

What topic do you want to hear about? A bilingual talking robot using English and Japanese Wikipedias

Graham Wilcock
CDM Interact, Finland
University of Helsinki, Finland
gw@cdminteract.com

Kristiina Jokinen
University of Tartu, Estonia
University of Helsinki, Finland
kristiina.jokinen@helsinki.fi

Seiichi Yamamoto
Doshisha University, Japan
seyamamo@mail.doshisha.ac.jp

Abstract

We demonstrate a bilingual robot application, WikiTalk, that can talk fluently in both English and Japanese about almost any topic using information from English and Japanese Wikipedias. The English version of the system has been demonstrated previously, but we now present a live demo with a Nao robot that speaks English and Japanese and switches language on request. The robot supports the verbal interaction with face-tracking, nodding and communicative gesturing. One of the key features of the WikiTalk system is that the robot can switch from the current topic to related topics during the interaction in order to navigate around Wikipedia following the user's individual interests.

1 Introduction

The WikiTalk system for Wikipedia-based spoken information access dialogues is described by Jokinen and Wilcock (2012a) who also presented information access with robots in a tutorial at COLING 2012 on *Open-domain conversations with humanoid robots* (Jokinen and Wilcock, 2012b). Different aspects of the implementation of WikiTalk on Nao robots (Figure 1) are discussed in several papers, including Csapo et al. (2012) on integration of the technologies, Meena et al. (2012) on the use of gestures in interaction, and Han et al. (2012) on the use of visual, sonar and other non-verbal information.

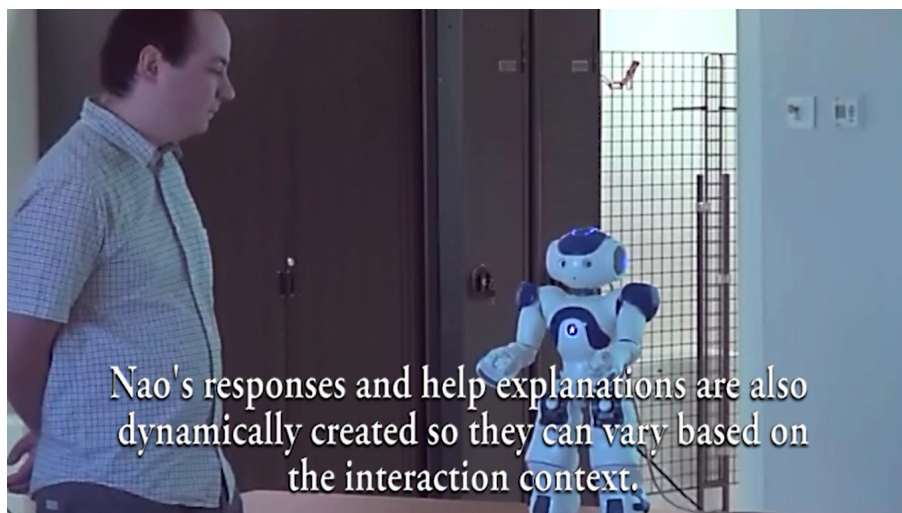


Figure 1: The first demo of English WikiTalk on a Nao robot at Supélec, Metz, in July 2012 (<https://drive.google.com/open?id=0B-D1kVqPMLKd0EcyS25nMWpjUG8>).

This work is licensed under a Creative Commons Attribution 4.0 International Licence.
Licence details: <http://creativecommons.org/licenses/by/4.0/>

WikiTalk was also demonstrated at SIGDIAL 2013, using a Nao robot for spoken information access dialogues with English Wikipedia (Jokinen and Wilcock, 2013). Although the speech recognition in 2012 and 2013 often gave low confidence scores, users were able to obtain spoken information from the robot about their desired topics and were able to navigate by speech from topic to topic. An evaluation of WikiTalk was published in 2013 by Anastasiou et al. (2013), showing that the robot was regarded as a lively and exciting interaction partner with future potential as an interesting agent interface, although the users' expectations about fluent speech interaction were higher than the robot's actual capabilities.

We subsequently developed multilingual capabilities for WikiTalk by adapting techniques for internationalisation and localisation of software systems to our spoken dialogue system, as described by Laxström et al. (2016). The first two localisations were for English and Finnish. A Finnish-speaking robot using WikiTalk was first demonstrated at EU Robotics Week 2014 in Helsinki. A video report by Iltalehti newspaper titled "This robot speaks Finnish and can tell you what is a robot" can be seen at http://www.iltalehti.fi/iltvdigi/201411290140927_v4.shtml.

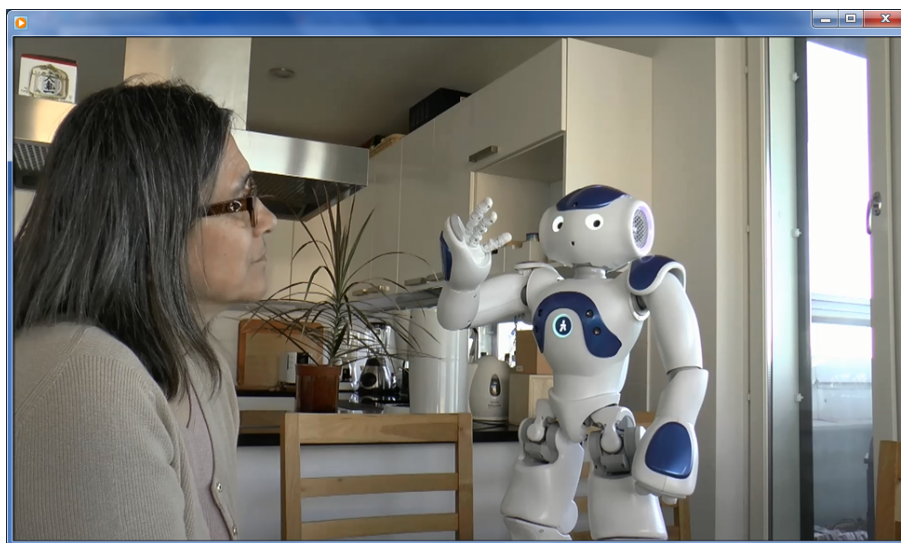


Figure 2: A robot talking in Finnish about a wide range of topics in a domestic setting (<https://drive.google.com/open?id=0B-D1kVqPm1KdY05JakMtMFJwRnc>).

The video in Figure 2 shows a robot in a domestic setting talking in Finnish about a wide range of topics using information from Finnish Wikipedia. Recent improvements in speech recognition can be seen by comparing the video in Figure 1 with the more recent videos (Figures 2 and 3). In the earlier video the robot often has low confidence scores for speech recognition and in that case is programmed to ask the user for confirmation, often asking for example "Did you mean *enough*?". In the later videos this almost never happens. Also in the earlier video the user leans forward to speak as close as possible to the microphone located in the robot's forehead, but this is not necessary in the more recent videos.

A localised Japanese version of WikiTalk developed in 2015 is described by Okonogi et al. (2015). At SIGDIAL 2015 we presented a video (Wilcock and Jokinen, 2015) showing a robot speaking English and Japanese, getting information from English and Japanese Wikipedias, and switching languages on demand. The dialogues with the robot in this video are described in Section 2.

At COLING 2016 we will demonstrate English and Japanese WikiTalk with a bilingual Nao robot, showing the improved speech recognition capabilities and focussing on unscripted user interaction and the system's ability to shift smoothly from the current topic to related topics to follow the individual user's interests. The demonstrated system is described in Section 3.

Future plans, including a system that can be configured for Nao and Pepper robots and also for robots that use ROS, the open source Robot Operating System, are briefly described in Section 4.

2 The robot dialogues in the video shown at SIGDIAL 2015

The video in Figure 3 shows information access dialogues with English and Japanese WikiTalk, using a bilingual Nao robot that switches language on demand. This video was shown at SIGDIAL 2015 and is described by Wilcock and Jokinen (2015).



Figure 3: Annotated video of an English-Japanese language-switching robot (<https://drive.google.com/open?id=0B-D1kVqPm1KdRD1kVHh4Z2tUTG8>).

The video lasts just over 14 minutes. The robot speaks English with an English-speaking user in the first 7 minutes, then speaks Japanese with a Japanese-speaking user for 5 minutes, and finally switches back to English for the same English-speaking user for the last 2 minutes.

At the beginning the robot identifies a human face and makes eye-contact. When the human moves, the robot uses face-tracking to maintain eye contact. It explains in English that it can talk about any topic in Wikipedia, and suggests some favourites such as Shakespeare and Manchester United. It briefly switches to Japanese to invite the user to select Japanese, but the human ignores the offer and the robot switches immediately back to English.

The user asks for “Shakespeare”, one of the suggested topics. The robot connects to Wikipedia via wifi, downloads the latest version of the article about the selected topic, processes the information to produce sentences suitable for speaking, and begins talking about Shakespeare. The robot continues talking about this topic for some time, but after completing a paragraph-sized chunk of information with no interruption by the user, the robot stops and asks explicitly whether to continue or not.

The user asks to “continue” and the robot continues telling more information about Shakespeare. After another paragraph-sized chunk of information about the same topic, the robot does not simply ask whether to continue, but explains some of the dialogue options by telling the user “You can change to other topics related to Shakespeare simply by saying them”. The user then asks about Shakespeare’s son Hamnet so the robot shifts topic and starts talking about Hamnet Shakespeare.

After the robot mentions Shakespeare’s play Julius Caesar, the human asks about “Julius Caesar” and the robot starts talking about the play. Interestingly, the robot mentions the historical person Julius Caesar while talking about the play with the same name. Next the human again asks about “Julius Caesar”, and this time the robot starts talking about Julius Caesar the person, not the play, as the person is more recently mentioned.

Soon the English-speaker stops interacting and goes away, and a Japanese-speaker approaches the robot and says “Nihongo” (the name of the Japanese language in Japanese). The robot switches to Japanese, makes eye-contact with the new person, and explains in Japanese that it can talk about any topic in Wikipedia, suggesting some favourite topics. The Japanese user also selects Shakespeare, and

this time the robot gets information about Shakespeare from Japanese Wikipedia.

The robot talks about Shakespeare in Japanese, and also explains the Japanese versions of various commands and interactions. The Japanese-speaking user asks in particular about Romeo and Juliet. After 5 minutes he decides to stop and then the English-speaker returns. He simply says “English” and the robot switches back to English speech. The video ends during this part of the interaction.

3 Description of the demonstrated system

The demonstrated system addresses the problem of open-domain interaction, i.e. how to enable robots to talk fluently about an unlimited range of topics. Given that companion-type interactive applications are expected to become more popular in the future there is a need for systems that can chat and entertain the human users on an unlimited range of topics, and the system’s ability to change topics fluently and find relevant information is important. The impact of multilingual robot agents which are capable of talking in such situations is huge, not only from the technological point of view but also considering how they affect human life: such interaction skills will make the world more complex but also extend human cognitive, physical and interaction capabilities.

Comparing the demo with existing systems, there are other systems that can read Wikipedia articles aloud, but WikiTalk also smoothly shifts topics in the middle of an article when prompted by the user. For example, as shown in the English-Japanese video (Figure 3), if the robot is talking about Japan and mentions “kanji” when explaining the Japanese name for Japan, the user can say “kanji?” and the system smoothly switches topics and starts talking about kanji after getting information from Wikipedia about this new topic. Details of the implementation of smooth topic shifting in WikiTalk are given by Wilcock (2012). In addition, WikiTalk switches languages smoothly on demand.

One novel aspect of the approach concerns internationalisation. Developers of devices where spoken dialogue systems are used, such as robots, can help internationalisation by providing better interfaces to enable better synchronisation of different modalities, for example audio and gestures or modules for detecting the gender of the user. This enables the robot system to address better the unique functional property, namely to talk about an unlimited range of topics using Wikipedia. Another aspect is that the system uses reliable and up-to-date information written and edited by humans in Wikipedia. Detailed discussions of internationalisation and localisation are given in (Laxström et al., 2016).

The system is also being applied in the revitalisation of endangered languages by the use of language and speech technologies. In Finland, the DigiSami project (Jokinen et al., 2016) is developing a Sami-speaking robot application based on WikiTalk, in order to encourage the North Sami language community in Lapland to view their language as a language with a future as well as a past. This SamiTalk application is described by Wilcock et al. (2016).

4 Future plans

Future versions of WikiTalk are likely to include new language localisations such as French, German and Dutch versions, which will use information from French, German and Dutch Wikipedias. WikiTalk will be developed by CDM Interact (www.cdminteract.com), a Finnish social robotics company.

WikiTalk will also be available for Pepper robots, which use the same Naoqi operating system used by Nao robots. Previously, Pepper robots were only available in Japan and at first they only spoke Japanese, but now Pepper robots are available in Europe and speak several European languages like Nao robots. Of course, the bilingual English and Japanese version of WikiTalk which is already available on Nao is very suitable for Pepper robots in Japan.

Future plans also include a version of WikiTalk for ROS, the open source Robot Operating System (Wikipedia, 2016), which is used by a wide range of robots from different manufacturers. A ROS version of WikiTalk will therefore be able to run on many different future robot models.

Although a ROS version of WikiTalk will not be restricted to Nao robots, it will still be usable with Nao by means of the `naoqi_bridge` interface which is part of ROS. This interface allows ROS components to invoke the functions of the Nao robot, to control for example its walking and talking using its Naoqi operating system. The same `naoqi_bridge` interface also means that ROS WikiTalk will be usable

with Pepper robots, which also use the Naoqi system. ROS WikiTalk will be able to use the robots' own face-tracking, nodding and gesturing capabilities to support interaction management and the presentation of new information on humanoid robots like Nao and Pepper.

There are many application opportunities for this type of system, where a talking robot is connected to internet-based digital information sources. For example, one area is in applying new technology to education, and another is in providing robot companions for elderly people.

Acknowledgements

We thank Niklas Laxström for his work on the internationalization of WikiTalk and the localized Finnish version, and Kenichi Okonogi for his help with the localized Japanese version.

References

- Dimitra Anastasiou, Kristiina Jokinen, and Graham Wilcock. 2013. Evaluation of WikiTalk - user studies of human-robot interaction. In *Proceedings of 15th International Conference on Human-Computer Interaction (HCI 2013)*, Las Vegas.
- Adam Csapo, Emer Gilmartin, Jonathan Grizou, JingGuang Han, Raveesh Meena, Dimitra Anastasiou, Kristiina Jokinen, and Graham Wilcock. 2012. Multimodal conversational interaction with a humanoid robot. In *Proceedings of 3rd IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2012)*, pages 667–672, Kosice.
- JingGuang Han, Nick Campbell, Kristiina Jokinen, and Graham Wilcock. 2012. Investigating the use of non-verbal cues in human-robot interaction with a Nao robot. In *Proceedings of 3rd IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2012)*, pages 679–683, Kosice.
- Kristiina Jokinen and Graham Wilcock. 2012a. Constructive interaction for talking about interesting topics. In *Proceedings of Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul.
- Kristiina Jokinen and Graham Wilcock. 2012b. Open-domain conversations with humanoid robots. In *COLING 2012 Tutorial*, 24th International Conference on Computational Linguistics (COLING 2012), Mumbai.
- Kristiina Jokinen and Graham Wilcock. 2013. Open-domain information access with talking robots. In *14th Annual SIGdial Meeting on Discourse and Dialogue: Proceedings of the SIGDIAL 2013 Conference*, pages 360–362, Metz.
- Kristiina Jokinen, Katri Hiovain, Niklas Laxström, Ilona Rauhala, and Graham Wilcock. 2016. DigiSami and digital natives: Interaction technology for the North Sami language. In *Proceedings of Seventh International Workshop on Spoken Dialogue Systems (IWSDS 2016)*, Saariselkä.
- Niklas Laxström, Graham Wilcock, and Kristiina Jokinen. 2016. Internationalisation and localisation of spoken dialogue systems. In *Proceedings of Seventh International Workshop on Spoken Dialogue Systems (IWSDS 2016)*, Saariselkä.
- Raveesh Meena, Kristiina Jokinen, and Graham Wilcock. 2012. Integration of gestures and speech in human-robot interaction. In *Proceedings of 3rd IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2012)*, pages 673–678, Kosice.
- Kenichi Okonogi, Graham Wilcock, and Seiichi Yamamoto. 2015. Nihongo WikiTalk no kaihatsu (Development of Japanese WikiTalk). In *Forum on Information Technology (FIT 2015)*, Matsuyama, Japan. (in Japanese).
- Wikipedia. 2016. Robot Operating System — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Robot_Operating_System [Online; accessed 24-August-2016].
- Graham Wilcock and Kristiina Jokinen. 2015. Multilingual WikiTalk: Wikipedia-based talking robots that switch languages. In *Proceedings of the 16th Annual SIGdial Meeting on Discourse and Dialogue*, Prague.
- Graham Wilcock, Niklas Laxström, Juho Leinonen, Peter Smit, Mikko Kurimo, and Kristiina Jokinen. 2016. Towards SamiTalk: a Sami-speaking Robot linked to Sami Wikipedia. In *Proceedings of Seventh International Workshop on Spoken Dialogue Systems (IWSDS 2016)*, Saariselkä.
- Graham Wilcock. 2012. WikiTalk: A spoken Wikipedia-based open-domain knowledge access system. In *Proceedings of the COLING 2012 Workshop on Question Answering for Complex Domains*, pages 57–69, Mumbai.