

Voice Driven Type Design

Matthias Wölfel • Tim Schlippe • Angelo Stitz



Information Loss

- × The richness of verbal communication is lost in text based communication.



Idea

- ✘ While the shape of a single typographical character has been treated as an unchangeable property until today, we let the shape of each single character adjust according to particular acoustic features in the spoken reference.





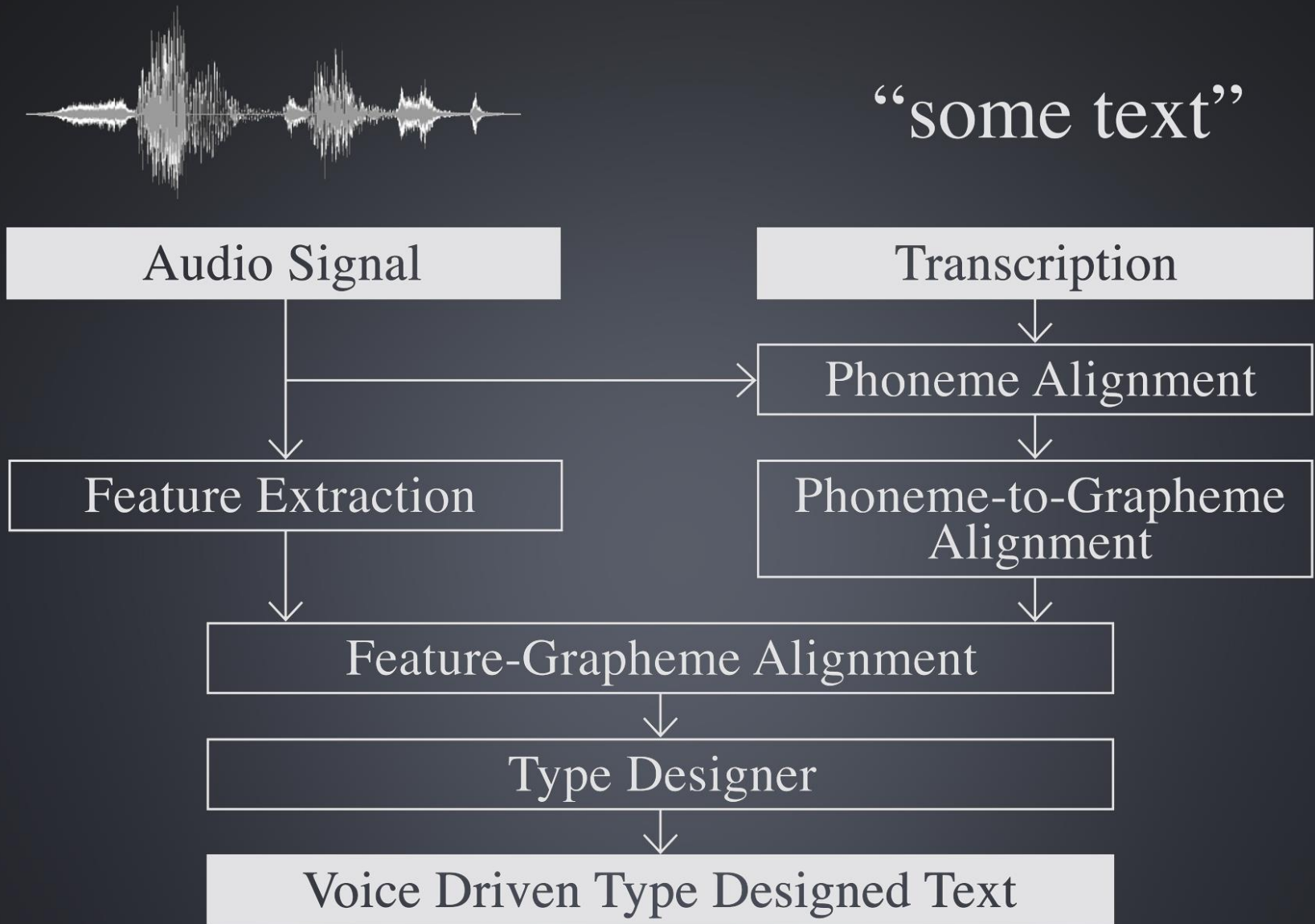
Voice Driven Type Design

- × Because the characteristics in the voice shape the design of each typographic character we call this approach *voice driven type design* (VDTD).

Using typography as a
stylistic device
has a very long tradition.

It can be:

- ◆ **static** as presented in printing books, posters and comics, where type represents content in a uniform and permanent way
- ◆ **dynamic** as in *kinetic typography* which is an animation technique to express ideas using text-based video animation
- ◆ **reactive** as in *responsive type* where the shape of each letter is adjusted according to the properties of the reader



some text

But which
property of the acoustic signal
should be mapped to which
property of the typographical character?

Loudness

- ◆ Producing loudness in speech amplifies the signal and is used to have the attention of a listener.
- ◆ To have the attention of the reader, bolder text is commonly used.
- ◆ To recognize each acoustic feature separately after the mapping on its visual representation, we decided to increase only the vertical stroke weight.

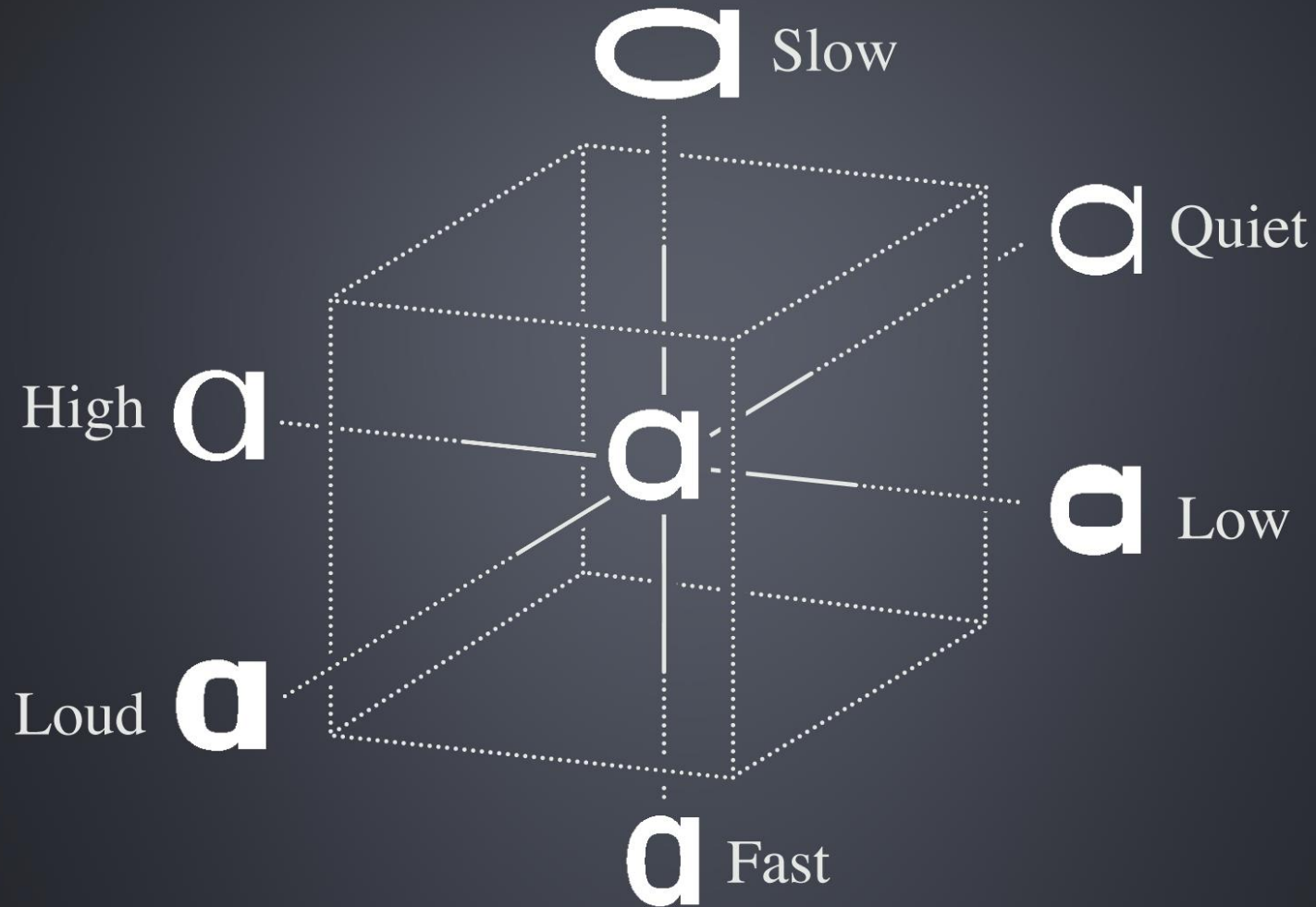
Pitch

- ◆ High pitch levels draw attention and express emotions.
- ◆ Medium pitch levels account for more neutral attitudes.
- ◆ We adapt the horizontal stroke weight depending on the pitch level since this modification increases the reader's curiosity.

Speed

- ◆ A reader usually jumps from a part of a word to a next part of a word.
- ◆ Increasing the character width extends this scanning process of the eyes.
- ◆ Therefore, we map the speed of the utterance to the character width.

Mapping of the Parameters



① Sei ruhig bleibe ruhig mein Kind

② Sei ruhig bleibe ruhig mein Kind

 Higher speed

③ Sei ruhig bleibe ruhig mein Kind

④ Sei ruhig bleibe ruhig mein Kind

 Less loudness

⑤ Sei ruhig bleibe ruhig mein Kind

⑥ Sei ruhig bleibe ruhig mein Kind

 Higher pitch

⑦ Sei ruhig bleibe ruhig mein Kind

⑧ Sei ruhig bleibe ruhig mein Kind

What possible
applications
can profit from the proposed approach?

atención

Language learning and speech-language pathology

Many people learning a new language have trouble doing the intonation and accentuation in a right way. A potential of VDTD is that it allows to illustrate intonation and accentuation in software and books for language learners.

Hints for deaf people

Deaf people profit from VDTD since it gives them hints on how something was spoken.

This is helpful in different situations:

- ◆ learning how to pronounce
- ◆ interpreting the intention

Hints for dyslexia

A reading disorder is primarily influenced by the so-called phonological awareness.

People with an unsatisfactory phonological awareness are not able to extract the correct orthographic word out of a spoken utterance.

VDTD offers a new way to visualize a comprehensible relationship between the spoken utterance and the written text.

Subtitles

Subtitles on television screens are very popular in places with either a lot of background noise or where different programs are simultaneously broadcasted.



Texting

People love texting – in the static text of the transcription, meta-information is lost and no emoticons help the receiver. VDTD can support such a scenario.





The image shows a musical staff in 4/4 time with a treble clef. The melody consists of the following notes: a quarter note G4, a quarter note A4 with a sharp sign, a quarter note B4, a quarter note C5, a half note D5, and a whole note E5. The lyrics 'We will we will rock you' are written below the notes, with 'We' under G, 'will' under A, 'we' under B, 'will' under C, 'rock' under D, and 'you' under E.

We will we will rock you

Singing and Karaoke

Voice-driven type designed songbooks can support singing songs without the requirement of being able to read notes since *loudness*, *pitch* and *speed* are illustrated in a more comprehensible form and provides additional information to those who are able to read and interpret notes.

We have
evaluated

our approach by asking 9 persons
(6 male, 3 female) various questions.

Acceptance

The question if VDTD is interesting for several applications has been answered by most participants positively – a result between strongly agree and agree.

Representation of Voice Characteristics

Subjects listened to one recording of the poem and graded how good each verse of five texts represents the loudness, pitch and speed of the recording.

Acoustic Feature	Approach				
	<i>VDTD</i>	<i>R1</i>	<i>R2</i>	<i>HTD</i>	
Loudness	2,46	2,71	2,81	3	
Pitch	2,46	2,58	2,4	3,1	
Speed	2,58	2,56	2,83	3,23	
Average	2,5	2,62	2,68	3,11	

Representation of Speakers' Characteristics

To determine if the visualization representation preserves individual characteristics of a speaker, standard deviations are calculated, between utterance pairs.

Acoustic Feature	Standard Deviation	
	Same Speaker	Different Speaker
Loudness	0.028	0.040
Pitch	0.032	0.045
Speed	0.019	0.027

