



# “HEY SIRI, DO YOU UNDERSTAND ME?”: VIRTUAL ASSISTANTS AND DYSARTHRIA

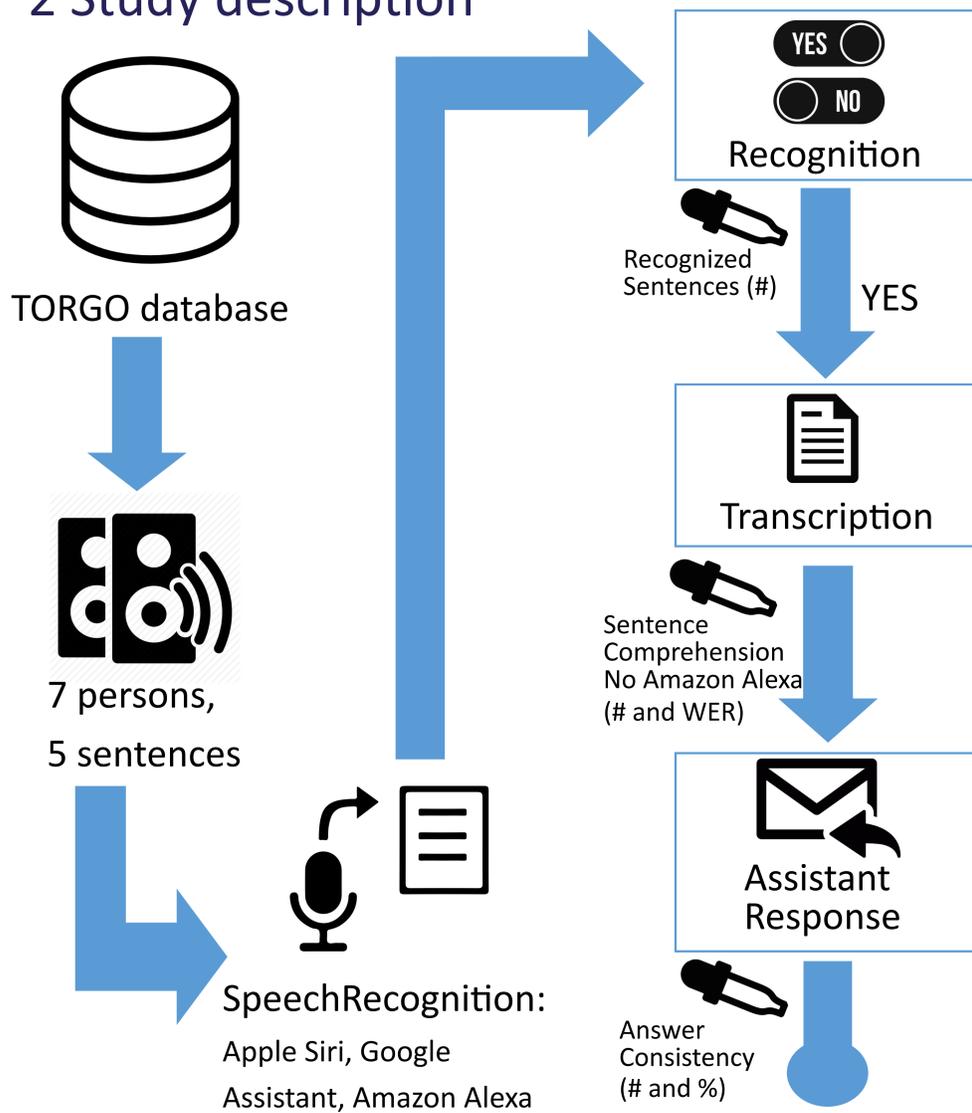
## 1 Motivation and Goal

The last five years have seen the spread of voice-controlled smart environments, powered by virtual assistants. Nowadays, people can speak to their smartphones, smart watches, smart homes...

A consequence of these devices embracing voice control is that people with dysarthria or other speech impairments may be unable to control their intelligent environments .

**Research question:** can people with different degrees of dysarthria easily access and be understood by voice-controlled devices?

## 2 Study description



## 3 Sentences from TORGO database

TORGO database

(<http://www.cs.toronto.edu/~complingweb/data/TORGO/torgo.html>)

Collection of dysarthric speech in English. The dataset contains audio files and transcriptions of sentences. We extracted 5 different sentences pronounced by 7 different speakers.

The five sentences:

S1. Some hotels are available nearby

S2. Start

S5. Please open the door quickly

S3. Play

S4. Today's date

## 4 Results

	Recognition Recognized Sentences	Sentence Comprehension		Answer Consistency	
		Correct Transcriptions	WER (mean)	Appropriate Responses (#)	Appropriate Responses (%)
Apple Siri	17 / 17	8 / 17	69.41%	10	58.82%
Google Assistant	13 / 17	10 / 17	15.38%	11	64.70%
Amazon Alexa	14 / 17	-	-	7	41.17%

Recognition

- Apple Siri tried to recognize all 17 sentences, by transcribing something. The other two assistants indicated, instead, that they were not able to recognize anything for some speech samples.

Sentence comprehension

- Almost equal level of recognition, for the two virtual assistants, with a percentage of correct transcriptions among 50 - 60%.

Answer Consistency

- The assistants were consistent in their answers. Apple Siri always tries to answer any request. Amazon Alexa and Google Assistant, instead, respond only if they recognize at least some words.

## 5 Conclusion

Preliminary results show that the three virtual assistants have comparable performance and similar behaviors for all three criteria. All the evaluated assistants have difficulties to understand dysarthric speech, this makes them poorly accessible from people with dysarthria.

Similar results were already found for contemporary ASR systems when used by people with other speech impairments, e.g., deaf people.

Future works will better assess such virtual assistants by characterizing their usefulness for different degrees of dysarthric speech.

Furthermore, we will use the outcome of this evaluation as a starting point to improve the accessibility and the recognition capabilities of such assistants.



Fulvio Corno

Luigi De Russis

Fabio Ballati

<https://elite.polito.it>