ENVOY: Enabling Natural-language Versatility & Opportunity

Evan Dietrich   Akhil Goel   Mark Goldwater   Kirsten Leong   Riley Martell   Ryan McGill   Erin Mitchell
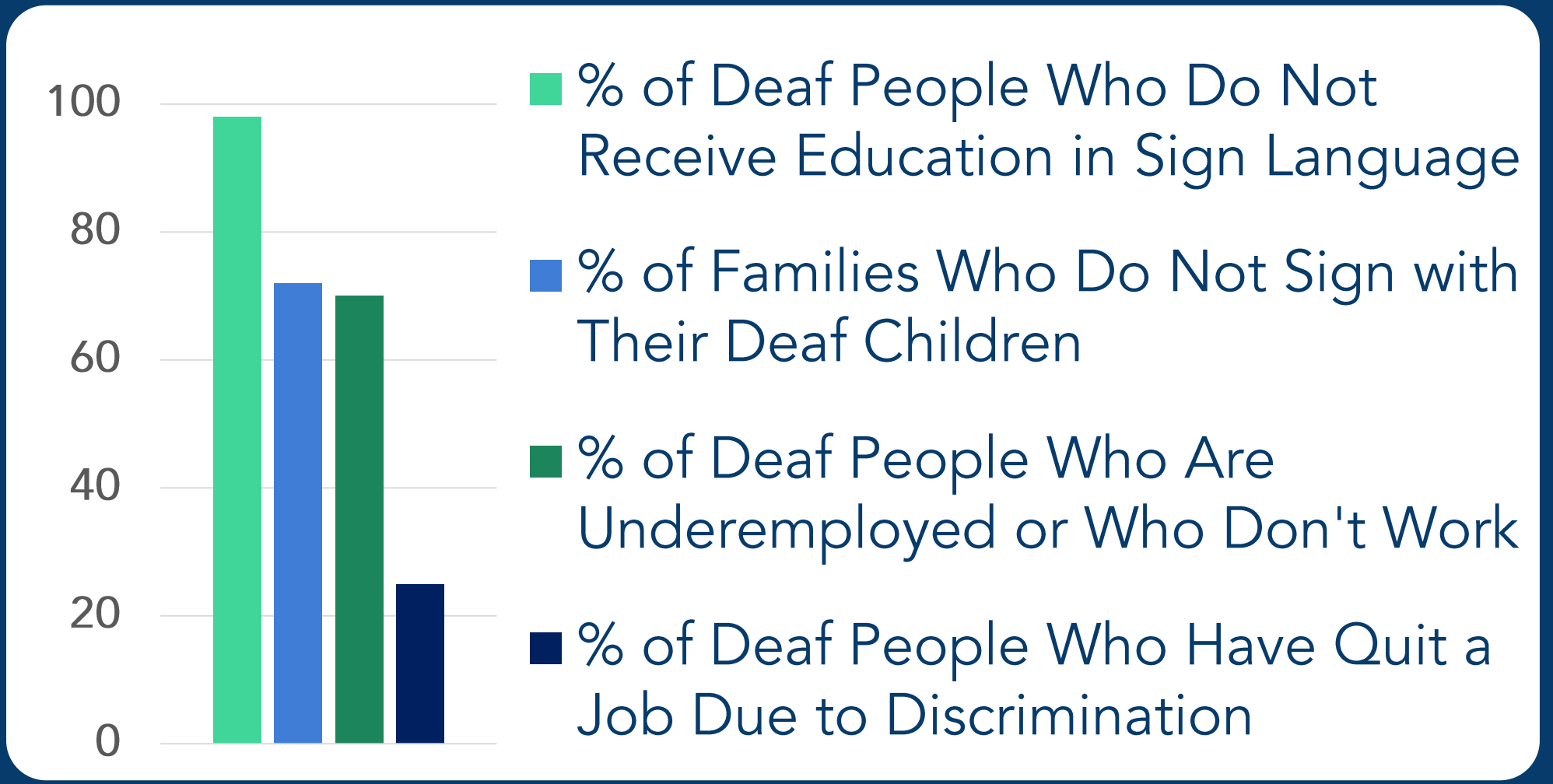
Overview

ENVOY is a modular communication system & teaching tool with the intent of enabling real-time conversations between Deaf ASL signers & English-speaking individuals

Motivation

Problems being Addressed

Bar chart legend:
- % of Deaf People Who Do Not Receive Education in Sign Language
- % of Families Who Do Not Sign with Their Deaf Children
- % of Deaf People Who Are Underemployed or Who Don't Work
- % of Deaf People Who Have Quit a Job Due to Discrimination
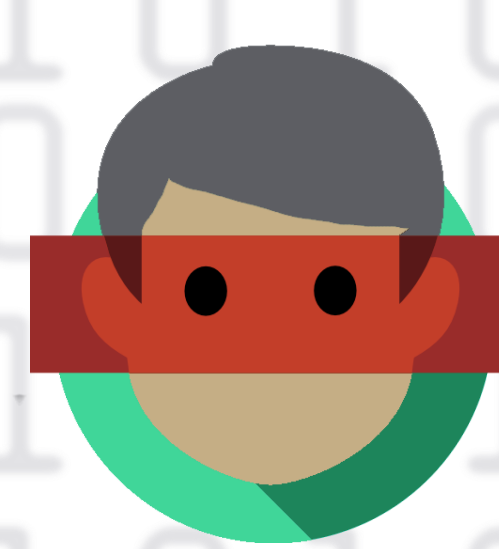
Impact of Communication Limitations
(Personal & Professional)

- Relationships w/ family & friends
- Taking advice & directions
- Collaboration w/ teachers & peers
- Interpreters rarely provided
- Exclusion from career networks
- Social isolation & societal stigma

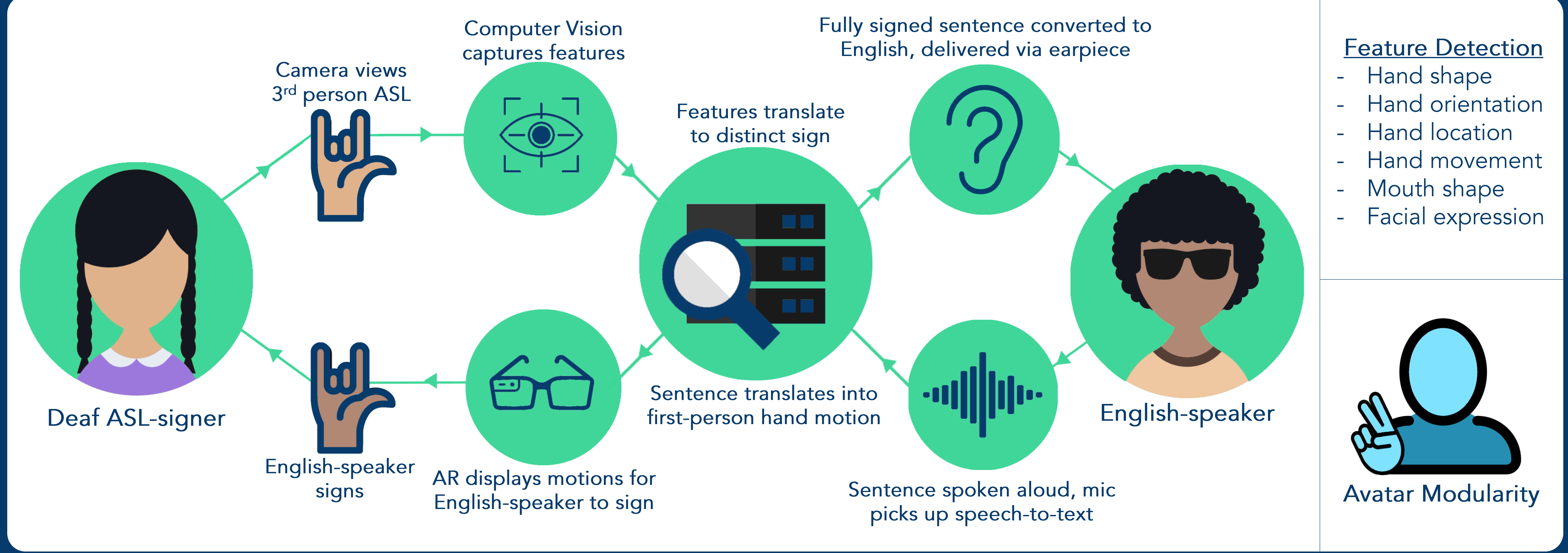### Major Gaps in Current Attempts

Device approach (gloves/wristband) is rooted in the preoccupations of hearing persons, instead of needs of Deaf signers

Most instruments neglect facial expression as ASL feature criteria, but without them, many signs are indistinguishable
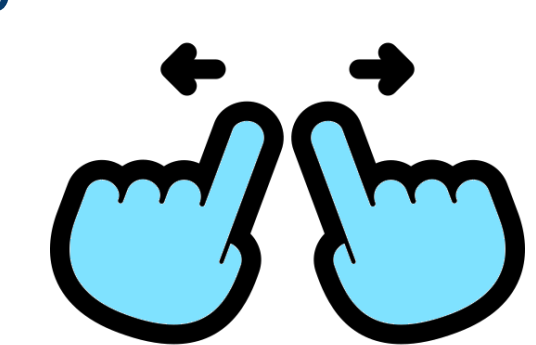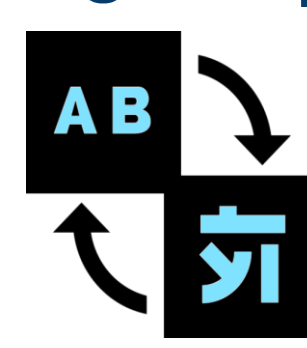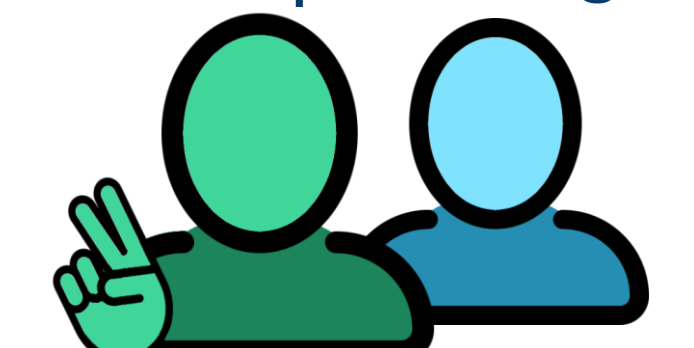
Proposal

System Overview

- Camera views 3rd person ASL
- Computer Vision captures features
- Features translate to distinct sign
- Fully signed sentence converted to English, delivered via earpiece
- Deaf ASL-signer
- English-speaker signs
- AR displays motions for English-speaker to sign
- Sentence translates into first-person hand motion
- Sentence spoken aloud, mic picks up speech-to-text
- English-speaker
- Avatar Modularity

Feature Detection
- Hand shape
- Hand orientation
- Hand location
- Hand movement
- Mouth shape
- Facial expression

Anticipated Challenges & Potential Solutions

Challenge	Potential Solution
Full language classification	Machine learning currently used but works only under a limited lexicon. Deep learning can improve accuracy & robustness of system
Data collection for deep learning	We will need to integrate existing datasets & build our own
Lack of diverse signers & variety of recordings (lighting/environment)	Data augmentation techniques can generate more data & under more diverse conditions
Variation in gesturing from user to user	User may opt-in for their signing to be used in ENVOY database & to personalize for individual tendencies
Accessibility & ease-of-use	Using affordable components such as the Kinect Sensor makes the system more accessible
Distinguishing between discrete signs can be hard with features alone	
No existing mapping of syntax & grammar between ASL & English	Integrating an ASL language model will allow us to use contextual clues to improve user experience & ensure users can sign naturally

Modularity Use Cases

Single-Learner Classroom
Follow words & phrases on display, camera checks progress

Non-ASL Speaker Skype Call
Encouraged to follow signs on-screen, avatar usage if unable

Multi-lingual Options
Translates between different preferred native languages

Group Meeting
Unique avatars for multiple speakers in same conversation

Future Path

While being our current focus, ASL is not a Universal Language

After tangible success, we hope to integrate all sign languages to maximize our potential impact for communication between 70 million+ Deaf individuals, their colleagues, & their loved ones