

One-way translation: an opportunity for NLG and MT research to interact

Sebastian Varges*

Information and Communication Technology

University of Trento

varges@dit.unitn.it

Abstract

‘One-way translation’ has been proposed by Ward (2002) as an alternative to full-blown speech-to-speech translation. A one-way translation system is a cross-language communication aid located on a wearable device that uses graphical or other input to generate a spoken utterance, for example when buying a train ticket in a foreign language. We argue that this type of application is an opportunity for NLG, MT (and also dialog and speech) research to interact.

1 Introduction: One-way translation

‘One-way translation’ has been proposed by Ward (2002) (although not called by this name) as a novel research area that avoids the unrealistically optimistic assumptions underlying research into speech-to-speech translation. It goes beyond making incremental improvements to existing approaches by proposing a new type of application that more directly focuses on usability. The idea is probably best explained by considering an example: when trying to buy a train ticket in a foreign country whose language the user does not speak, he or she would graphically compose a question on a mobile device, which is verbalized by a text-to-speech system (TTS) and played to the travel agent (“Excuse me, I need to use this device because I do not speak your language. I would like to go to ...”). Since obviously the user will not understand the agent’s an-

The author was supported by European Commission Marie Curie Excellence Grant contract No. 022593 (ADAMACH project).

swer in full, this same device is also used to compose follow-up questions or to ask the agent to write down the departure time. Crucially, the user is assumed to be able to pick up non-verbal cues that indicate, for example, a negative response, even if the details are not understood. In other words, one-way translation takes advantage of the various communication channels that exist between people when interacting directly, in contrast to methods that are intended to operate fully automatically.

2 Research opportunities

One-way translation, as outlined above, is at its core based on natural language generation. However, the application scenario is cross-language communication, a topic usually addressed by research on machine translation. One-way translation also involves aspects of dialogue and speech processing (see below) as well as HCI issues, in particular wearable/mobile devices: the original design of Patokallio and Ward (2001) uses a head-mounted display that does not force the user to look down onto some device, which would interrupt non-verbal face-to-face communication. Thus, the research issues likely to arise are not necessarily found at the straightforward intersection of MT and NLG (which may focus on issues such as large-scale text processing in an offline/batch setting). In the following, we outline some of these research issues.

2.1 The (usual) question of the input

A persistent and crucial issue in research on NLG is the question about the source of the input to the generator (McDonald, 1993). This question arises much

more sharply than in Machine Translation, where the source language input is known. In one-way translation, there seems to be a wide range of sources that could constitute the input to the generator, which in turn proposes sentences/turns to the user for selection. The input to the generator could be supplied semi-automatically, with a focus on reducing the burden of selecting the next user turn:

1. application/domain content selection, esp. at the beginning of a dialog, possibly using menu-based form-filling;
2. dialog move information to suggest possible follow-up moves;
3. sentiment detection in follow-up dialog turns: the user could select negative/positive/unknown response types. Alternatively, this may be supplied by specialized classifiers analysing the ‘source language response’. Such a classifier may also attempt dialogue act tagging (Bangalore et al., 2006) – however, since one-way translation is a semi-automatic approach, this is not strictly required;
4. specialized number and Named Entity speech recognition in the source language, to be incorporated into requests for confirmation.

As these examples show, the input to the generator which drives the TTS system can be varied, including propositional content and dialog moves, and may be partially probabilistic or using confidence scores (for classifier and partial speech recognition). Ultimately, the user selects from a range of options. Patokallio and Ward (2001) assume that the user selects sentences/turns from a menu of no more than 5 items. In our view, this could be a staged and interactive process with a larger range of options (on a larger tablet-like display without using a head-mounted display, for example). Presenting those options in a suitable form requires forms of language generation. To reduce the need for entering information during human-human dialog, there could be an initial planning phase, as suggested in (Patokallio and Ward, 2001). Negative feedback from the dialog participant could trigger the generation of options for various dialogue repair strategies.

3 Discussion: any Machine Translation?

While it seems obvious that one-way translation is based to a substantial part on NLG, the connection to MT is, despite the name, less clear. This is true particularly of statistical MT which does not use any dedicated interlingua representations or NLG phase, and maps source language to target language more or less directly. However, even in a one-way translation scenario such statistical mappings may be useful, for example for identifying possible words to use, given some cues about the source language input. Since data collection for one-way translation seems difficult, the ability to make use of standard aligned corpora could be of great benefit.

To our knowledge, there have only been very few attempts at one-way translation, and it is thus too early to draw conclusions on either the practical feasibility of the approach or its scalability: what happens, for example, when domains and dialogues become more complex and the range of options for the user to choose from multiplies? However, we believe that one-way translation of the form outlined above offers an interesting opportunity for several research fields, including NLG and MT, to interact. It challenges some of the standard assumptions of these fields, for example about the homogeneity of the input. Furthermore, we believe that systems – and research agendas – should be built bottom-up starting from small but working systems, but with a high-level vision in mind.

References

- Srinivas Bangalore, Giuseppe Di Fabbrizio, and Amanda Stent. 2006. Learning the structure of task-driven human-human dialogs. In *Proceedings of COLING/ACL*, pages 201–208, Sydney, Australia, July.
- David D. McDonald. 1993. Issues in the Choice of a Source for Natural Language Generation. *Computational Linguistics*, 19:191–197.
- Jani Patokallio and Nigel Ward. 2001. A Wearable Cross-language Communication Aid. In *Proceedings of the Fifth International Symposium on Wearable Computers (ISWC'01)*, page 176.
- Nigel Ward. 2002. Machine Translation in the Mobile and Wearable Age. In *Proceedings of the 9th International Conference on Theoretical and Methodological Issues in Machine Translation*.