ASPREX Fact Sheet

Gesture to voice technology

Camera-based system to identify motion gestures (e.g., hand movements, facial expressions, and sign language, used by deaf persons to communicate), recognize them and translate them in real-time into speech. They can also perform the reverse function of translating from speech to gestures (sign language) communicated by an animated avatar on the screen. Internet pages, with integrated text to sign language translation technology options, may facilitate access to information for deaf and hearing-impaired signers who have no personal assistive technology.

Possible variants concern the recognition technique (which may also be a non-vision technology – based on sensors placed on the body or on a garment of the person, e.g. a special glove, that analyze the movement of the hand and arm – or an audio-based technology for speech to gesture conversion), the conversion from gesture to voice (which may also be from gesture to text), the conversion from voice to gesture (which may also be from speech to text, or from speech to text and gesture), the desk top version of the system (which can also be mobile version, on mobile phone, or tablet), the software (which can also be a plug-in for web pages, for the translation from text to sign language, include the alternate text images), the web cloud based support (which can also be missing, in which case the recognition system is installed on the device), and the bi-directionality (which can also be one-way conversion, from gesture to text, from gesture to spoken language, from text to sign language, or from spoken language to sign language).

- Product Classification
  - APL (WHO Assistive Product Priority List): 14 (Gesture to voice technology)
  - ISO 9999:2022: 222109 (Dialogue units)
- Possible configuration variants

  None specified.
- Possible accessories or optional components

  None specified.
- Product goals

  Activities or functions the product is mainly intended to support, according to WHO ICF Classification:
  - Understanding gestures or sign language [d320].
  - Producing messages in sign language [d340].
  - Reading [d166].
- Indicated impairments

  Difficulties the product is mainly intended to address, according to the WHO ICF Classification:
  - Hearing [b230].
  - Listening [d115].
  - Understanding spoken words [d310].
- Contraindicated impairments

  Difficulties for which the product may be inappropriate:
  - Inability to understand sign language.
  - Deafblindness.
  - Motor difficulty that makes it difficult to interpret the gesture, the facial expression, or the sign language.

- Indicated environments

  Specific environments in which the product should be used:
  None specified.
- Contraindicated environments

  Environments in which the product may be inappropriate:
  - Noisy places (the accuracy of voice to gesture or voice to text recognition may be difficult in noisy environments).
  - Places with bad visual conditions (bad lighting, uneven background, and little contrast between the speakers' hands and face and the background).
- Other indicated factors

  Other factors or situations the product is intended to address:
  - Supporting conversation between people who use sign language and people who do not know it.
- Other contraindicated factors

  Other factors or situations in which the product may be inappropriate:
  None specified.
- Points to be considered in product selection
  - Ensure that the user can adequately communicate with sign language, including fingerspelling.
  - Ensure that the gesture to text or to voice technology is available in the sign language of the user; be aware that the spoken language does not reflect the sign language (e.g., American Sign Language, ASL, is quite different from British Sign Language, BSL, even though English is the spoken language of both countries).
  - A mobile gesture to voice technology is suitable for signers that want to communicate with a talking person in the marketplace, workplace, schools, health care, and civic centers.
  - In case of cloud-based systems and mobile use, make sure that the place you want to use gesture to voice technology is covered by wireless internet access.
  - Select a voice or text to gesture system for deaf and hearing-impaired signers who have difficulties in reading and understanding written language.
  - The accuracy of the translation process of text to gesture technologies is very high, because it is less susceptible to negative interference of contextual factors like light, contrast, noise, or sound quality.
  - For mobile use of text to gesture technology, the user's device may be touched by unknown people and the user may be exposed to the virus.
  - Speech to text technologies are recommended if the user can read and understand written texts; free apps are easily available, also in different languages.
  - In case of video or telecommunication, select a software (ISO code 222424, telecommunication, and telematics software) with integrated voice to text translation and transcription (closed capturing) into the language the user can read and understand.
  - For gesture to text or to speech technologies, a hands-free communication is necessary, and the device camera must frame the hands; mobile use may be difficult, if there is no appropriate placement for the device.
- Points to be considered in product fitting
  - Ensure that the user can activate his gesture to voice technology.
  - In case of gesture to voice or to text technology to produce communication, ensure that the translation is accurate, if not, increase the environmental brightness, change the background, or try to sign slower.
  - In case of machine learning system, correct misinterpretations and teach the system to better interpret your personal sign language.
  - In case of voice to gesture technology to receive communication, check if the translation is accurate, if not, reduce the environmental noises, bring the device microphone nearer to the speaker, increase the microphone sensibility, and ask the speaker to speak louder, slower and one at a time.
  - In case of text to gesture technology to receive communication, check if the translation is accurate, if not, check if synonyms exist and add new words or synonym words to your personal system.
  - In case of voice to text technology to receive communication, check if the translation is accurate, if not, manually modify the misinterpretation and add the new words to your personal device system.
- Points to be considered in product use

- If the voice to gesture technology is not working perfectly, try to change environmental factors to increase the functionally of your system.
- If possible, continue to correct the misinterpretations and teach your system to make it work more accurately.
- If you change your device on which you have installed the voice to gesture technology, make sure that all your personal adaptations will be transferred to the new device.
- Place the camera or your device with a camera in a stable and safe place and ensure that it is properly oriented to frame your gestures.
- For mobile use, ensure that your device is charged.

- Points to be considered in product maintenance / follow-up
  - Regularly clean the camera of your voice to gesture technology.
  - Frequent up-date of the software is recommendable.
  - Frequently up-date and back-up your personal adaptations to the voice to gestures technology.
  - In case of mobile device, control and charge the batteries when they are low.
- Examples of products available on the market
  - Live product search in the EASTIN website https://www.eastin.eu/en/searches/products/list?iso=222109

Source

This Fact Sheet was compiled in 2021 by an international team of experts, to provide the initial knowledge base for a project ("An online system to assist the selection of assistive product") supported by the World Health Organization in 2020-2021 within the GATE Initiative (Global collaboration on Assistive Product). Fact Sheets were compiled for each of the 50 types of products included in the WHO APL (Assistive Product Priority List).

The team was composed of Renzo Andrich (Italy, group leader), Natasha Layton (Australia), Stefan von Prondzinski (Italy), Jerry Weisman (USA), Silvana Contepomi (Argentina) and Hasan Minto (Pakistan).

The project led to a prototype online tool called ASPREX (ASsistive PRoduct EXplorer). At the end of the project, it was transferred to a WHO collaborating center (the Global Disability Hub in the UK), in view of possible future developments.