



# Natural Language Speech Synthesis

## Making Voice Your Biggest Brand Opportunity

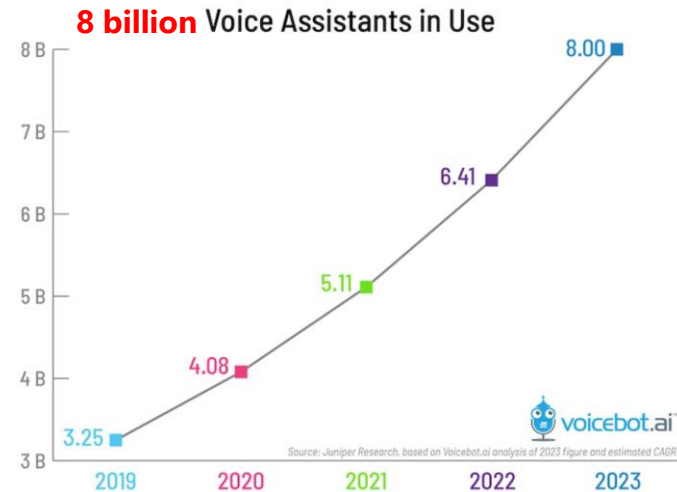
Shing Pan, VP Marketing and Business Development  
[shing@speechmorphing.com](mailto:shing@speechmorphing.com)

February 2020

# Customer Experience (CX) Market Drives Demand and Sets New Standard for Speech Synthesis



- **Digital:** More self-service and Digital/Virtual Assistants with increased use of voice
- **Conversational:** Open Dialogue, Dynamic Scripts
- **AI-enabled:** Sentiment analytics; situational/contextual-based interactions
- **Personal:** Personalized engagement
- **Branded:** Digital voice is an extension of the company brand. It's the personality and tone of your business.



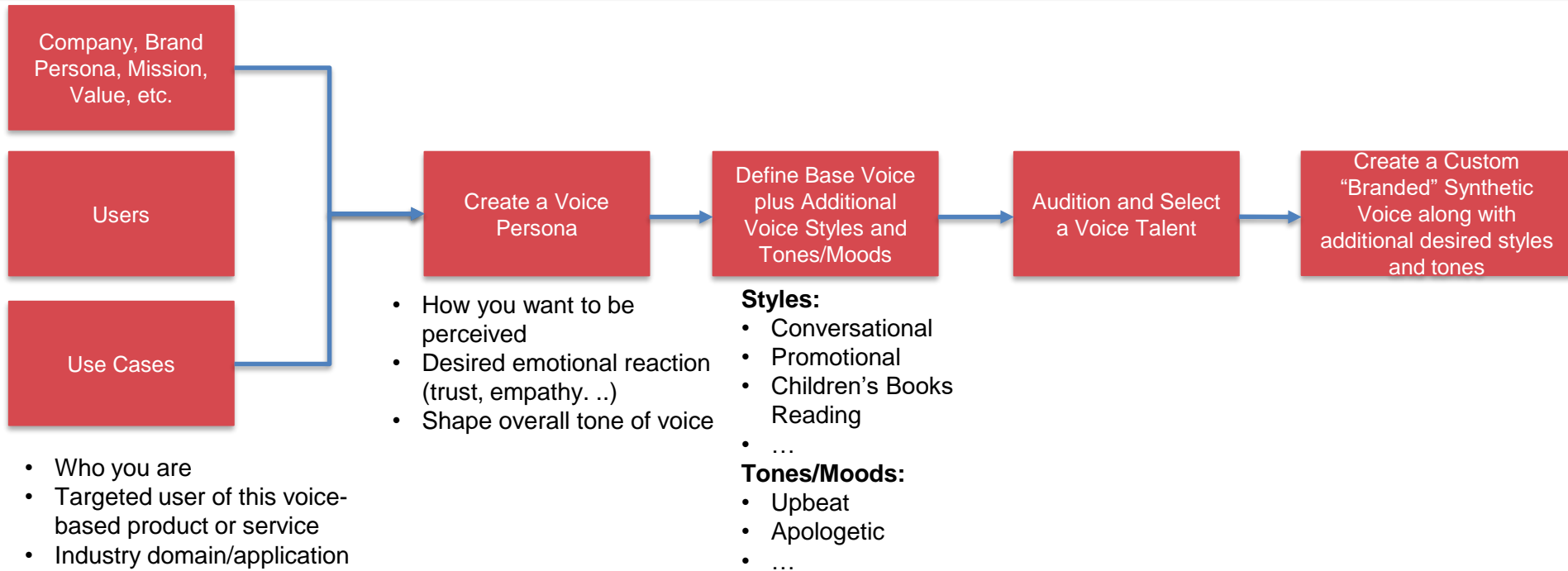
# Designing a Purposeful and Consistent Brand Voice is More Critical Than Ever



- A company brand is **unique with distinct personality**.
- Brand personality is **reflected in every customer touch point and experience**.
- Brand voice, written or spoken, is **an integral part of the company brand design**.
- Brand voice is **unique**.
- An established brand already has a brand proposition and a **recognized tone of voice**. They are the foundations for voice persona creation.
  - For example, Nike's tone of voice is urgent, motivational, and encouraging; Coca Cola's tone of voice is friendly and passionate.



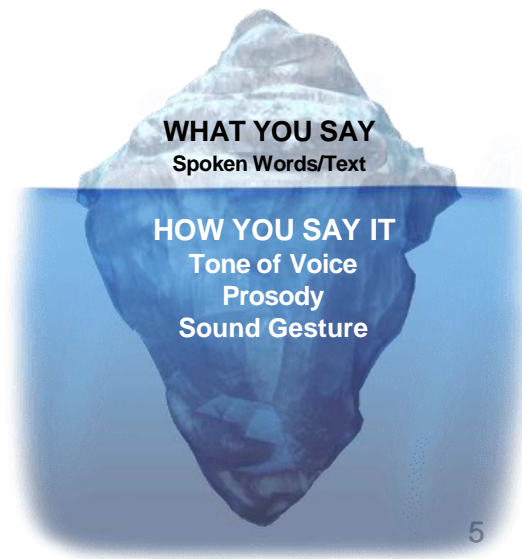
# The Journey of Your Digital Voice Design Starts from Your Company Brand



**Brand continuity can extend an established and trusted relationship to new voice-based products and services.**



- **Speechmorphing *unique*:**
  - Synthesizes not only *what* to say but *how*
  - ... with multilingual analysis, synthesis of vocal style and **explicit prosodic modeling**
- **Most current text-to-speech systems DO NOT explicitly model prosody**
  - **Cannot control** or customize exactly how the speech should sound
  - May yield **monotonous** speech



# Voice Personalization and Customization are Now Essential



**smorph™**  
ANY VOICE RENDERING

John Doe, Male, Older Adult, English

**Voice Selection**

- Search Voice
- Gender
- Age
- Language
- Select Voice

**Domain Selection**

- Search Domain
- Domain

Style: Neutral, Sympathetic, Customer Service, Sales Pitch

Style: Neutral | Gesture | Mood | **B** | *I* | OPEN | SAVE

Volume | Speed | Pitch

Audio waveform visualization

# Best-in-Class Voice Tuning Capabilities Are Desired to Further Improve Interaction Experience and Ensure Most Natural Dialogue



## Control Over:

- Phonemes (IPA)
- Variable Fields
- Sensitivity (Gender, ...)
- Break
- Boundary Tone
- Focus
- Compressed f0
- Tone
- Prosody and Intonations

Template Editor 1.1.7

Language: US English | US Default | Project: System | admin smi | Logout

Do you want the transaction history for your checking account or your savings account ?  
 Do you want to transfer money to your savings account, or to a different account ?  
 Hi, this is Alex from the Bank of the West. How can I help you ?  
 I can definitely get you over to one of our human agents - but the wait time is currently twelve minutes .  
 I can help with most requests, even transferring money or checking your balance .  
 That branch is open from eight to six Monday through Friday, and from ten to four on Saturdays .  
 Would you like me to try helping you first ?  
 Your wait time will be twelve minutes .

Did you read him his rights ?  
 Did you read him his rights ?  
 Did you read him his rights ?  
 Did you read him his rights ?  
 Did you read him his rights ?

Did you read him his rights ?  
 Do you want information about its regular opening hours or its exact address ?  
 Do you want the transaction history for your checking account or your savings account ?  
 Do you want the transaction history for your checking account or your savings account ?  
 Freud probably contributed more than anyone else to the understanding of dreams, enabling us to recognize their equivalents in our wakeful thoughts .  
 Greg, whose glasses are these ?  
 Greg, whose glasses are these ?  
 Hello!  
 Hello!  
 Hello!  
 Is that alright with you?

Word	Phonemes (IPA)	Field Type	Sensitivity	Break	Boundary Tone	Focus	Compressed f0	Tone
Did	d'id	<Not A Field>	Next Word	B_1	None	Strong	None	N/A
you	'ju	<Not A Field>	Gender	B_1	None	None	CF0	N/A
read	r'id	<Not A Field>	Gender	B_1	None	None	CF0	N/A
him	h'im	<Not A Field>	Gender	B_1	None	None	CF0	N/A
his	h'iz	<Not A Field>	Gender	B_1	None	None	CF0	N/A
rights	r'arts	<Not A Field>	Gender	B_4	IH-H%	None	CF0	N/A
?		<Not A Field>	Gender	B_0	None	None	None	N/A
	SIL	<Not A Field>	Gender	B_1	None	None	None	N/A

Hadar Conversational 20190 | Director | Save As... | Did\_you\_read\_him\_his\_rights2 | Make Public | Share with Project | Suggest New OVC

# Intent is Important to NLU. It is Important to Speech Synthesis as well. Ensure Your Virtual Agents Speak with Clear Intentions



## ASR+NLU+Dialog Management

- **Recognize you**
- **Know what you did before**
- **Anticipate your needs**
- **Read your emotions**
- **Personalize a response**



## Speechmorphing's Text-to-Speech

- **Use appropriate tones, styles**
- **Speak with clear intentions**
- **Exhibit emotional intelligence**
- **Preserve brand image**
- **Deliver natural, personable, intelligent speech**

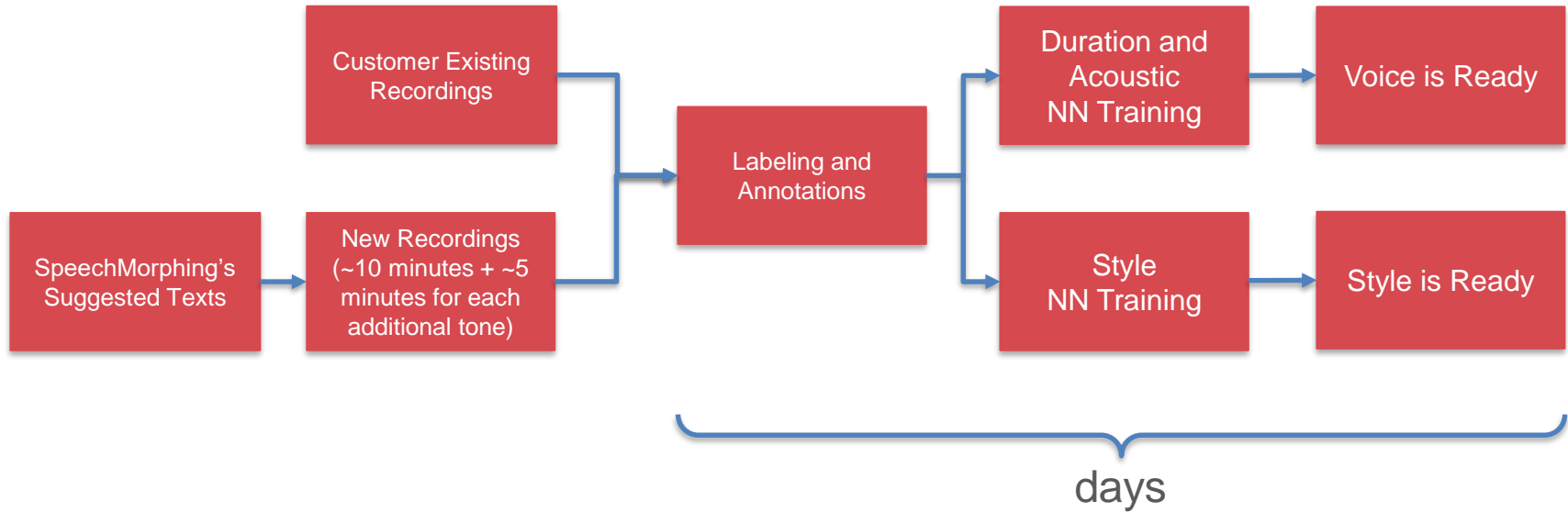


# Speechmorphing: Natural Language Speech Synthesis



- Quickly and cost-effectively produce **high-resemblance, high-quality custom “branded” voice**
  - **5-10 minutes** of voice recordings...and **days** (not months!) of voice production
- Expressive Speech with **varied styles and tones** for more personalized interactions
  - Voices are explicitly trained with detailed information about **prosody**
  - Varied styles and tones can be arbitrary mixed during synthesis to support **situational** dialogue and context
- **Best-in-class voice tuning** capabilities for perfect prosody and natural dialogue
  - Highly **customizable voices**
  - **Non-stitched** voice prompts with **desired prosody** and **self-defined variable fields** (names, time, location, address, ...)
- Powering leading **conversational AI and voice platforms** for improved CX

# Create Custom Voice in Days (Not Months!)



# Rapid Development of High Resemblance, High Quality Custom Voice Created in Days with 5-10 minutes of Audio Material



**Resemblance**



**Conversational Dialogue**



**Natural Prosody**



**Accent**

# Expressive Voice with Varied Tones and Styles



 Angry	 Stern
 Promotional	 Empathetic
 Praising	 Scolding
 Presentational	 Conversational

# Speechmorphing at a Glance



- Leading speech synthesis technology for improving human-machine interaction
- Expertise in computational linguistics, speech synthesis, digital signal processing, deep neural networks, and machine learning
- Founded in 2010 with headquarters in San Jose, CA and offices in Israel and Europe
- Powering leading conversational AI and voice platforms



*“Most natural dialogues”*

*“A game-changing TTS”*

*“Voices with emotional intelligence”*

...



Thank You!

