A System to Control Language for Oral Communication

Laurent Spaggiari[1] and Sylviane Cardey[2]

[1] Human Factors Dept., EDYDNX, section 527, M0151/0, Airbus Operations SAS, 316 route de Bayonne, F-31060 Toulouse, France
[2] Centre Tesnière, Université de Franche-Comté, UFR SLHS, 30 rue Mégevand, F-25030 Besançon Cedex, France
laurent.SPAGGIARI@airbus.com, sylviane.cardey@univ-fcomte.fr

Abstract. In this paper we discuss the use of controlled languages not for written texts but for oral communication which has never been done before, and this in safety critical domains. Interference between languages could effectively cause accidents due to misunderstanding of messages whatever they are. We discuss how firstly we could automatically detect eventual possibilities of misunderstanding due to mispronunciation or bad interpretation and secondly how to prevent these problems by using controlled languages. We show that our methodology, which is intensional in nature, is much more productive than working in extension.

Keywords: Controlled language, oral communication, language interferences.

1 Introduction

In this paper we discuss the use of controlled languages not for written texts but for oral communication which has never been done before, and this in safety critical domains. Interference between languages could effectively cause accidents due to misunderstanding of messages whatever they are. Though research concerning interferences between different languages has been carried out at Airbus Operations SAS and in Centre Tesnière, this has been written-text oriented and not oral-text oriented. This paper is concerned with Airbus Operations SAS and Centre Tesnières' research into helping in the design of controlled languages which are optimized for computer-human communication.

We discuss how firstly we could automatically detect eventual possibilities of misunderstanding due to mispronunciation or bad interpretation and secondly how to prevent these problems by using controlled languages. We show that our methodology, which is intensional in nature, is much more productive than working in extension.

2 Natural Language

Whether it be oral, written or signed, a communication will be considered as successful and efficient when the received message complies with the mental process used for reconstructing and interpreting the information within the message. However, natural language not only allows everyone to create many variations for the same

H. Loftsson, E. Rögnvaldsson, S. Helgadóttir (Eds.): IceTAL 2010, LNAI 6233, pp. 393–400, 2010.

expression – *Paul sold the car* versus *The car was sold by Paul* – but it is also intrinsically ambiguous from the following points of view:

- semantics (word): *Parle-moi de ta nouvelle pièce*
  (Tell me about your new room/coin/play)
- semantics (sentence): *Paul donne un os à son chien pour s'amuser*
  (Paul gives his dog a bone for his own fun/for the dog to play with)
- phonetics (word): *il est assis sur [letal*O~*]*
  (he is riding (sitting on) the stallion/he is squatting on his heels)
- phonetics (sentence): *[sEtwazola]*
  (you are Zola/that guy/that bird)
- syntax (word): *Paul préfère les gâteaux au chocolat*
  (Paul prefers chocolate cakes/Paul prefers cakes to chocolate)
- syntax (sentence): *Flying planes can be dangerous*
  (this can be a dangerous work/ aircraft above your head can be dangerous)
- pragmatics (word): *Le magasin est ouvert le dimanche*
  (the store is open only/even on Sunday)
- pragmatics (sentence): *Paul a dit qu'il viendrait*
  (Paul said he (Paul/Jean) would come)

In certain highly technical domains such as nuclear/surgery/chemistry/aeronautics, etc. safety is crucial and some situations require from the operators immediate corrective actions and to succeed in their tasks, they needs to understand fully and immediately the situation, i.e. what they are expected to do, the consequences, etc. In this respect, natural language, too broad, too variable cannot be used.

3 Controlled Languages

Contrary to natural language, controlled languages (CLs) are favoured by industry because they refer to systems that limit the number of core vocabulary words, of applicable grammar and stylistic rules. "Industry does not need Shakespeare or Chaucer; industry needs clear, concise communicative writing – in one word Controlled Language" [1]. Their objective is to reduce ambiguities, complexity, colloquialisms and synonyms in order to improve consistency, readability, translatability and retrieval of information [2]. "Consistency is one of the most basic usability principles [...] the same information [...] should be formatted in the same way to facilitate recognition" [3].

This concept of CL is not really new. Indeed, in the 1930s, C. K. Ogden, a linguist, developed British American Scientific International Commercial English to help students write in a clearer way and to make the non Anglophone students' training easier [4]. Since then, many of CLs have been developed for different purposes, e.g. TAUM meteo, Air/Sea/police speak, Douglas Aircraft, ScaniaSwedish, Caterpillar Fundamental/Technical English, Kodak International Service Language, and, the most widely used for writing aircraft maintenance procedures, ASD-STE 100 (formerly known as AECMA Simplified English [5]); for a more complete overview, refer to [6]. These CLs are not "simple" or "baby" English but true simplified English.

Although English is a very productive natural language for the creation of CLs as it is the current international language used for trade and science, other languages such as German, Chinese, Swedish and French are also used for the creation of CLs.

4 The Oral Aspect

Creators of CLs usually base their grammar restrictions on well-established writing principles (e.g. "write short sentences with only one topic; avoid passive form ..."). Furthermore, despite the fact that these languages do not have many rules in common [7], they do share one main characteristic: they deal with the written aspects of language, and not with the oral.

It is exactly this oral aspect that we address here. The fact is that messages are not only read but can also be heard using synthetic/recorded voices in nuclear plants and airports for example. Because the receiver of the message may not have the same mother tongue as the one he/she hears and because one cannot expect him/her to master it, a syntactically and lexically controlled message may not be sufficient. Indeed, when looking at the following pairs, one can easily imagine the potential consequences in case of a misunderstanding:

– *increase the temperature* versus *decrease the temperature*
– *the gear is uplocked* versus *the gear is unlocked*
– In French for an Anglophone *"dessus" and "dessous"* will sound the same

As a further example, ambiguities can result from English phonemes not present in Thai, as is illustrated in Table 1 [8]. We do not enter into the complexity of the matter here, but we can already see that *half* for example becomes *harp* as well as *ball,* which is pronounced *born*.

Table 1. Ambiguities resulting from English phonemes not present in Thai (the phonetic transcriptions are in SAMPA)

English phonemes	Sound in Thai	Ambiguities
[T]	[t], [d]	birth → bird
[D]	[d]	they → day
[f]	[p]	half → harp
[v]	[w]	vine → wine
[s]	[d]	bus → bud
[z]	[t]	buzz → but
[l]	[n]	ball → born
[r]	[l]	free → flee

5 A System to Help When Creating Sentences That Are to Be Pronounced

Based on the observations in the previous section, we have devised a system that can help when creating sentences that are to be pronounced. This system has the ability to detect within a list not only all the homophones (e.g. *night knight*) and the minimal pairs (e.g. *brake brain*) but also quasi homophones (e.g. *increase decrease*) for a proposed word, according to the source language (North American English).

5.1 The Database

The database we used for checking the pronunciation is the Carnegie Mellon University Pronouncing Dictionary, also known as 'cmudict'. This dictionary is a public domain machine-readable pronunciation dictionary for North American English that contains over 130,000 words and their phonetic transcriptions. The pronunciation of the words is encoded using a modified form of the Arpabet system, each phoneme having a unique code (e.g. *ABRACADABRA AE2 B R AH0 K AH0 D AE1 B R AH0).* This dictionary is used in different projects such as the Festival speech synthesis system and also the CMU Sphinx speech recognition system.

As a result of this database, our system was able to retrieve all the words with the same pronunciation. However, we also wanted to obtain quasi homophones (e.g. *increase decrease*) from this list. So we devised an algorithm that looks at the phonetic differences between words.

5.2 The Algorithm

The algorithm performs the following steps:

- Calculation of the number of phonemes for the submitted word (11 for *abracadabra*)
- Retrieval from the database of all the words that have:
  - The same number of phonemes
  - The same number +1 of phonemes
  - The same number -1 of phonemes
- Calculation of the similarity (number of different phonemes in the same order) between the submitted word and the retrieved words
- Calculates the proximity between the two phonetic strings using the Levenshtein distance function.

Levenshtein distance is defined as the minimum number of necessary characters to be changed, inserted, or modified for transforming a string into another one. It is commonly used for spelling checking, speech recognition, DNA analysis, etc.

The system works by requesting a check for a specific word, one by one. So, it is not possible to get statistics on the numbers of pairs of words retrieved for a specific language. The algorithm is able to retrieve any string of any length (monosyllabic to very long words). The number of retrievals decreases with the length of the submitted word. Also, the words retrieved are sometimes irrelevant (different enough to be not mistaken). This is due to the fact that we weakened the constraints. Indeed, a difference of 2 phonemes is considered in our algorithm. Also, we did consider the phoneme itself as a whole and not as a sum of phonological features. We think the retrievals will be more relevant when we will look at the divergences of phonological features and only consider one phoneme of difference instead of 2 as at present.

5.3 Oral Communication Involving Different Mother Tongues

The receiver of the message can have a different mother tongue from English. Consequently, some phonemes may not exist for him or her. In this case, what will the receiver understand? Assuming the fact that he/she will reconstruct words using existing phonemes in their own language, the system should replace the phonemes by those existing in another language and check in the database if homophones or quasi homophones exist.

To do this, we devised a resource that gives, for each language (Arabic, French and Chinese for the moment), the list of non-existing English phonemes and their counterpart(s). Table 2 illustrates an extract from this table.

Table 2. Non-existing English phonemes and their counterpart(s) (extract)

From English to		
Phonemes	**Replaced by**	**Language**
Ax	Aa	Arabic
Ax	Ae	Arabic
Ax	Ao	Arabic

As a result of this resource, our system is now able, depending on the language selected, to reconstruct the pronunciation and then retrieve all the words with the same pronunciation in English. An illustration of the system's interface showing the results for tomato is shown in Fig. 1.

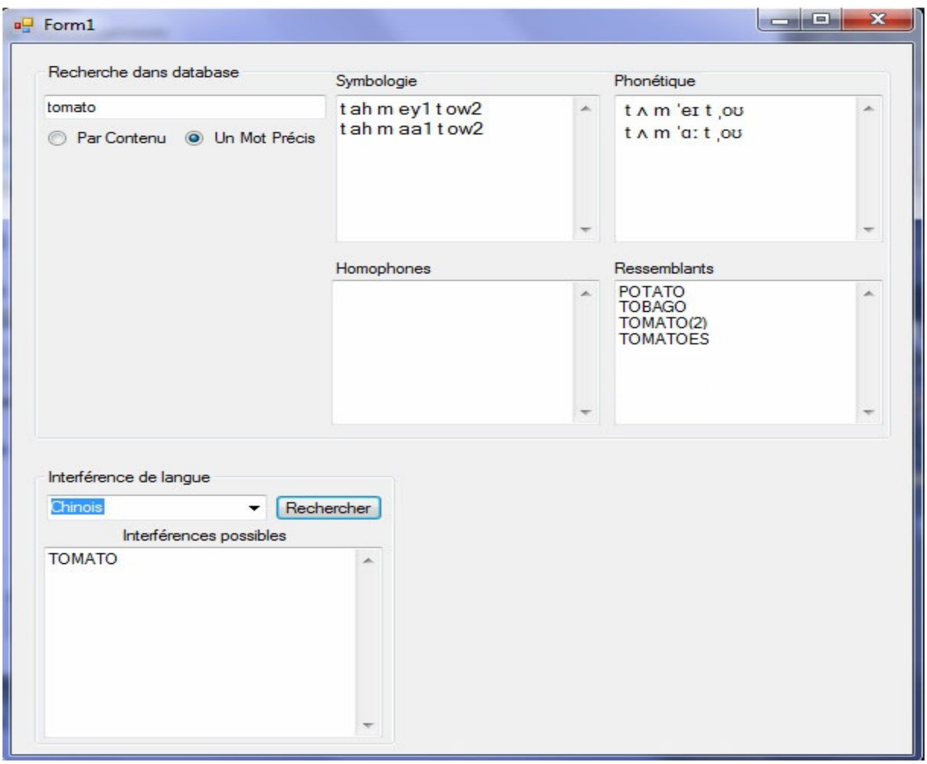

Fig. 1. Screen shot of the 'System to Help when Creating Sentences that are to be Pronounced' for the word *tomato*

6 Results and Improvements

The system behaves as intended because:

– it retrieves all the homophones for a term (*feel* (*F IY1 L*) → *fiel, feil, foell*).
– it retrieves all the minimal pairs for a term (*feel* (*F IY1 L*) → *fail, fall, feat, feed, fees, fell, file, fill, foal, foil, fool, foul, fowl, full, peal, peel*).
– it retrieves quasi homophones for a term depending on the language selected (*thought* (*TH AO1 T*) → for Chinese: *fought, sawed, sod, sought* and for French: *fought, sought, taught, taut, tot*)

Using this information, one can easily decide, when creating a spoken message, if one can use a word or if one should change it (e.g. use *reduce* instead of *decrease*) or even reformulate the whole sentence.

However, when looking just at English, some of the words that were retrieved could be avoided as they are different enough not to be mistaken. We think that the reasons for this lie in the constraints (insufficient) we have applied for calculating the similarity. Indeed, we performed this calculation by counting the number of different phonemes between two words. Also, to be able to retrieve for example *increase decrease*, we had to consider an acceptable number of 2 differences whatever they are. As a consequence, many words with 2 differences are retrieved. A better way would be to take into account for each phoneme its phonological features, and to count the different ones to get a much more precise result. For example, we would continue to take a difference of 2 phonemes as a maximum, but, reducing this time the maximum number of differences allowed between features, we would reduce the numbers of results. The 2 different phonemes in *increase decrease* share the same phonological features except one: nasal vs. oral.

Another improvement would consist in considering the whole sentence, that is to say, to consider the assimilations that occur between the words once these are put together.

Finally, the phonetics of the reconstructed sentences could be saved in an ssml file [9]:

```
<speak version="1.0"
xmlns="http://www.w3.org/2001/10/synthesis"
xml:lang="en-US">
<phoneme ph="tʌm'eɪt,oʊ">tomato</phoneme>
</speak>
```

This file can easily be enriched with plenty of information concerning prosody and style (voice, emphasis, break, pitch, speaking rate and volume of the speech output), text structure etc. This would allow us to use this file as an entry to obtain these phonetic strings pronounced by a synthetic voice (for example Microsoft US English Anna). An illustration of the system's interface enriched for voice synthesis (word by word only) is shown in Fig. 2.

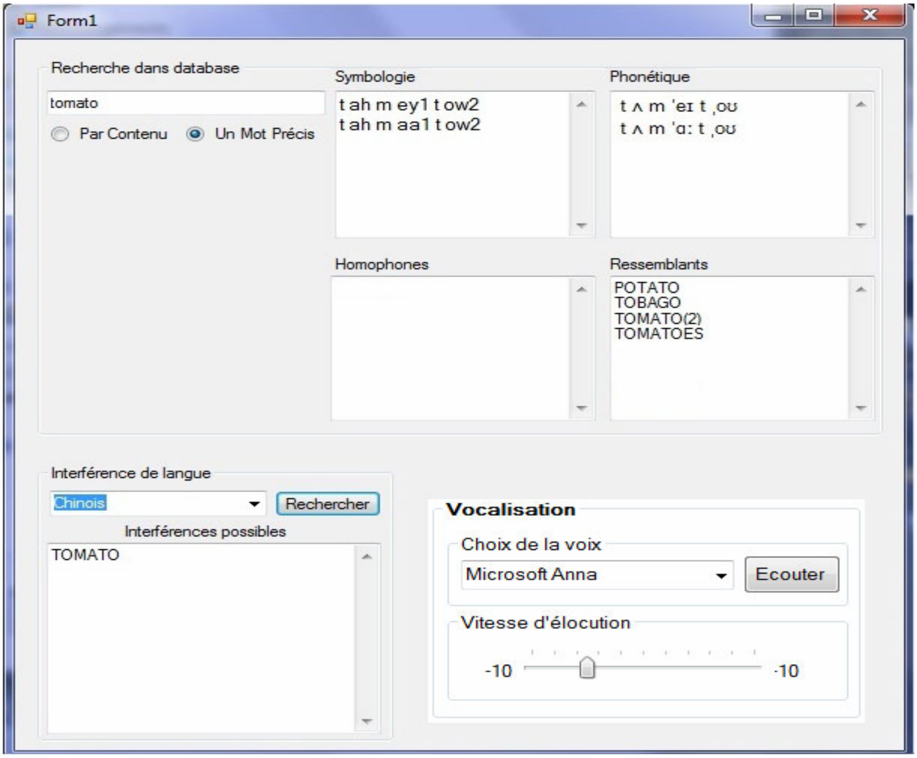

Fig. 2. Screen shot of the 'System to Help when Creating Sentences that are to be Pronounced' enriched for voice synthesis

7 Conclusion

We have seen in this paper why language interferences have to be avoided and we have proposed a methodology and a system to find and solve these problems. Some work about interference had already been done in our laboratory with native speakers but our system automatically detecting possible interference revealed itself much more efficient. Much work still has to be done at the level of the boundaries between words but the methodology which consists in working at the level of phonemes and distinctive features rather than trying to find individual words seems to be more productive and easier to generalise for solving the problems of interferences which are due to bad pronunciation or bad interpretation. The methodology used allows tracing back to the cause of the problems which is essential in safety critical applications. The results of this research can be applied to different domains. This is because the specific data (i.e. lexicon by domain) is tested against the general pronunciation dictionary as we cannot know the level of English people have. So the methodology is not domain dependant and can be applied to any specific domain as long as the domain has its own dictionary.

References

1. Goyvaerts, P.: Controlled English, Curse or Blessing? A User's Perspective. In: Proceedings of CLAW 1996, Leuven, Belgium, March 26-27, pp. 137–142 (1996)
2. Huijsen, W.O.: Controlled Language: An Introduction. In: Proceedings of CLAW 1998, Pittsburg, USA, May 21-22, pp. 1–15 (1998)
3. Nielsen, J.: Usability Engineering. Academic Press Ltd., London (1993)
4. Ogden, C.K.: Basic English: A General Introduction with Rules and Grammar, 8th edn. (1940); PSYCHE 29, K. Paul, London, England, (1930)
5. AECMA Simplified English Association Européenne des Constructeurs de Matériel Aérospatial. A Guide for The Preparation of Aircraft Maintenance Documentation. The International Aerospace Maintenance Language', Gulledelle, Brussels, Belgium (1995)
6. Gavieiro-Villatte, E., et al.: Open-Ended Overview of Controlled Languages. In: BULAG 24, Besançon, France (1999)
7. O'Brien, S.: Controlling Controlled English: An Analysis of Several Controlled Language Rule Sets. In: Proceedings of EAMT-CLAW 2003, Dublin, Ireland, May 15-17, pp. 105–114 (2003)
8. Cardey, S.: How to avoid interferences with other languages when constructing a spoken controlled language. In: Proceedings of the International Conference « La comunicazione parlata/Spoken Communication », Naples, Italy, February 23 -25 (2006)
9. Speech Synthesis Markup Language (SSML) Version 1.1, `http://www.w3.org/TR/2010/PR-speech-synthesis11-20100223/`