-design of the prototype and make changes so the tool better fits the cognitive, emotional, and physical context of the users.

The third activity focused on co-designing the decision aid (see Figure 4). It integrated the thinking that resulted from previous activities. The goal was to foster the collective framing of the problem from the perspective of people with different medical roles and to design a better tool.

Table 1. Summary of co-design activities and tools and their roles in orchestrating collaboration

Activities	Goals	Tools & roles
1. Process mapping	• To identify the decision-making process participants follow when assessing for possible UTI and ASB. • To assist clinicians to situate and articulate their current practices. • To recognize barriers and areas for improvement.	Dialogue prompter, markers, quotes, scenarios, and dialogue help participants situate themselves in daily practice. To engage in a simulation to identify the decision-making process. The dialogue prompter helps to pre-structure the mapping activity and encourages participants to think collectively about the problem. To recognize different roles, diverse practices, and interrelationships in the process.
Discussion	• To learn from others, sharing and gaining understanding. • To identify and articulate existing knowledge.	Filled-in dialogue prompters, dialogue. To encourage collective and reflective listening. To discover different ways of practicing, knowledge gaps, and biases. To start re-framing the problem collectively.
2. Co-analysis of the existing algorithm (decision aid)	• To explore whether the algorithm assists the decision-making process. • To identify information needs.	Second dialogue prompter, scenarios, a provotype (in the first workshop), prototypes (in the second and third workshop), dialogue. The dialogue prompter guides reflection. The Provo/prototype fosters reasoning, providing a model to evaluate and mentally simulate the use of the tool to improve clinicians' decision-making. It promotes critical thinking, assessing decision points and their consequences. It helps engage in problem-solving, identifying constraints, flaws, and ways to overcome them. Both tools help express diverse needs, perspectives, and fears.
3. Co-design of the decision aid The mapping and simulation, the listening, analysis, and exploration from the previous activities served to d the basis for the co-design of a new decision aid.	• To engage clinicians in the design of the decision aid. • To collectively establish what the tool should help achieve and how to achieve it. • To design a tool that fits the users' cognitive, emotional and physical context.	The dialogue prompters employed in the previous activities, dialogue. The second dialogue prompter with Provo/prototype helps reason through the problem and generate ideas to design a decision aid that supports better practice processes. The prototype fosters reasoning by model manipulation and adaptation (the notes and changes suggested reflect that). It helps explore hypotheses (problem-solving) and draw while thinking (creative thinking). The prototype facilitates reasoning, problem-solving and helps design a new decision aid.
Discussion	• To identify different needs, opinions, perspectives. • To share and gain understanding.	The completed filled-in dialogue prompters, dialogue. They help articulate understanding: what was modified and what was gained.

Data analysis

We took verbatim notes during the workshops and aggregated participants' comments about the tools. Three of the co-authors participated in an iterative process to group, label, and code the data. We analyzed the data through affinity diagramming, allowing theme identification (Hartson et al., 2012). Affinity diagrams can reveal the scope of participants' problems and help develop performance requirements (Beyer & Holzblatt, 2017).

We used the findings from the affinity diagramming to develop performance requirements and guide design decisions.

This process transformed our research findings – from what people did and why they did it – to what they needed to do and what the design should do to get them there.

3. Findings

The co-design activities helped participants to understand the project's goals and flesh out their own decision-making process. Participants uncovered knowledge gaps, identified points of resistance to the message, and brought to the surface fears about implementing the new process. Participants pointed to content that needed improvement and to information requirements for making an informed decision. Immersing participants in the activities allowed us to co-design a tool that fits clinicians' capacities and needs.

Our findings shed light on how co-design facilitated (1) thinking through the problem space while mobilizing participants' knowledge and skills, and (2) developing an actionable decision aid for diverse healthcare providers and settings.

Thinking through the design space

Co-design helped participants navigate the problem by creating a safe space for collaboration, uncovering tacit knowledge, seeing and thinking through the problem, and recognizing fears and beliefs about diagnostic decisions.

Design facilitation: Creating a safe space

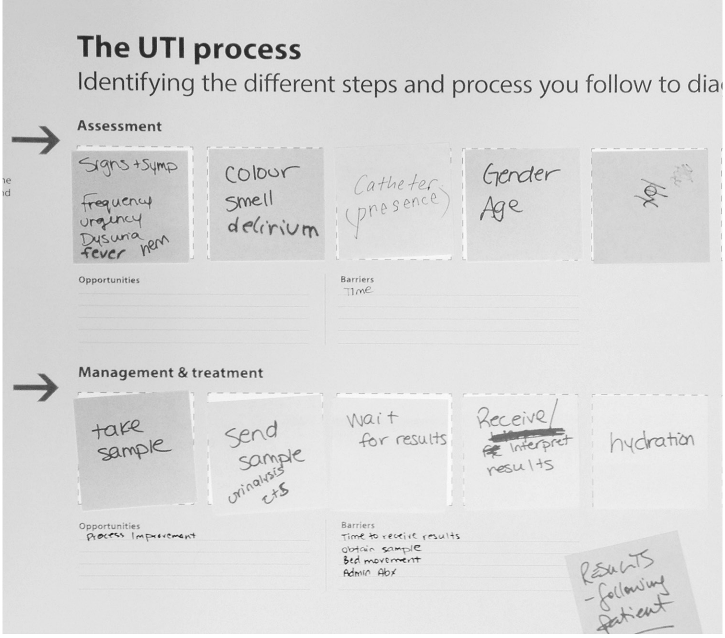
Asking participants for their help was key in setting an appropriate atmosphere for co-design. We made it clear that each participant's expertise was relevant to building an actionable decision aid. Introducing the problem from both medical and design perspectives helped us build a common understanding. Articulating the goals for each activity established a structure for working together. By making goals explicit and allowing participants to see how their ideas would form the basis for the decision aid, we set up the conditions for participants to feel comfortable about voicing their opinions and collaborating. We created a safe space.

Process mapping: Uncovering tacit knowledge

Process mapping allowed us to uncover participants' tacit knowledge. The maps revealed variations in how UTI and ASB are diagnosed and treated. Participants indicated a need for a standardized practice, which became a motivator. The maps raised awareness about current practices and knowledge gaps. For example, that some care teams considered cloudy or smelly urine as indicative of UTI. Other teams used urinalysis to rule-out UTI, while others used urine dipsticks (a diagnostic tool used to determine changes in urine) (see Figures 5 & 6). These variations confirmed that some practices were inconsistent with those suggested by best practices (as shown in Figure 5).

Figure 5.

For this care team, the second step in the process was to check changes in a patients' urine for cloudiness or smell. However, recent evidence shows that they are not necessarily indicative of UTI.



The content needed improvement in two areas: to educate about the benign nature of ASB; and to clarify that cloudy or smelly urine is not indicative of UTI. Other aspects that needed improvement included:

— To prompt providers to look for other potential causes of non-specific changes, such as loss of consciousness, agitation, lethargy, and falls in older adults.

— To improve processes, to ensure that test results follow the patient during the hospital journey.

— To change the laboratory requisition to include the reason for urine testing.

Participants identified that the time it took to collect and receive the lab results, particularly in the emergency department, are barriers to appropriate assessment. In emergency departments, wait times and length of admission are considered benchmarks of quality. Other barriers were:

- Delirium in older patients is perceived as a UTI symptom; this is an area of contention. Physicians are afraid of missing something.
- Lack of information about urinary catheters, particularly regarding whether the catheter is new, intermittent, or chronic.
- Difficulties obtaining a quality urine sample.

Process mapping helped us to understand the decision-making process clinicians followed. It led us to uncover current practices, collectively identify areas for improving information, and at the same time, recognize barriers to appropriate care.

Dialogue prompters: Seeing and thinking through the problem

The dialogue prompter fostered co-analysis of the provotypes and prototypes (Figure 3) by drawing on participants' unique expertise as part of a shared problem space. Rather than quickly focusing on the solution, "the dialogue process attempts to slow down the conversation to allow participants to reflect" (Schein, 2017, p. 111). The prompter provided a visual representation for discussing problems clinicians faced.

Participants pointed out that the decision aid was not entirely about ASB; it was also about appropriate urine testing. They felt the tool needed to specify the intended audience. As a consequence, we added a heading to help users quickly identify the intended readers.

Participants found that the decision aid contained two types of information: content to guide decision-making and content to educate healthcare providers. These types of content served competing purposes. On the one hand, content for decision-making encouraged users to draw on working memory and act. On the other, content for learning was intended to activate users' long-term and working memory – triggering users' prior knowledge, fostering reflection, and promoting the integration of old and new knowledge. Because participants felt that content directed at learning did not suggest immediate action, they recommended that this educational content be placed elsewhere in the toolkit. We moved the educational information to a "physician tool" to focus the redesign exclusively on content for decision-making.

Participants also indicated that not all of the content presented in the algorithm's decision boxes actually guided them to a satisfactory decision. For example, the advice "Hydrate and reassess in 24 hours" was problematic for the emergency department clinicians, given that the patient might not stay there for 24 hours.

The advice to assess changes in a patient's mental state created resistance. Clinicians felt that asking emergency department personnel to assess mental states negatively influenced their willingness to use the algorithm. This led us to remove the advice.

Prototypes: Recognizing fears and beliefs about diagnostic decisions

The visual nature of the prototypes created an enabling context that helped to recognize fears and beliefs related to diagnosing UTI. Parts of the decision were revised to alleviate participants' reservations about the guidance. For example, participants suggested that the advice relating to 'No UTI symptoms' was problematic because physicians may have fears of missing a diagnosis of sepsis (a life-threatening condition). As Collins (2018) explained, "when humans are given a choice between their own judgment and that of a demonstrably superior algorithm, they will generally choose the former" (Paragraph 9). It became apparent that the problem was not merely to guide decision-making but to show that the process we were proposing would *lead to safer outcomes*.

Three sets of requirements for the decision aid

Results from the co-design workshops provided a basis upon which to develop performance requirements. Co-design also guided our understanding of how to organize the content so that it followed a typical decision path a healthcare provider might pursue. We summarize these requirements in Table 2.

Uncovering clinicians' needs through iterative co-design

Participants thought that the title of the algorithm needed to communicate that the advice was, in fact, evidence-based. Consequently, the title went through a series of iterations – from "Urinary Tract Infection (UTI) Acute Care Algorithm" – to "Evidence-Based Criteria for Urinary Infection Testing."

Participants also wanted a visible and prominent sign that reminded users that cloudy and smelly urine is not indicative of UTI. We added a warning sign to address their concern.

In the first co-design workshop, participants stated that to match the users' mental process, the tool needed to start with the patient. They felt the algorithm should present different scenarios the healthcare team considers when diagnosing UTI and ASB. Participants suggested organizing the algorithm into four main groups: sepsis, UTI symptoms, no UTI symptoms, and older adults with no symptoms.

Table 2. Requirements for the decision aid.

User requirements	1. It should be easy to identify the intended audience. 2. It should work for both emergency and acute care clinicians. 3. It should avoid resistance and promote adoption. 4. It should not promote other decision tools.
Writing requirements	5. Reading-to-do and reading-to-learn should be separated. 6. The title should clearly state that the decision aid is based on evidence. 7. The title should clearly communicate the function of the tool. 8. It should use terms and language familiar to clinicians (Schriver, 2017). 9. It should present only necessary information and avoid being text-heavy. 10. Advice should be clear, for example: Do not test urine for changes in color or smell. 11. It should clearly indicate to send urine for testing only if there are UTI symptoms. 12. In the case of delirium, it should indicate the patient's need for hydration.
Structure requirements	13. The reading order should be very apparent to users: left to right, or top to bottom. 14. It should be easy to identify the different scenarios: symptoms, no symptoms, others. 15. Every box should lead to a decision. 16. The box for sepsis should be part of the main symptoms to consider.

thought it would be more effective to have two main groups: patients presenting UTI symptoms and patients not presenting UTI symptoms, as shown in the bottom right prototype of Figure 6.

Twenty-two iterations were made before arriving at the final prototype. Clinicians were involved in all iterations (as co-design participants or as members of the clinical committee). The final document was approved by the ASAB Steering Committee, the group responsible for oversight of the content. Figure 6 shows the main iterations of the decision aid over the co-design process. Figure 7 presents the final prototype.

Figure 6.

Transformations of the decision aid—from original to provotype to prototypes.

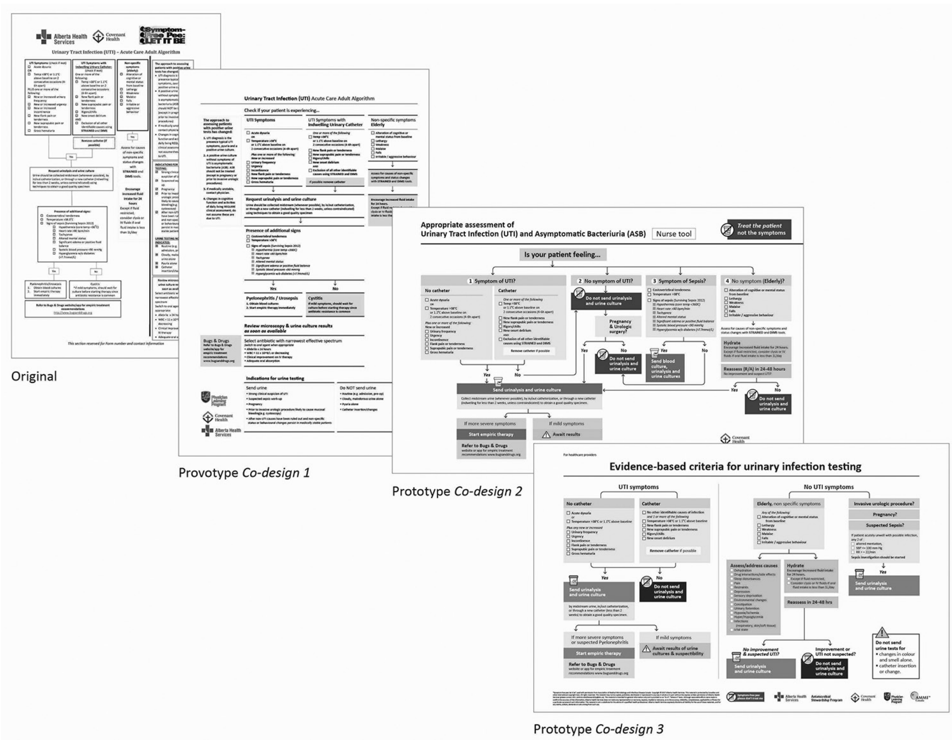
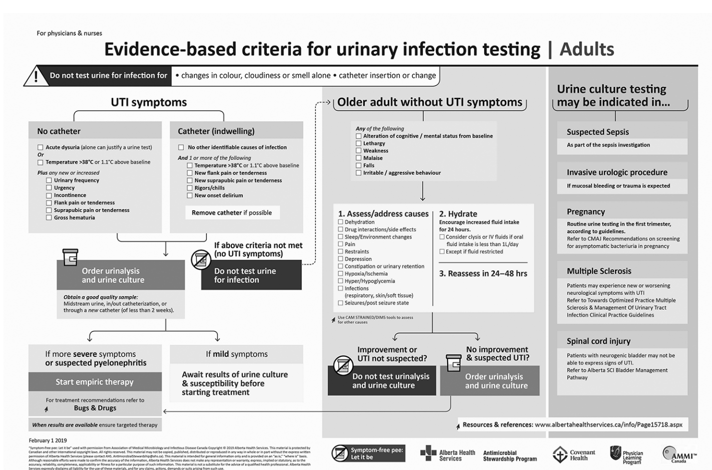


Figure 7.

The final prototype of the decision aid.



4. Discussion

Our findings show that co-design both helped participants navigate the problem space and enabled them to develop a human-centered decision aid. But our research also led us to realize that the goals we were pursuing: to examine the problems associated with assessing and treating UTI and ASB and the evidence regarding when to order urine testing, was not always viewed by clinicians as a problem. We found that some participants needed to accept the evidence before they could believe that "overusing urine testing" was even an issue. As Lefebvre (2013) points out, before one offers a solution to a problem, people need to be aware of it and embrace the idea that it actually needs to be improved.

Generating problem awareness through co-design

As we collaborated with participants over the course of the study, we grew to recognize that even if we developed a good evidence-based tool, clinicians might still resist adopting the decision aid because their prior experience might conflict with recent evidence. Some participants did not agree with what the evidence said the "facts" were. For example, the idea that "cloudy or smelly urine is not necessarily indicative of UTI" was a recurring point of discussion—even of contention among some participants. As one nurse said as she read our quotes on the wall, "This is not true." Another participant added, "This is not what I was taught in school." As participants listened to the experiences of the group, they gradually developed an awareness that it was a genuine problem and eventually bought into the idea that it needed improvement.

Building shared vocabulary during co-design

We identified that the language of UTI and ASB was foreign to some participants. As one participant pointed out, "I don't even know how to pronounce *asymptomatic bacteriuria*." We saw that what was straightforward to experts could alienate those trying to improve clinical practice. This acknowledgment led us to avoid using unfamiliar terms. It also made us realize that reflecting on one's practices requires using language to explore what one knows and does not know (for how humans operate in a semantic domain shaped by languages, see Maturana & Varela, 1987, pp. 211-213). We concluded that language could be an impediment to co-design if all participants do not share similar understandings of the key terms being used during discussion. In fact, a lack of shared vocabulary can hinder productive engagement in co-design (Von Krogh et al., 2000, p. 118).

Drawing on external representations to promote co-design

We found that when participants collaborated through discussing an external representation of the problem (mapping practices), they were often prompted to move beyond personal experience. The tools (i.e., dialogue prompters) encouraged participants to reflect on their own practices, attend to the experiences of others, and jointly consider what could be done differently. For example, using the provotype led some participants to argue that it would be useful to know how long urine samples wait in the emergency department before being sent to the lab for analysis. The provotype served to catalyze ideas about what is typically done and allowed participants to better conceptualize ways to improve the process.

External representations offered participants concrete visual artifacts to respond to. As Larkin and Simon (1987) observed, diagrammatic representations allow people to recognize features and make inferences about information that may be otherwise unavailable to them.

We found that visual representations of the problem played a key role in constructing and modifying understanding and in exploring possibilities. Planning for effective collaboration involved thinking about what could emerge during each session, imagining where conversations might lead and how the tools might drive interaction and collective problem solving. Designers have been shifting their role—from facilitator to orchestrator—anticipating what could happen (Aguirre, Agudelo, & Romm, 2017). Facilitators organize and lead a workshop. Orchestrators anticipate the flow of the collective event and “materialize their intent through contextually designed facilitation tools” (Aguirre, Agudelo, & Romm, 2017, p. 207). When designing conversations, there is uncertainty in the collective learning process; dealing with uncertainty requires planned flexibility.

Benefitting from collective learning during co-design

Our co-design process led to an unexpected and favorable outcome. We found that as participants conversed with one another and became more aware of the evidence about UTI and ASB, their resistance to messages that ran counter to their personal experience lessened. By working together on the decision aid, participants who struggled with believing in the evidence changed their minds. Participants who challenged recent “facts” became champions of those newly acquired facts and proved to be instrumental in fostering the adoption of safer care processes.

During the workshops, participants shared situations and problems they experienced that helped group members re-consider their current understanding of UTI and ASB. Over time, their collective knowledge grew more integrated, and their individual knowledge was reconfigured (Huang & Yang-Chieh 2018). The exchange of beliefs, experiences, and medical evidence through "designed conversations" fostered collective group teaching and learning. As Zhao and colleagues (2004) elaborate, "a group [who]…work together demonstrate their common and inter-personal knowledge, while the members of the recipient community share, integrate and synthesize their learning among themselves" (p. 139). During the workshops, for example, participants made comments such as: "It's a nice way of learning," "It's a place to address knowledge gaps," and "It helped me to think critically by having a conversation."

That participants acknowledge knowing more about UTI and ASB as a result of co-design showed that learning can be a consequence of a robust co-design process. As Larson and Christensen (1993) suggest, "the knowledge that something was learned can be quite valuable because it provides a cue for ensuring that what was learned actually came out in discussion" (p. 17). Co-design activities can help teams translate and share their knowledge (Greenhalgh, 2018; Erwin & Krishnan, 2017; Sanematsu & Cripe, 2017; Langley et al., 2018).

The ASAB initiative materials are continuously promoted by the organization's website [see previously shared URL] and as part of ongoing clinicians' education. The tools were used as a pilot education process at a large urban hospital emergency department, which consisted of a blitz of 1:1 education of emergency department nurses, pharmacists, and physicians. After the introduction of the tools, urine culture rates decreased by 17% and were sustained for at least a one-year period.

The tools were also used as part of a quality improvement project at geriatric psychiatric units located in a psychiatric hospital. Coupled with staff and physician education, implementation of the evidence-based tools which resulted in reductions in urine cultures (34%), antibiotic prescriptions (28%), and antibiotic utilization (27%). The toolkit has been promoted by the Association of Medical Microbiology and Infectious Disease of Canada, and Choosing Wisely Canada.

6. Implications & conclusions

This case study shows that co-design can be an effective method toward improving healthcare practices—helping teams to exchange knowledge and develop human-centered tools that work. Even though designers can play a key role in tapping team members' tacit knowledge and in transforming that knowledge into useful tools for particular audiences, the role of the designer in orchestrating co-design is not well understood. Future research should address this knowledge gap. Too often, designers in healthcare are perceived as professionals who make information pretty—who tidy things up at the end of thinking—rather than as interpreters of complex problems, who can help in implementing new evidence and in humanizing healthcare systems.

A key role of the designer when facilitating is empowering participants to determine the nature of their assumptions behind beliefs and processes, to foster dialogue, reflection, and listening.

> The key to initiating this kind of dialogic conversation is to create a setting in which participants feel secure enough to suspend their need to win arguments, to clarify everything they say, and to challenge each other every time they disagree (Schein, 2017, p. 111).

One of the key roles of the designer in this project was to create a safe space for dialogue, knowledge sharing, collective learning, and collaboration to find out how to improve diagnostic testing and treatment of UTI and ASB to provide better care. "Improving the quality of care delivery processes necessarily requires different viewpoints, each grounded in deep knowledge of a different aspect of the process" (Nembhard & Edmondson, 2006, p. 943). This was achieved through the development of activities and tools thought to engage overworked clinicians in practice reflection and exploration, an iterative process of questioning and re-thinking.

This study allowed us to draw a number of implications about applying co-design to effectively engage busy clinicians in practice reflection, learning, and improvement. We suggest the following:

Frame the interaction and set expectations

1. Invite people with diverse disciplinary backgrounds, professional skills, roles, working cultures, and responsibilities to participate.

2. Ensure the co-design goals are relevant to participants. Connect the goals with what participants bring to the table. If known, articulate the human's point of view, not just the technical.

3. Take the rhetorical stance of asking participants for their help. Show that you value their diverse perceptions about the problem. Ensure all participants take part in formulating alternative solutions.

4. Design activities to structure participants' co-identification of the problem, its conceptualization, and co-design of the tool.

Frame the problem, facilitate collective thinking and problem exploration through visual reasoning

5. Present the perception of the problem in a clear and precise manner. Use plain language and avoid ambiguity.

6. Encourage participants to reflect on what, how, when, and why something occurs. Use a process map to investigate this.

7. Inspire participants to discuss options and confront alternative ideas. Give them a sense of the power of co-analysis.

8. Use methods that focus participants on discussing an external representation of the problem, for example, dialogue prompters.

9. Ensure that the route to reason through the problem is workable. When appropriate, offer explicit procedures the group can take.

Be aware of your stance: sense, listen, learn, and orchestrate

10. Structure co-design activities so participants can work at a reasonable pace. Create opportunities for reaching milestones along the way.

11. Make human-centered design a guiding principle in how you interact with participants. Try not to lead the group, but orchestrate it.

When designers effectively structure and implement co-design in healthcare settings, clinicians may not only change their ideas about the value of designers, but they may also more fully recognize the human in human-centered design when devising healthcare solutions. Human-centered design is moving beyond making things easier to use. As Whitney and Nogueira (2020) elaborated, "Imagine that design included broader dimensions of what makes us human, such as happiness and health, in its frameworks and methods" (p. 149).

The perspective of the lead physician on this project highlights the value of seeing and becoming aware, of immersing in new contexts that matter (Scharmer, 2018, p. 84), "Observing how users interacted with the prototypes was very illuminating. The developing group and content experts had not anticipated some of the issues that created confusion in users or how some phrases and concepts were being understood. Observing the sessions also illuminated ways in which the algorithm had to be adapted to fit different healthcare scenarios and to address other educational needs. It was clear that what we thought we were saying and what they were actually reading and understanding were much farther apart than we ever would have guessed. In retrospect, I am now nervous about how other clinical guidance documents and tools I have worked on are being interpreted!"

This study adds to the body of knowledge of the role of human-centered design to achieve evidence-based healthcare practices. The knowledge we gained about the overuse of urine testing and unnecessary antibiotic prescriptions allowed us to take action on a pervasive problem of inpatient care with measurable impacts on clinical practice. We hope this knowledge on the role of co-design in evidence implementation for clinical practice improvement will be of value to other teams of clinicians and designers.

References

Abbo, L. M., & Hooton, T. M. (2014). Antimicrobial stewardship and urinary tract infections. *Antibiotics*, *3*(2), 174–192.

Adam, M., McMahon, S. A., Prober, C., & Bärnighausen, T. (2019). Human-centered design of video-based health education: an iterative, collaborative, community-based approach. *Journal of medical Internet research, 21*(1), e12128.

Aguirre, M., Agudelo, N., & Romm, J. (2017). Design facilitation as emerging practice: Analyzing how designers support multi-stakeholder co-design. *She Ji: The Journal of Design, Economics, and Innovation, 3*(3), 198–209.

American College of Radiology. (2012). Five things physicians and patients should question. *The Journal of the Oklahoma State Medical Association, 105*(12), 482.

Barlam, T. F., Cosgrove, S. E., Abbo, L. M., MacDougall, C., Schuetz, A. N., Septimus, E. J., ... & Hamilton, C. W. (2016). Implementing an antibiotic stewardship program: guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. *Clinical Infectious Diseases, 62*(10), e51–e77.

Beckford-Ball, J. (2006). Management of suspected bacterial urinary tract infection. *Nursing Times*, *102*(32), 25–26.

Betsch, Cornelia, Constanze Rossmann, Mathias W. Pletz, Horst C. Vollmar, Antje Freytag, Ole Wichmann, Regina Hanke et al. (2018). Increasing influenza and pneumococcal vaccine uptake in the elderly: study protocol for the multi-methods prospective intervention study Vaccination60+. *BMC Public Health, 18*(1), 885–889.

Blakiston, M., & Zaman, S. (2014). Nosocomial bacteriuria in elderly inpatients may be leading to considerable antibiotic overuse: an audit of current management practice in a secondary level care hospital in New Zealand. *Infection and Drug Resistance, 7*, 301–308.

Boer, L., & Donovan, J. (2012). Provotypes for participatory innovation. *Proceedings of the designing interactive systems conference*, New York, NY: ACM, 388–397. https://dl.acm.org/doi/10.1145/2317956.2318014

Breslin, M., Mullan, R. J., & Montori, V. M. (2008). The design of a decision aid about diabetes medications for use during the consultation with patients with type 2 diabetes. *Patient education and counseling, 73*(3), 465–472.

Britton, G. (2017). *Co-design and social innovation: Connections, tensions and opportunities.* New York, NY: Taylor & Francis.

Collins, J. (2018). What to do when algorithms rule. *Behavioral Scientist.* http://behavioralscientist.org/what-to-do-when-algorithms-rule/?utm_source=feedly&utm_medium=webfeeds

Erwin, K., & Krishnan, J. A. (2016). Using design methods to provide the care that people want and need. *Journal of Comparative Effectiveness Research,* 5(1), 13–15.

Erwin, K., & Krishnan, J. A. (2017). Welcome to the beginning. *Information Design Jour*nal, *23*(3), 357–359.

File Jr, T. M., Srinivasan, A., & Bartlett, J. G. (2014). Antimicrobial stewardship: importance for patient and public health. *Clinical Infectious Diseases, 59*(3), S93–S96.

Finucane, T. E. (2017). Urinary tract infection"—requiem for a heavyweight. *Journal of the American Geriatrics Society, 65*(8), 1650–1655.

Fleming-Dutra, K. E., Hersh, A. L., Shapiro, D. J., Bartoces, M., Enns, E. A., File, T. M., ... & Lynfield, R. (2016). Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011. *JAMA, 315*(17), 1864–1873.

Flokas, M. E., Andreatos, N., Alevizakos, M., Kalbasi, A., Onur, P., & Mylonakis, E. (2017). Inappropriate management of asymptomatic patients with positive urine cultures: A systematic review and meta-analysis. *Open Forum Infectious Diseases, 4*(4), ofx207. https://doi.org/10.1093/ofid/ofx207

Flores, F. (2012). *Conversations for action and collected essays: Instilling a culture of commitment in working relationships*. CreateSpace Independent Publishing Platform.

Formoso, G., Paltrinieri, B., Marata, A. M., Gagliotti, C., Pan, A., Moro, M. L., ... & Magrini, N. (2013). Feasibility and effectiveness of a low-cost campaign on antibiotic prescribing in Italy: community level, controlled, non-randomised trial. *BMJ,* 347, f5391.

Frascara, J. (Ed.). (2006). *Designing effective communications: Creating contexts for clarity and meaning.* New York, NY: Allworth Press.

Frascara, J., & Noël, G. (2010). Evaluation and design of a blood components transfusion request form. *Information Design Journal, 18*(3), 241–249.

Frascara, J. (2017). Design, and design education: How can they get together? *Art, Design & Communication in Higher Education,16*(1), 125–131.

Garvelink, M. M., Emond, J., Menear, M., Brière, N., Freitas, A., Boland, L., ... & Légaré, F. (2016). *Research Involvement and Engagement, 2*(1), 26–42. https://doi.org/10.1186/s40900-016-0040-0

Greenhalgh, T., Jackson, C., Shaw, S., & Janamian, T. (2016). Achieving research impact through co-creation in community-based health services: Literature review and case study. *The Milbank Quarterly, 94*(2), 392-429.

Greenhalgh, T. (2018). *How to implement evidence-based healthcare*. UK: John Wiley & Sons.

Gould, C. V., Umscheid, C. A., Agarwal, R. K., Kuntz, G., & Pegues, D. A. (2009). Healthcare infection control practices advisory committee (HICPAC). Guideline for prevention of catheter-associated urinary tract infections. *Healthcare Infection Control Practices Advisory Committee,* Atlanta. https://stacks.cdc.gov/view/cdc/49910

Hargraves, I. (2018). Care and capacities of human-centered design. *Design Issues,* 34(3), 76–88.

Hartson, R., & Pyla, P. S. (2012). *The UX book: Process and guidelines for ensuring a quality user experience*. Waltham, MA: Elsevier.

Huang, Y. C., & Chin, Y. C. (2018). Transforming collective knowledge into team intelligence: the role of collective teaching. *Journal of Knowledge Management, 22*(6), 1243–1263.

Juthani-Mehta, M. (2015). Changing clinicians' behavior: To order or not to order a urine culture. *JAMA internal medicine, 175*(7), 1127–1129.

Langley, J., Wolstenholme, D., & Cooke, J. (2018). 'Collective making' as knowledge mobilisation: the contribution of participatory design in the co-creation of knowledge in healthcare. *BMC health services research, 18*(1), 585.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–99.

Larson, J. R., & Christensen, C. (1993). Groups as problem-solving units: toward a new meaning of social cognition. *British Journal of Social Psychology,* 32(1), 5–30.

Lefebvre, R. C. (2013). *Social marketing and social change: Strategies and tools for improving health, well-being, and the environment*. San Francisco, CA: John Wiley & Sons.

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human Understanding.* Boston, MA: Shambhala Press.

Meyer, M. W., & Norman, D. (2020). Changing design education for the 21st century. *She Ji: The Journal of Design, Economics, and Innovation, 6*(1), 13–49.

Meyer, M. W., & Norman, D. (in press). From DT to DD via DDT: Design Thinking, Design Doing, and Design-Driven Transformation. *Journal of Product Innovation Management.*

Nembhard, I. M., & Edmondson, A. C. (2006). Making it safe: The effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams. *Journal of Organizational Behavior, 27*(7), 941–966.

NHS Institute for Innovation and Improvement (Great Britain). (2005). Managing the Human Dimensions of Change. *NHS Institute for Innovation and Improvement*. https://www.england.nhs.uk/improvement-hub/publication/improvement-leaders-guide-managing-the-human-dimensions-of-change-personal-and-organisational-development/

Nicolle, L. E., Gupta, K., Bradley, S. F., Colgan, R., DeMuri, G. P., Drekonja, D., ... & Siemieniuk, R. (2019). Clinical practice guideline for the management of asymptomatic bacteriuria: 2019 update by the Infectious Diseases Society of America. *Clinical Infectious Diseases, 68*(10), e83-e110.

Noël, G., Luig, T., Heatherington, M., & Campbell-Scherer, D. (2018). Developing tools to support patients and healthcare providers when in conversation about obesity: The 5As Team program. *Information Design Journal, 24*(2), 131-150.

Noël, G., Frascara, J., & Wong, C. (2019). Designing bowel preparation patient instructions to improve colon cancer detection: Evidence-based design criteria for patients' documents. *Information Design Journal, 25*(1), 110 –121.

Plaisance, A., Witteman, H. O., Heyland, D. K., Ebell, M. H., Dupuis, A., Lavoie-Bérard, C. A., ... & Archambault, P. M. (2016). Development of a decision aid for cardiopulmonary resuscitation involving intensive care unit patients' and health professionals' participation using user-centered design and a wiki platform for rapid prototyping: a research protocol. *JMIR Research Protocols, 5*(1), e–24.

Roberts, J. P., Fisher, T. R., Trowbridge, M. J., & Bent, C. (2016). A design thinking framework for healthcare management and innovation. *Healthcare 4*(1), 11–14).

Rowe, T. A., & Juthani-Mehta, M. (2014). Diagnosis and management of urinary tract infection in older adults. *Infectious disease clinics of North America,* 28(1), 75.

Ragouzeos, D., Gandrup, J., Berrean, B., Li, J., Murphy, M., Trupin, L., ... & Schmajuk, G. (2019). "Am I OK?" using human centered design to empower rheumatoid arthritis patients through patient reported outcomes. *Patient education and counseling, 102*(3), 503-510.

Rumelhart, D. E. (1992). Towards a microstructural account of human reasoning. In *Connectionism: Theory and practice*, S. Davis, (Ed.), 69–83. New York, NY: Oxford University Press.

Sanders, E., & Stappers, P. J. (2008). Co-design and the new landscapes of design. *Co-design, 4*(1), 5–18.

Sanders, E. B. N., & Stappers, P. J. (2012). *Convivial design toolbox: Generative research for the front end of design.* BIS Publishers.

Sanematsu, H., & Cripe, L. D. (2017). The impact of design on research teams in health services. *Information Design Journal, 23*(3), 310–318.

Scharmer, O. (2018). *The essentials of theory U: Core principles and applications*. Berrett-Koehler Publishers.

Schein, E. H. (2017). *Organizational culture and leadership.* John Wiley & Sons.

Schriver, K. A. (2017). Plain language in the United States gains momentum: 1940–2015. *IEEE Transactions on Professional Communication, 60*(4), 343–383. http://ieeexplore.ieee.org/document/8115322/

Scottish Intercollegiate Guidelines Network (SIGN). 2012. *Management of suspected bacterial urinary tract infection in adults*. Edinburgh: SIGN; 2012. (SIGN publication no. 88). http://www.sign.ac.uk

Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization*. Random House.

Soong, C., Lee, T. C., Leis, J. A., Okrainec, K., & McDonald, E. G. (2016). A point prevalence study of urinary catheter use among teaching hospitals with and without reduction programs. *Journal of Hospital Medicine, 11(*11), 799–800.

Thornhill, B., Potter, E., Nakarada-Kordic, I., & Reay, S. (2017). Designing clinical guidelines to support effective decision-making in emergency surgery. *Information Design Journal, 23*(3), 268–289.

Von Krogh, G., Ichijo, K., & Nonaka, I. (2000). *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*. Oxford University Press on Demand.

Voûte, E., Stappers, P. J., Giaccardi, E., Mooij, S., & van Boeijen, A. (2020). Innovating a large design education program at a university of technology. *She Ji: The Journal of Design, Economics, and Innovation, 6*(1), 50–66.

Whitney, P., & Nogueira, A. (2020). Cutting cubes out of fog: The whole view of design. *She Ji: The Journal of Design, Economics, and Innovation, 6*(2), 129–156.

Witteman, H. O., Dansokho, S. C., Colquhoun, H., Coulter, A., Dugas, M., Fagerlin, A., ... & Ivers, N. (2015). User-centered design and the development of patient decision aids: Protocol for a systematic review. *Systematic Reviews, 4*(1), 11.

Zhao, Z., Anand, J., & Mitchell, W. (2004). Transferring collective knowledge: teaching and learning in the Chinese auto industry. *Strategic Organization, 2*(2), 133–167.

Zeballos-Palacios, C. L., Hargraves, I. G., Noseworthy, P. A., Branda, M. E., Kunneman, M., Burnett, B., ... & Montori, V. M. (2019). Developing a conversation aid to support shared decision making: Reflections on designing anticoagulation choice. *Mayo Clinic Proceedings. 94*(4*)*, 686–696.

Authors

Guillermina Noël

Dr. Guillermina Noël holds a PhD in Design Sciences (University IUAV of Venice, Italy). She applies a human-centred, evidence-based, and outcomes-oriented design approach to improve healthcare. At the Physician Learning Program in the Faculty of Medicine & Dentistry, University of Alberta, Canada, Guillermina supported healthcare teams by helping them to move beyond problem's symptoms, understand the people's capacities and needs, and co-designing to achieve better health outcomes. She is also the Director of the Health Design Network. Since 2019, she is the Head of Design Management International programme at the Lucerne University of Applied Sciences and Arts.

Darren Pasay

Darren Pasay is the Drug Stewardship Pharmacist for Alberta Health Services, Central Zone. His interests in Antimicrobial Stewardship, geriatrics, knowledge translation, transformation change and research lead him into the role as the Project Lead for the Appropriateness of Care: Asymptomatic Bacteriuria project. He previously spent 13 years as a clinical pharmacist and pharmacy manager in rural hospitals in Alberta and 3 years as a Clinical Associate Professor at the Faculty of Pharmacy and Pharmaceutical Sciences at the University of Alberta.

Denise Campbell-Scherer

Dr. Denise Campbell-Scherer is a Professor in the Department of Family Medicine, and the Associate Dean of the Office of Lifelong Learning and Physician Learning Program at the University of Alberta. With a background in evidence-based clinical practice, she has been active internationally in education of multidisciplinary learners and is an Associate Editor of BMJ-Evidence-Based Medicine (http://ebm.bmj.com).Dr. Campbell-Scherer's research leads and interdisciplinary research team focused on innovations to advance clinical practice.

Lynora Saxinger

Dr. Lynora Saxinger is an academic Infectious Diseases specialist, her career outside patient care is focused on antimicrobial stewardship and antimicrobial resistance, with involvement in hospital, health system and national collaborations to improve antibiotic stewardship practice, resistance and utilization surveillance. She is a founding member and past chair of the Antimicrobial Resistance and Stewardship Committee of the Association of Medical Microbiology and Infectious Diseases Canada. She is Associate Professor of Medicine, University of Alberta, and Medical Lead of Antimicrobial Stewardship for Northern Alberta, with Alberta Health Services.