

Designing an Artificial Robotic Interaction Language*

Omar Mubin

Department of Industrial Design, Eindhoven University of Technology, The Netherlands
{o.mubin,c.bartneck,l.m.g.feijs}@tue.nl

Abstract. The project described hereunder focuses on the design and implementation of a “Artificial Robotic Interaction Language”, where the research goal is to find a balance between the effort necessary from the user to learn a new language and the resulting benefit of optimized automatic speech recognition for a robot or a machine.

Keywords: Artificial Languages, Speech Interaction, Automatic Speech Recognition.

1 Introduction

Speech is a natural means of information exchange for humans. Therefore, improving speech interaction technology in Human Computer Interaction could lead to a more pleasant interaction. Speech Interaction is confronted by various issues, such as: ambiguity in natural dialogue, un-robust speech recognition and un-synchronization between software and hardware [1]. Recent attempts to improve the quality of the technology of automatic speech recognition for machines have not advanced enough. The limitations prevailing in current speech recognition technology for natural language is a major obstacle behind the unanimous acceptance of Speech Interfaces. Existing speech recognition is just not good enough for it to be deployed in natural environments, where the ambience also influences its performance.

Generally in speech interfaces the focus is on using natural language and given their unsuitability, it is time to find a different balance in the form of a new language. Recent research in speech interaction is already moving in this direction, as stated in [2], constraining language is a plausible method of improving recognition accuracy. In [3] the user experience of an artificially constrained language (“Speech Graffiti”) was evaluated and it was concluded that 74% of the users found it more satisfactory than natural language and also more efficient in terms of time. The field of handwriting recognition has followed a similar road map. The first recognition systems for handheld devices, such as Apple's Newton were nearly unusable. Palm solved the problem by inventing a simplified alphabet called Graffiti which was easy to learn for users and easy to recognize for the device. Therefore we aim to construct an “Artificial Robotic Interaction Language” where an artificial language is a *language deliberately invented or constructed, especially as a means of communication in computing* [4]. In linguistics, there are numerous artificial languages which address a user

* Advisors: Christoph Bartneck and Loe Feijs.

perspective by making communication between humans easier and/or universal; however there has been little or no attempt to optimize a spoken artificial language for automatic speech recognition. Therefore the main goal of our research is constructed on the basis of two sub goals. Firstly the language should be learnable by the user and secondly, the language should be optimized for efficient automatic speech recognition by a machine or a robot.

2 Overview of Artificial Languages

As a first step of our research an overview of artificial languages has been carried out [5], to ascertain what we could learn from existing Artificial Languages, especially in reference to what could be easier to learn for humans. The overview was carried out across two aspects, namely morphology or grammar and phonology. Various encyclopedias [6] define the major properties of a language of which morphology and phonology are two key aspects.

2.1 Morphological Overview

In summary, there are two major approaches of morphological design amongst artificial languages: The first is to have very few grammatical markings, leaving it to the interpretation of the speakers and the context, and secondly, some languages have inflections but their grammatical rules are consistent across all words. The question that emerges is that which grammar type would be easier to learn and which would be less ambiguous. Shown below (see Table 1) are the two main approaches of grammar design amongst artificial languages. Grammar-I has very few grammatical markings (as indicated by the 'X') and therefore interpretation is determined by the speakers, the context or the word order. Grammar-II has inflections but the grammatical rules are consistent across all words within each category.

Table 1. Two main grammar types in artificial languages

Grammatical Category	Grammar I	Grammar II
Aspect	X	X
Case	X	Basic Levels: for e.g. Possessive, Nominative
Gender	X	Male, Female
Mood/Modality	X	X
Numbering	X	Singular, Plural
Person References	First, Second, Third	First, Second, Third
Tense	X	Past, Present, Future
Polarity	Positive, Negative	Positive, Negative
Voice	X	Active, Passive
Definiteness (Articles)	X	Definite (the), Indefinite (a)
Word Order	SVO	SVO

2.2 Phonological Overview

Another important metric upon which artificial languages could be overviewed was determined to be the domain of Phonetics. Deciding the manner in which to use phonemes as part of our phonological classification of artificial language was an important decision. Extending from our research goal of designing an interaction language that is easy to learn for humans, we extracted a set of the most common phonemes/segments present in the major languages of the world, based on the number of speakers of a language as indicated in the Ethnologue [7]. The overview utilized the UCLA Phonological Segment Inventory Database (UPSID), see [8] and [9]. The database provides a large inventory of all the existing phonemes of 451 different languages of the world. The number of phonemes documented in the database amount to 919. What we were seeking for was a list of phonemes found in only the major languages of the world. This resulted in a net total of 23 phonemes [5]. Interesting trends were observed; certain dental consonants were not found in any of the artificial languages that we overviewed. One reason why this might have occurred is that most artificial languages stem from Germanic or Western languages, whereas the dental consonants are found in Indic or Asian languages only. Trends that have been observed in natural languages with regards to the most common consonants (for e.g. 'm', 'p') were to some extent replicated for the case of artificial languages. The mirroring effect between natural and artificial languages extended to vowels as well. The vowels 'a', 'o' and 'u' were found in all the artificial languages that we classified and the vowels 'i' and 'e' were absent in only one artificial language.

3 The Design of the Artificial Interaction Language

We have presented a morphological overview of artificial languages where, two primary grammar types were discussed. In the future, we aim to evaluate which of the mentioned grammar types will be easier to learn for our intended artificial language and which will be less ambiguous, using methods as advocated in [10]. Moreover, our phonological overview has revealed a set of phonemes that might be desirable to include in the artificial language to render it conducive for human learnability. However for both aspects of morphology and phonology what also needs to be determined is how both could contribute to improve speech recognition. For example unique phonemes that have less confusion amongst them would be easier to recognize [11]. Similarly, selecting a particular grammar type could also influence the quality of speech recognition, and we aim to determine this in the future. Another interesting variable is word length, as shorter words tend to be confused with each other in automatic speech recognition. Therefore we also aim to evaluate the role the length of a word could play in improving the quality of speech recognition.

As a first step in the design process we aim to inherit the vocabulary set or word concepts of the simple artificial language Toki Pona [12]. It has 118 word concepts and sufficiently caters for the needs of a simple language. We aim to adapt the pronunciation of the words of Toki Pona based on the requirements of word length and phonetic information. For example, given that Toki Pona is a simple language it has some words which are very short; of course to be easier to learn for humans. However to assist speech recognition, some of its words will need to be elongated based on a specific methodology, which will attempt to improve the phonetic discernability of

words and also be scalable to allow the generation of new words. Additionally, we aim to start the design from Grammar Type II and gradually remove grammatical markings and rules to move towards Grammar Type I, an evolutionary trend that can also be noticed in natural languages [13].

In the future, we shall also investigate speech recognition engines to ascertain “*the exact criteria that makes speech easy for machines to recognize*”. As a combination of our initial endeavors we will then move towards designing the “Artificial Robotic Interaction Language”. Our intention is to carry out future research in the form of several cycles as a spiral model. Each cycle typically would have four phases: requirements, design, implementation and evaluation.

Contribution to HCI. We intend to deploy our interaction language within the domain of robotics, however our proposed interaction language does not necessarily have to be restricted to robots only, but it could be applied to any behavioral products that employ speech interaction. Moreover, HCI is moving towards the domain of Ambient Intelligence, where technology is invisible in the background but there are still objects that mediate/interact, for e.g. robots. Therefore the design of an Artificial Robotic Interaction Language would lie in the heart of next generation HCI.

References

1. Kulyukin, V.A.: On natural language dialogue with assistive robots. In: ACM conference on Human-robot interaction, pp. 164–171 (2006)
2. Rosenfeld, R., Olsen, D., Rudnicky, A.: Universal speech interfaces. *Interactions* 8, 34–44 (2001)
3. Tomko, S., Rosenfeld, R.: Speech Graffiti vs. Natural Language: Assessing the User Experience. In: Proceedings of HLT/NAACL (2004)
4. A Dictionary of Psychology. In: Colman, A.M. (ed.) Oxford Reference Online. Oxford University Press, Oxford (2006)
5. Mubin, O., Bartneck, C., Feijs, L.: A Morphological and Phonological Overview of Artificial Languages. In: Interspeech 2009 (2009) (submitted)
6. David, C.: The Cambridge encyclopedia of language. Cambridge University Press, Cambridge (1997)
7. Gordon, R., Grimes, B.: *Ethnologue: Languages of the World*. Summer Institute of Linguistics, SIL International (2005)
8. UPSID Info., http://web.phonetik.uni-frankfurt.de/upsid_info.html
9. Maddieson, I.: *Patterns of Sounds*. Cambridge University Press, Cambridge (1984)
10. Mubin, O., Shahid, S., Bartneck, C., Krahmer, E., Swerts, M., Feijs, L.: Using Language Tests and Emotional Expressions to Determine the Learnability of Artificial Languages. In: CHI 2009 Conference on Human Factors in Computing Systems, pp. 4075–4080 (2009)
11. Hinde, S., Belrose, G.: Computer Pidgin Language: A new language to talk to your computer? Hewlett-Packard Laboratories
12. Toki Pona - the language of good, <http://www.tokipona.org/>
13. Grudin, J., Norman, D.: Language evolution and human-computer interaction. In: Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, pp. 611–616 (1991)