Sing2Text

Concept for an instant messenger leveraging the voice for expressive typography

Graeme Zinck

Asynchronous messaging

Text messages

Voice messages

Asynchronous messaging

Text messages

- Easy to revisit old messages

Voice messages

- Difficult to revisit old messages

Asynchronous messaging

Text messages

- Easy to revisit old messages
- Hands and eyes needed

Voice messages

- Difficult to revisit old messages
- Hands- and eyes-free

Asynchronous messaging

Text messages

- Easy to revisit old messages
- Hands and eyes needed
- Expression through emojis & typographic properties

Voice messages

- Difficult to revisit old messages
- Hands- and eyes-free
- Expression through vocal properties

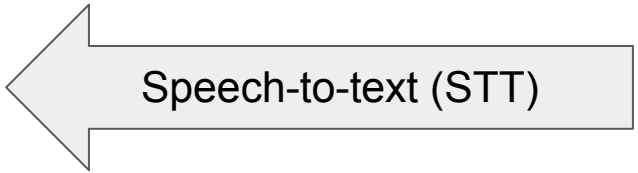
Asynchronous messaging

Text messages

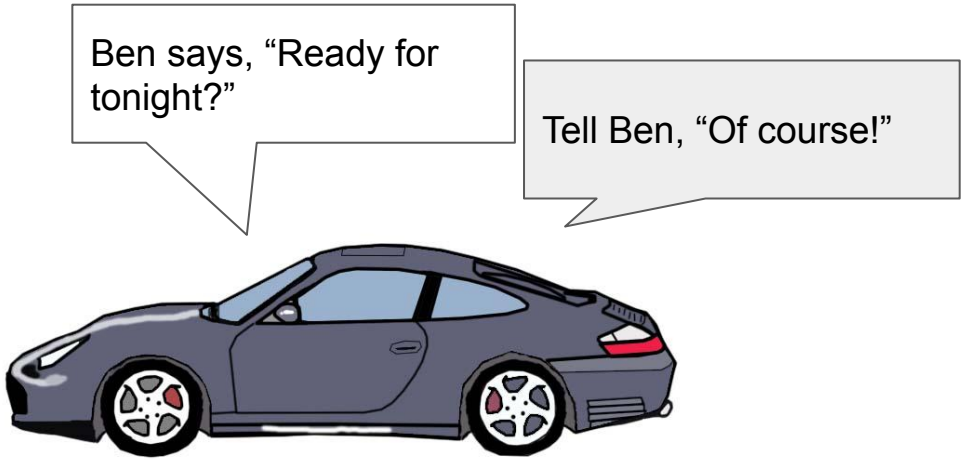
- Easy to revisit old messages
- Hands and eyes needed
- Expression through emojis & typographic properties

Voice messages

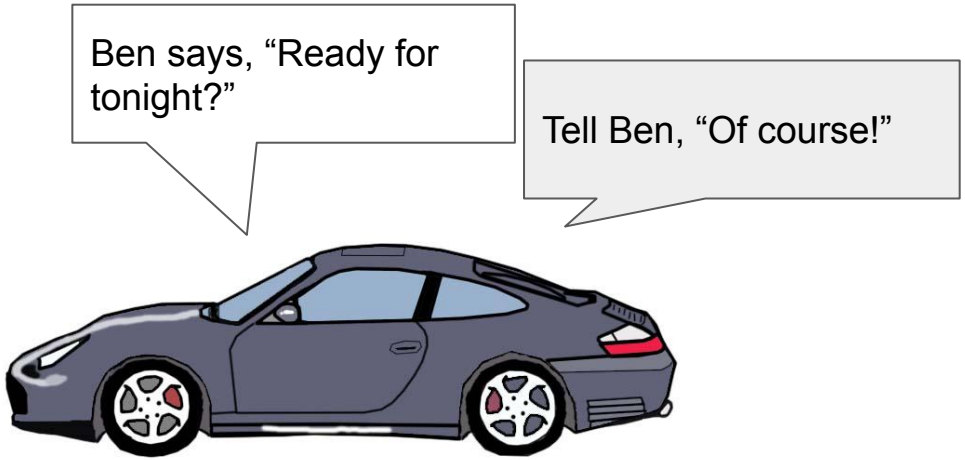
- Difficult to revisit old messages
- Hands- and eyes-free
- Expression through vocal properties



Sing2Text



Sing2Text

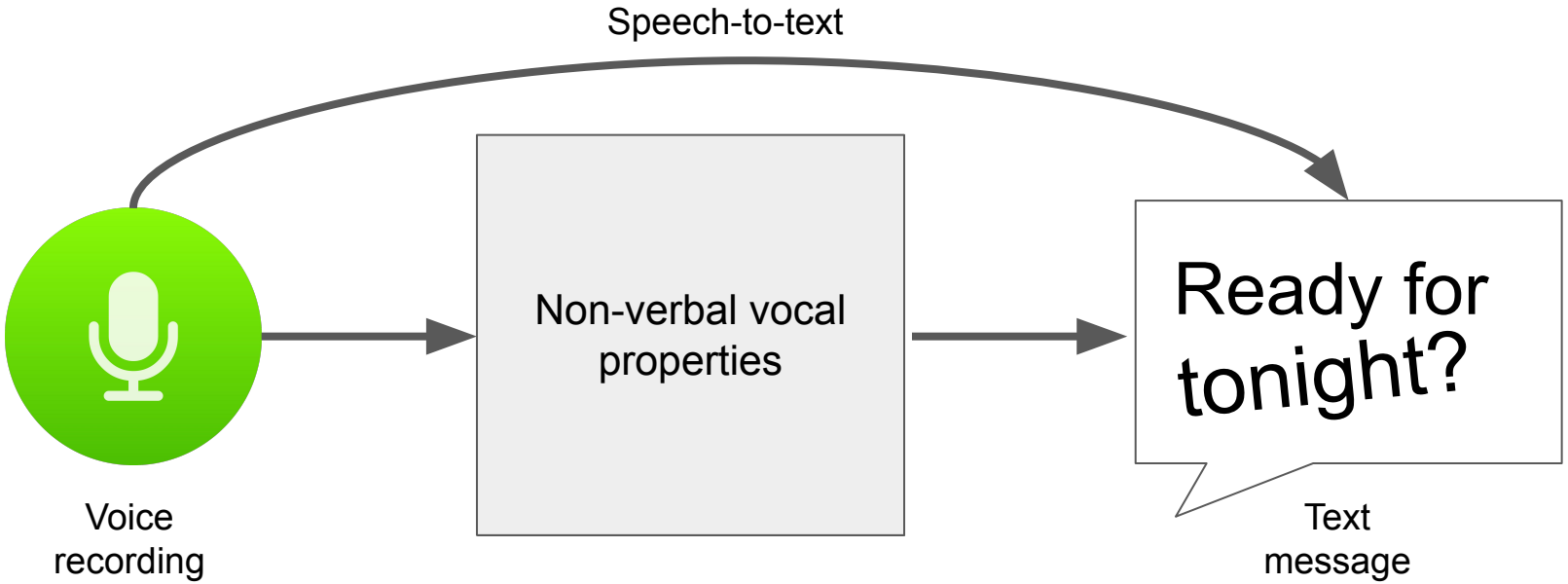

Ben says, “Ready for tonight?”
Tell Ben, “Of course!”


Ready for tonight?
Of *course*!

Sing2Text is...

- Mobile & voice assistant app
- Send messages by voice
- Receive by audio or expressive typography

Sing2Text


Speech-to-text
Non-verbal vocal properties
Ready for tonight?
Voice recording
Text message

Overview

1. Related work
2. Mapping vocal properties to typographic properties
3. Designing Sing2Text

1. Related Work

Non-verbal vocal interactions	Kinetic typography messengers	Sound-based typography manipulation

a) Non-verbal vocal interactions (NVVIs)

NVVIs: vocal interactions without using words

- **Polacek & Mikovek** (2012) use hissing
- **Vocal Joystick** (2006) uses vowels
- **Voodle** (2017) uses pitch and rhythm
- **Igarashi and Hughes** (2001) use words followed by vowel sounds

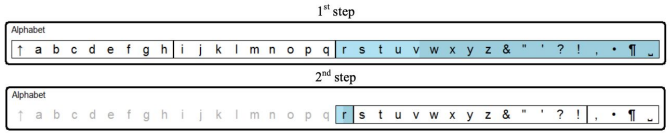
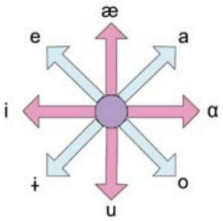
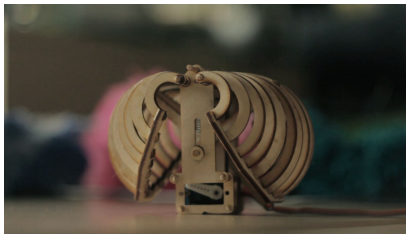
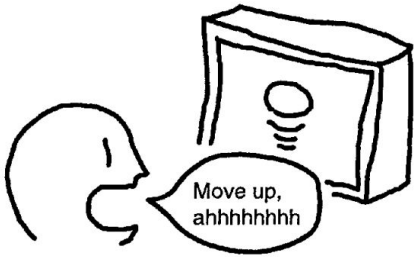


Figure 3. Ternary search scanning keyboard, typing "*r*" after "*Text ent*"



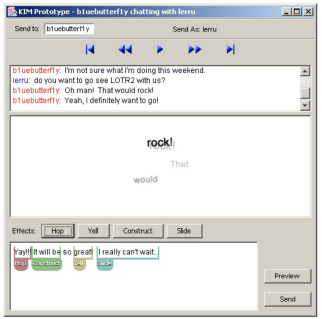
(b)

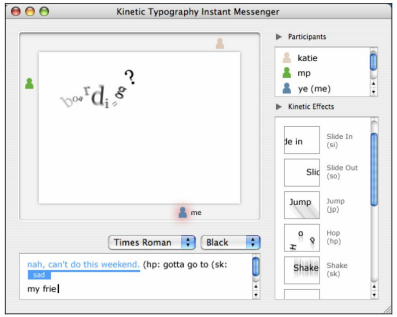




b) Kinetic typography messengers

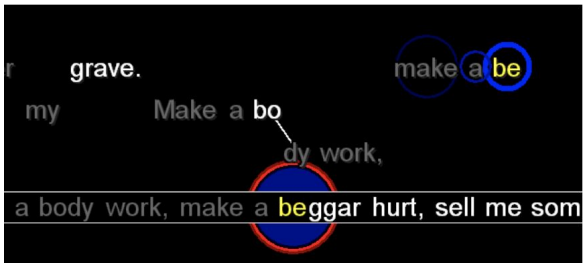
- **Wang et al.** (2004) animate typography using physiological sensors
- **Bodine & Pignol** (2003) and **Lee et al.** (2006) use kinetic typography editors
- **Yeo** (2008) uses keyword matching with emotional words

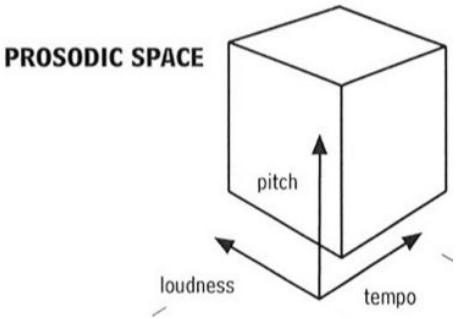




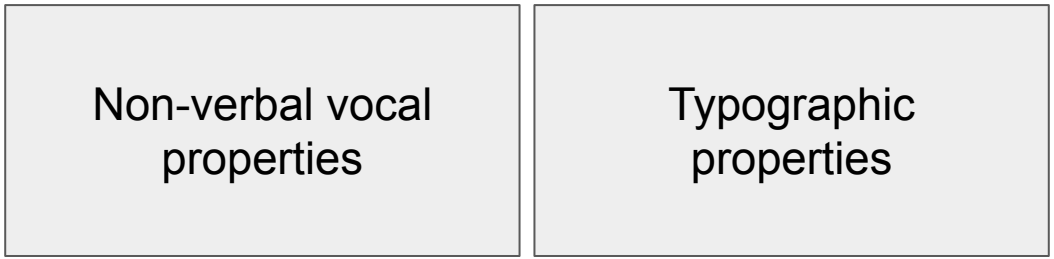
c) Sound-based typography manipulation

- **LyricText** (2008) use pitch, intensity, and rhythm of music
- **Mealing** (2003) suggests idea of using pitch, intensity, and tempo of speech





2. Mapping typographic properties to vocal properties



a) Non-verbal vocal properties

Constraints

- Independent of words spoken
- Independent of other vocal properties in mapping
- Independent of vocal quality and background noise
- Easily controllable

a) Non-verbal vocal properties

Constraints

- Independent of words spoken
- Independent of other vocal properties in mapping
- Independent of vocal quality and background noise
- Easily controllable

Properties

Word duration
Silence duration
Pitch
Intensity
Others

a) Non-verbal vocal properties

Word duration

- Easy to control in daily speech
- Dependent on words
- Measure as time per syllable

Properties

Word duration
Silence duration
Pitch
Intensity
Others

a) Non-verbal vocal properties

Silence duration

- Time between words
- Independent of words, dependent on punctuation
- Independent of all other properties

Properties

Word duration
Silence duration
Pitch
Intensity
Others

a) Non-verbal vocal properties

Pitch

- Fundamental frequency of the voice
- Continuous measure at every segment of time
  - Average frequency of word
  - First & last frequency of word
  - Character-by-character frequency
- Consideration: vocal range

Properties

Word duration
Silence duration
Pitch
Intensity
Others

a) Non-verbal vocal properties

Intensity

- Continuous measure at every point in time
- Consideration: hard consonants
- Dependent on pitch in some cases

Properties

Word duration
Silence duration
Pitch
Intensity
Others

a) Non-verbal vocal properties

Others

- Speech mode
  - Dependent on many other properties
  - Harder to control
- Emotion
  - Dependent on many other properties
  - Harder to control
- Frequencies present
  - Dependent on background noise, words spoken, vocal quality

Properties

Word duration
Silence duration
Pitch
Intensity
Others

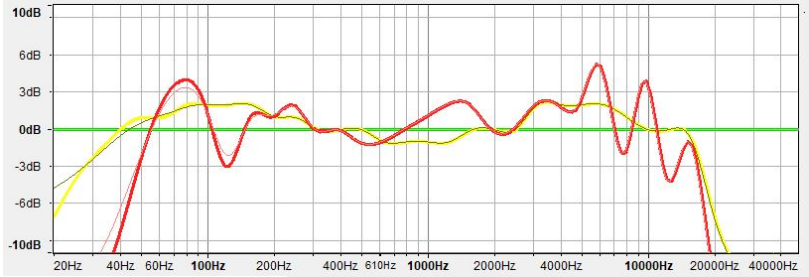
a) Non-verbal vocal properties

Others

- Speech mode
  - Dependent on many other properties
  - Harder to control
- Emotion
  - Dependent on many other properties
  - Harder to control
- Frequencies present
  - Dependent on background noise, words spoken, vocal quality

Properties

Word duration
Silence duration
Pitch
Intensity
Others



b) Typographic properties

Constraints

- Readable
- Moderate range of levels
- Easy to understand mapping

b) Typographic properties

Constraints

- Readable
- Moderate range of levels
- Easy to understand mapping

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Dimensions

- Font size
  - Readability concerns and whitespace
  - Intensity or pitch
- x-stretch
  - Word duration
- y-stretch
  - Pitch

Properties

Dimensions

Spacing

Colour

Emphasis

Discrete attr.

b) Typographic properties

Dimensions

- Font size
  - Readability concerns and whitespace
  - Intensity or pitch
- x-stretch
  - Word duration
- y-stretch
  - Pitch

Properties

Dimensions

Spacing

Colour

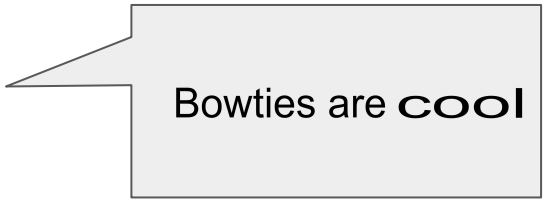
Emphasis

Discrete attr.

b) Typographic properties

Dimensions

- Font size
  - Readability concerns and whitespace
  - Intensity or pitch
- x-stretch
  - Word duration
- y-stretch
  - Pitch



Properties

Dimensions

Spacing

Colour

Emphasis

Discrete attr.

b) Typographic properties

Dimensions

- Font size
  - Readability concerns and whitespace
  - Intensity or pitch
- x-stretch
  - Word duration
- y-stretch
  - Pitch

Properties

Dimensions

Spacing

Colour

Emphasis

Discrete attr.

b) Typographic properties

Spacing

- Letter-spacing
- Word-spacing
  - Silence between words
- Leading

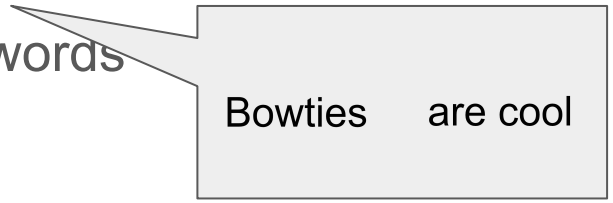
Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Spacing

- Letter-spacing
- Word-spacing
  - Silence between words
- Leading



Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Colour

- RGB, HSB… or linear gradient
  - Pitch
  - Considerations regarding synesthesia

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Colour

- RGB, HSB… or linear gradient
  - Pitch
  - Considerations regarding synesthesia colour-blind

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Continuous emphasis

- Font weight
  - Intensity
- Slant
  - Word duration
- Rotation
  - Readability considerations
  - Pitch

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Continuous emphasis

- Font weight
  - Intensity
- Slant
  - Word duration
- Rotation
  - Readability considerations
  - Pitch

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Continuous emphasis

- Font weight
  - Intensity
- Slant
  - Word duration
- Rotation
  - Readability considerations
  - Pitch

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Continuous emphasis

- Font weight
  - Intensity
- Slant
  - Word duration
- Rotation
  - Readability considerations
  - Pitch

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Discrete typographic attributes

- Cursive/roman
- Serif/sans serif/font family
- Capitalization
- Underline

Less intuitive, minimal granularity

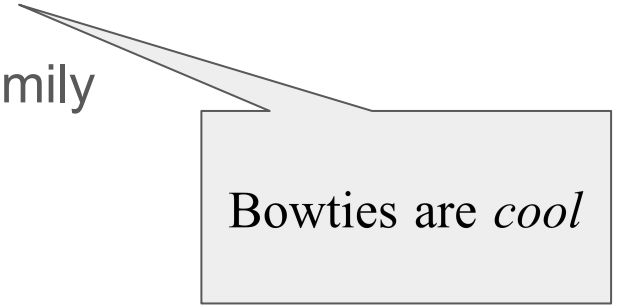
Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Discrete typographic attributes

- Cursive/roman
- Serif/sans serif/font family
- Capitalization
- Underline



Less intuitive, minimal granularity

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Discrete typographic attributes

- Cursive/roman
- Serif/sans serif/font family
- Capitalization
- Underline

Less intuitive, minimal granularity

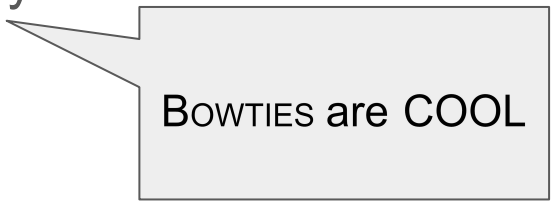
Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Discrete typographic attributes

- Cursive/roman
- Serif/sans serif/font family
- Capitalization
- Underline



Less intuitive, minimal granularity

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

b) Typographic properties

Discrete typographic attributes

- Cursive/roman
- Serif/sans serif/font family
- Capitalization
- Underline

Less intuitive, minimal granularity

Properties

Dimensions
Spacing
Colour
Emphasis
Discrete attr.

3. Designing Sing2Text

Sing2Text high-level requirements

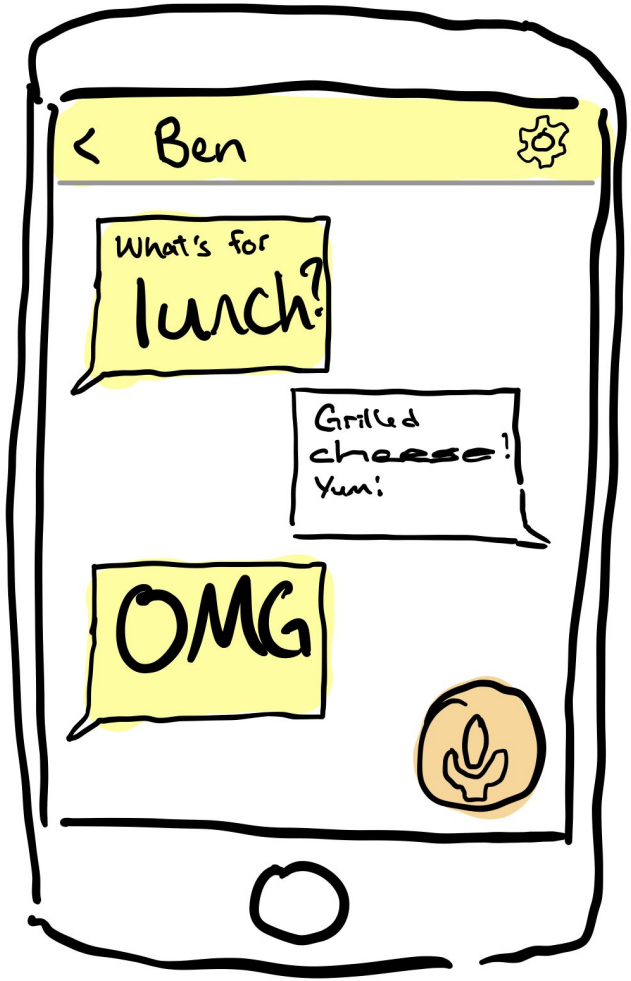
1. Accessible via a mobile app or voice assistant
2. Compose expressive messages quickly
3. Able to review messages quickly, like with text messages*
4. Options to customize mapping

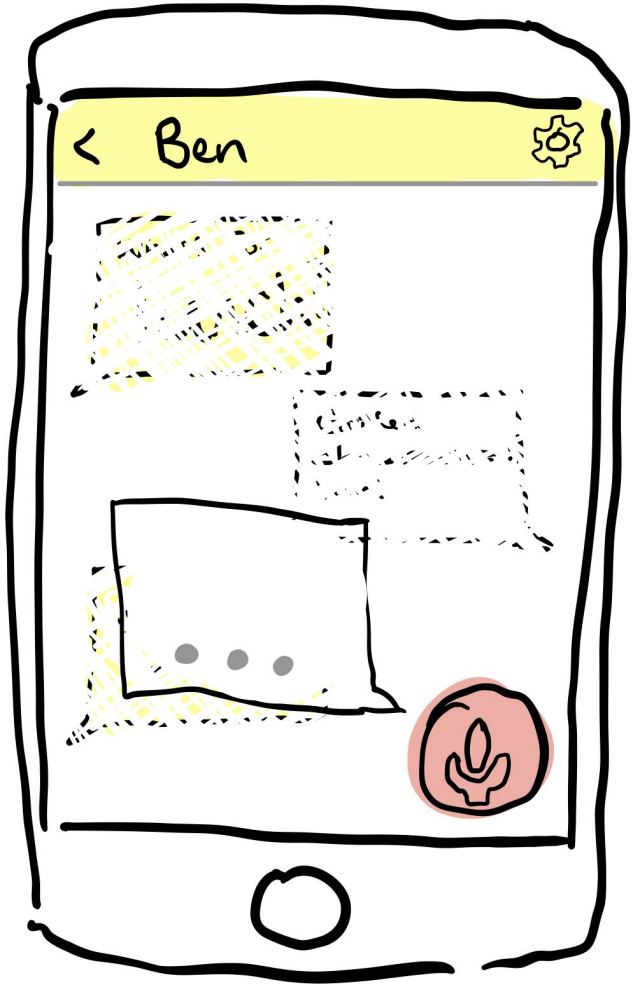
Sing2Text voice assistant app

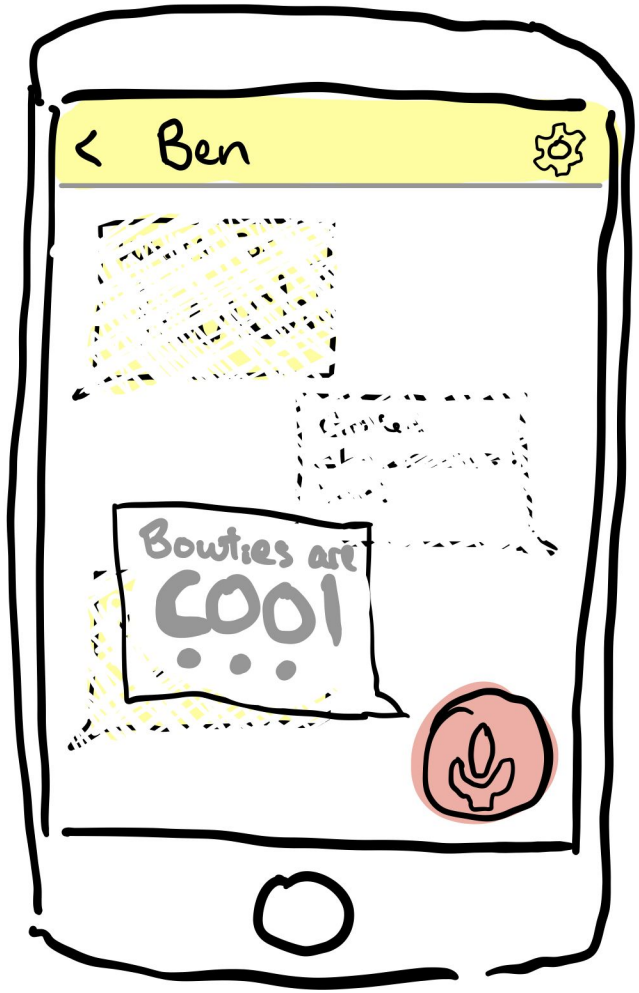
- Play back messages as they arrive
- Play back the most recent message from person X
- Record and send a message to person X

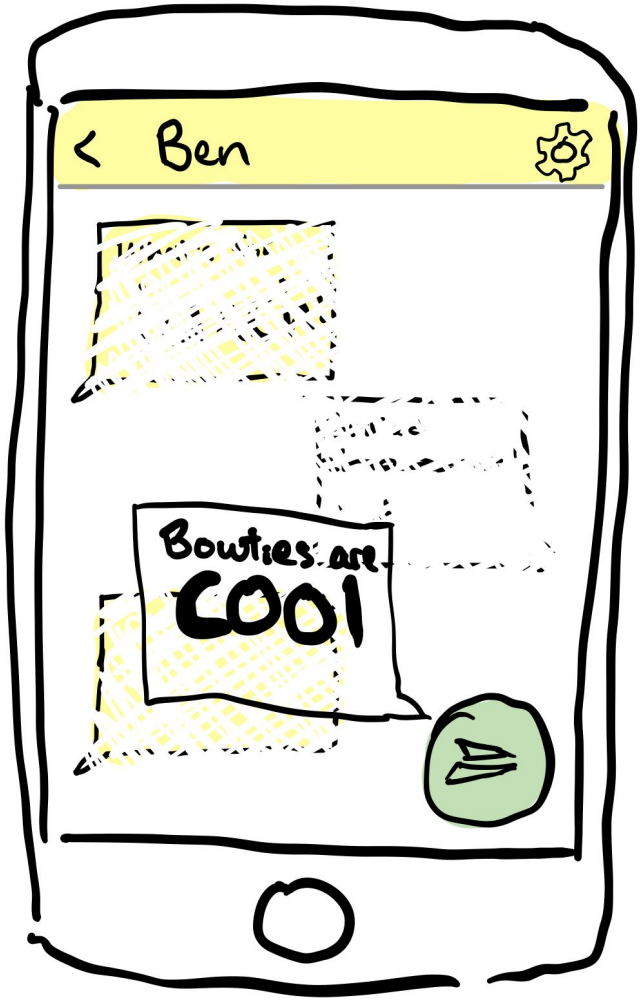
Sing2Text mobile app

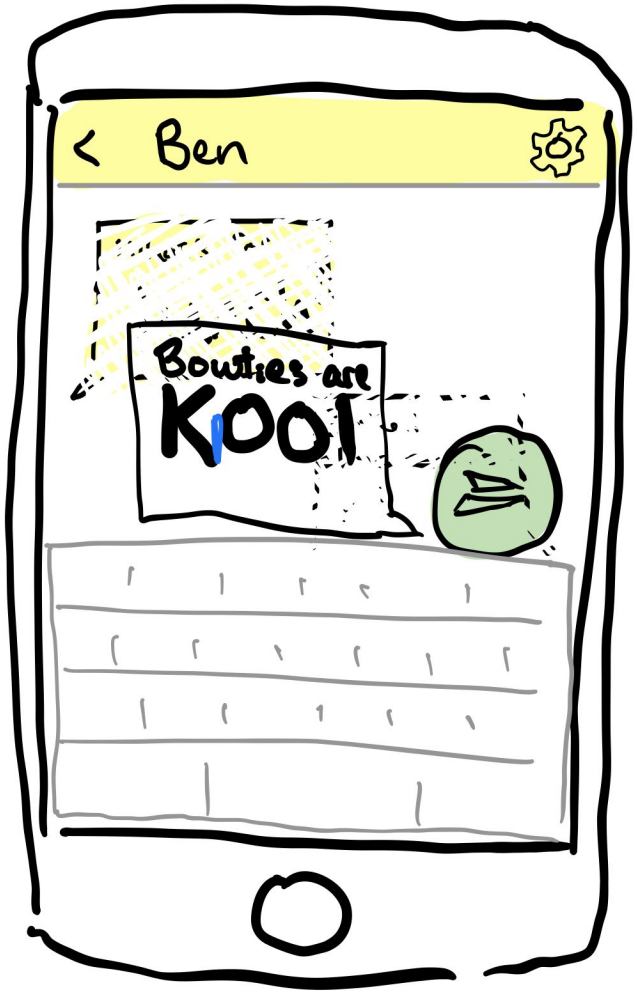
- View/play back all messages from person X
- Record and send a message for person X
- Edit the message for typos
- Customize the mapping between vocal and typographic properties

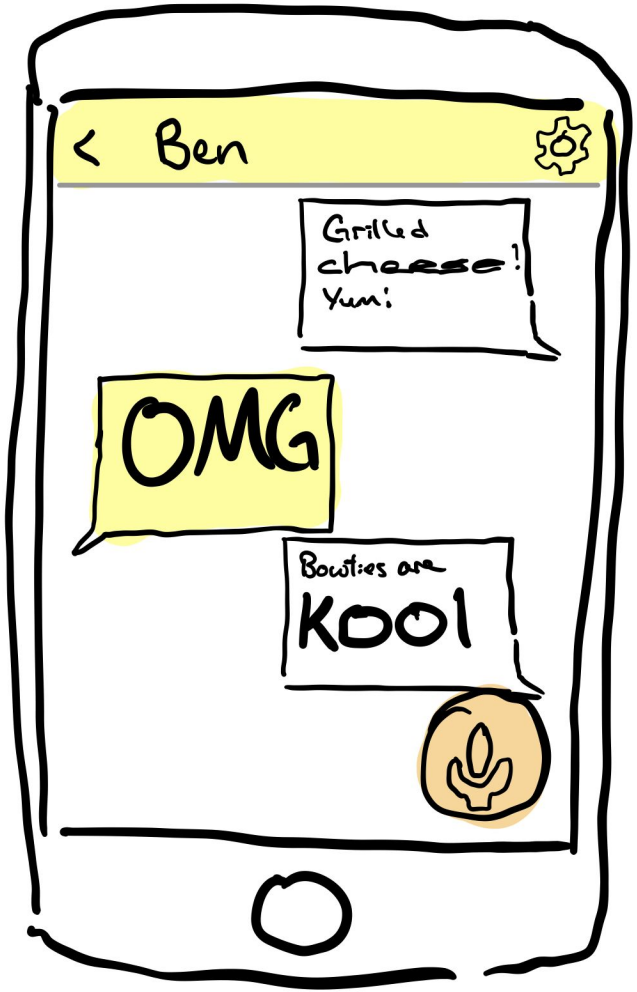

< Ben
What's for lunch?
Grilled cheese! Yum!
OMG

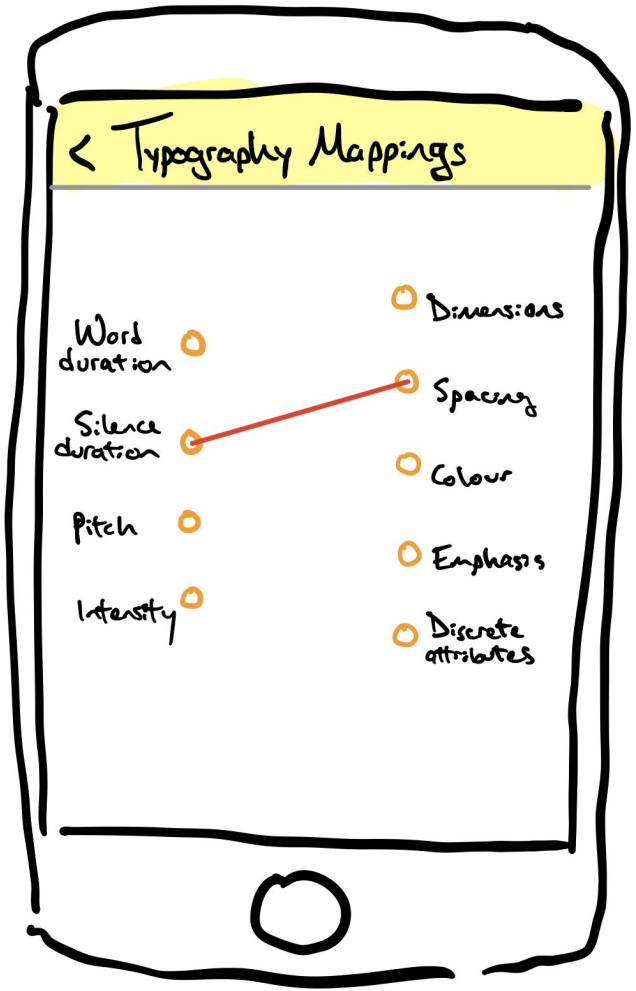

< Ben

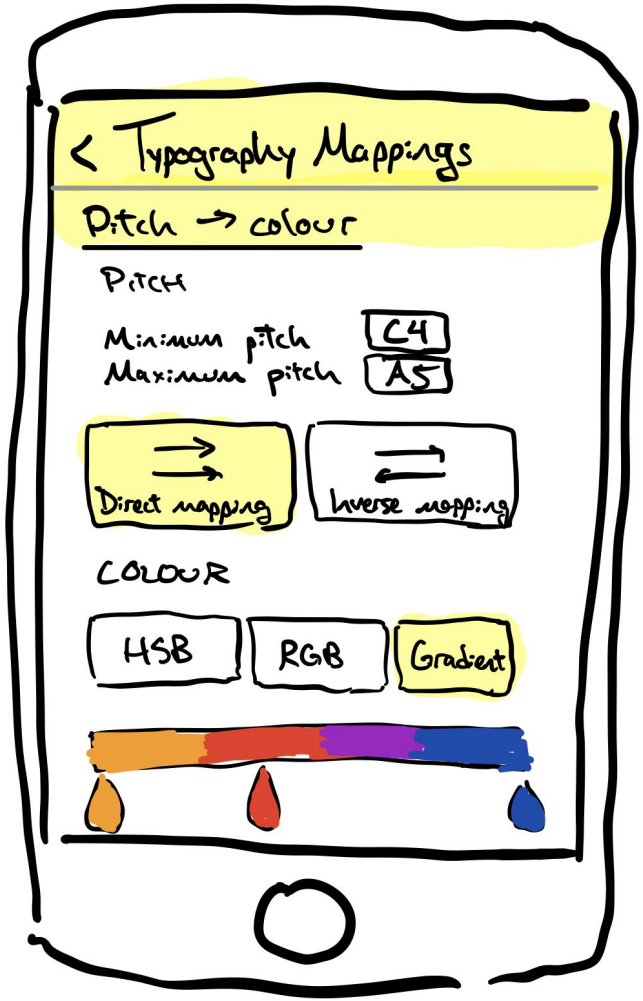

< Ben
Bowties are
cool

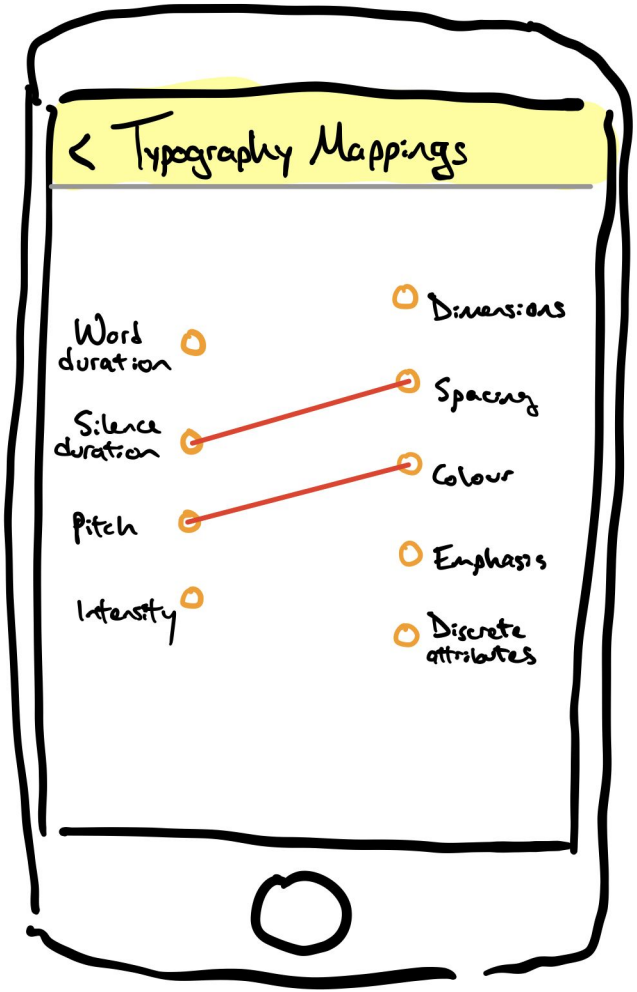

< Ben
Bowties are
COOL


< Ben
Bowties are
KOOL


< Ben
Grilled cheese! Yum!
OMG
Bowties are KOOL


< Typography Mappings
Word duration
Silence duration
Pitch
Intensity
Dimensions
Spacing
Colour
Emphasis
Discrete attributes


< Typography Mappings
Pitch → colour
PITCH
Minimum pitch
Maximum pitch
C4
A5
Direct mapping
Inverse mapping
COLOUR
HSB
RGB
Gradient


< Typography Mappings
Word duration
Silence duration
Pitch
Intensity
Dimensions
Spacing
Colour
Emphasis
Discrete attributes

Sing2Text

Concept for an instant messenger leveraging the voice for expressive typography

Graeme Zinck