Building a Natural Language Hindi Speech Interface to Access Market Information

Ahmed Imran and Sunil Kumar Kopparapu
TCS Innovation Labs - Mumbai
Tata Consultancy Services
Yantra Park, Pokharan Road 2, Thane (West), India
Email: {Ahmed.Imran, SunilKumar.Kopparapu}@TCS.Com

***Abstract*—It is a well known fact that majority of rural India earns its livelihood from agriculture and farming. Although India is a net exporter of various agricultural products, the farmer who happens to be the primary producer, has remained *information poor* which puts him at a disadvantage. With little or no knowledge of prices at the markets, farmers have no leverage to negotiate better prices for their produce. Speech based solution can address this issue of market price information availibility to farmers. Speech based solutions are increasingly being used for transaction but they are both (a) restricted to menu based type interactions where a series of interactions are required for the transaction to take place and (b) primarily built for the English literate population synonymously urban population. Paradoxically, the benefit of a speech based solution is best reaped by the rural folks speaking their native language (very often non-English) because the other modes of transactions are either not readily available to them or if available difficult to use. In this paper, we develop a natural language Hindi speech interface to enable Hindi speaking population access market prices of commodities.**

***Index Terms*—Mandi Bhav Jankari; Natural Language; Hindi Speech Interface;**

I. Introduction

To quote World Agriculture Group [1], in order to have the best advantage in selling the produce, one of the basic facilities that the farmer should have is access to information regarding the market conditions as well as about the ruling prices, enabling better returns to the farmers. It is a well known fact that majority of rural India earns its livelihood from agriculture and farming. Although India is a net exporter of various agricultural products, the farmer who happens to be the primary producer, has remained *information poor* which puts him at a disadvantage. With little or no knowledge of prices at other Mandi, farmers have no leverage to negotiate better prices, this is a major disadvantage. In this scenario the farmer is forced to accept the prices dictated by the middlemen and purchasing agents at Mandi to sell their produce without an option to sell at a different Mandi where they might get a better price. This information-gap aspect has not gone unnoticed and several corporates have pitched in as part of their corporate social responsibility to bridge the information gap. As a fallout, in the last couple of years there have been a couple of technological initiatives to address this problem, e-Choupal [2] being one of them. It was launched in 2000 with the installation of Internet enabled terminals in various villages that let farmers check the current market prices of their goods, negotiate and ultimately sell it. In early 2009 several telecom operators enabled a commodity price sharing application, called *Mandi Bhav*, to enable access to prices of commodities on mobile phones [3]. While these are commendable initiatives they are loaded against the rural population because (a) there is a need for a company to set up a Internet kiosk in the village and the actually working of that infrastructure [2], (b) there is a need to have access to a GPRS enabled mobile phone for a WAP based *Mandi Bhav* and (c) the need to send an SMS in local Indian language to get commodity prices in case of an SMS based *Mandi Bhav*. [4] presents a speech recognition enabled *Mandi Bhav* informaion system in Telugu language and similarly, [5] presents a speech-driven application designed and deployed in a community center in rural Tamil Nadu, India, to disseminate agricultural information to village farmers. However, these speech enabled systems have a menu driven IVR (Interactive Voice Response) interface opposed to a natural language speech interface. Primarily there has not been much focus on the ease of use (usability) of Indian language speech enabled interaction systems for the rural farmers.

In this paper we discuss a natural language Hindi speech interface for accessing commodity prices called *Mandi Bhav Jankari* (*Mandi Price Information* in English). The speech enabled IVR system has capability to understand naturally spoken Hindi queries related to market prices of commodities. The rest of the paper is organized as follows. In Section II we describe the need of a *Mandi Bhav Jankari* system followed by a discussion on the implementation in Section III focusing on building capability of the system to recognize spoken queries in natural language. Section IV presents personalizing the system to (a) enhance the speech recognition accuracy and more importantly (b) the use experience of the farmers, we conclude in Section V.

II. The Need: *Mandi Bhav Jankari*

Traditionally, in the Indian context, the produce by farmers are purchased by agents from the local Mandi at a throw away price and a long chain of intermediaries is involved in buying the produce from farmers and moving it to the Mandi's. Often, the middlemen set the price of the produce based on their judgment which is often unfair and not favoring the farmers.

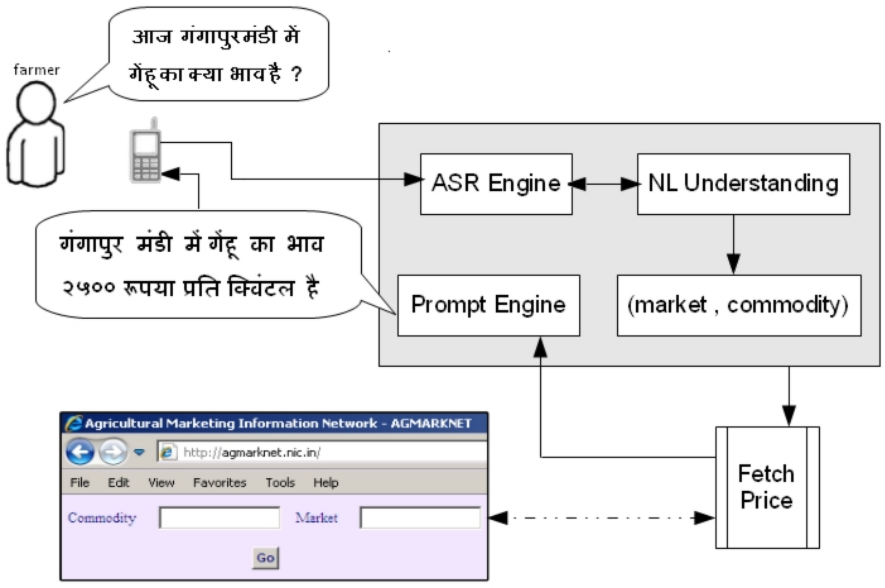

Fig. 1. Speech Enabled Mandi Bhav Jankari System Implementation.

Lack of easy access to the *actual* price of the commodity put him in a position of no choice but to sell at the asked price and is forced to sell his produce in the local Mandi at the poor prices set by the purchasing agents. One way of avoiding this unfavorable situation is to have information about commodity prices be made easily available to the farmers.

Clearly there is a need for a system that gives (a) the farmers access to the commodity prices at different Mandi's and (b) enables easy use of the system. Initiatives like e-Choupal cater to enabling access to information but is not easy because the farmer has to visit a kiosk to get the information and might not have actual access to the information when he is actually at the Mandi making the sell. Additionally, the interface of computer Internet terminals, its use and acceptance among the rural farmers and infrastructure bottlenecks such as frequent electricity outage and unreliable Internet connectivity are major factor hindering its effectiveness. With the penetration of mobile phones in rural India telecom operators have introduced facilities that enables farmers on the move get access to information through Mandi Bhav information service, however the ease of use remains a constraint especially because of the need to communicate to the farmer in his local language. A much easier and probably a more natural interface for the rural folks in general and the farmers in particular is a speech enabled system that answers queries of the farmers in their local language spoken in a natural way. [4] presents a speech recognition enabled *Mandi Bhav* informaion system in Telugu language and similarly, [5] presents a speech-driven application designed and deployed in a community center in rural Tamil Nadu, India, to disseminate agricultural information to village farmers. However, these speech enabled systems have a menu driven IVR (Interactive Voice Response) interface opposed to a natural language speech interface.

III. *Mandi Bhav Jankari* IMPLEMENTATION

A Mandi price information system typically is the intersection of two inputs, namely, (a) Commodity name (Cabbage, Tomato) and (b) the name of the Mandi (Kharghar, Thane). Given these two specific dimensional (commodity, location) inputs, the price of the commodity at that location can be fetched from a Mandi price database, which is regularly updated. Thus the simplest method to construct a speech enabled *Mandi Bhav Jankari* is to create an IVR system[1] which would in two steps gather the required inputs namely the commodity name and the Mandi location. Such systems have been demonstarted in [4] and [5]. However, such a two step system would make the interaction less friendly and unnatural for the rural farmers [6]. A more natural way of interaction would be to enable a system which understands the natural spoken query of the form "<loc: Gangapur> mandi mei <com: gehun> ka kya bhav hai?" ("What is the price of <com: wheat> at <loc: Gangapur> mandi?" in English).

Figure 1 shows a high level architecture of the system used to build a Natural Language Hindi Speech Interface to Speech enabled *Mandi Bhav Jankari* system. Typically, a farmer initiates the call and is greeted by the system and expects the farmer to speak his query in a natural way. The spoken query is recognized by the ASR (Automatic Speech Recognition) engine with an express intent of recognizing the name of the commodity and the name of the location. The Natural Language Understanding (NLU) block consists of a specifically designed *grammar* which enables the ASR engine to interpret the natural spoken query and extract the two key required fields. Internally the ASR uses a finite state grammar (FSG) which is commonly used in all speech solutions. The location name and the commodity name is then passed to the

[1]called menu based system

TABLE I
GRAMMAR TO ENABLE PROCESSING NATURAL LANGUAGE QUERY (PHRASES IN [] ARE OPTIONAL, <> DENOTES A LIST)

[mujhe] [humko] [aaj]	[yeh] [jaanna hai] [ke] [yeh] [bataiye] [ki]	<loc>	[ki] [mandi] mei [ke] [market] mei	<com>	ka [kitna] bhav ka [kya] bhav	[kitna] [kya]	[chal raha] [hai]	[yeh] [jaanna hai] [yeh] [bataiye]
		<com>	ka bhav ka price	<loc>	[ki] [mandi] mei [ke] [market] mei			

Fetch Price block which fetches the price of the commodity at the given market, from the Agricultural Marketing Information Network (Agmarknet) website [7] in real time. The output (price or other information) is processed and the price is then played back using a text to speech engine (TTS) to the farmer by the prompt engine in Hindi.

In general recognition of a spoken utterance, requires a speech recognition engine. A speech recognition engine consists of two components, namely, (a) the acoustic recognition aided by the pronunciation dictionary and (b) the speech grammar. We used CMU Sphinx [8] speech recognition engine in *Mandi Bhav Jankari*. This engine comes with acoustic models for American English language and a built in pronunciation dictionary of all major English words. Recognizing English with Sphinx is not very difficult especially if the speaker is an American. However *Mandi Bhav Jankari* system needs to recognize (a) Hindi words and names of commodities and Mandi's which are not present in the dictionary provided with the speech engine and (b) speech from Indian farmers in Hindi. Since the emphasis is on building a Natural Language Speech Interface the *Mandi Bhav Jankari* system was designed using the default acoustic models. We achieve the speech recognition by building an appropriate FSG grammar and a phonetic dictionary of Hindi words using the English phonemes. This enables us to use the default acoustic models. The FSG grammar is specifically designed to enable the ASR engine to interpret the natural spoken query and extract the two key required fields of commodity and location. In the following subsection we present this NL speech recognition grammar for market price information interaction.

A. Grammar for Natural Language Speech Recognition

The interface to *Mandi Bhav Jankari* system is a naturally spoken local language interface. Meaning the system can understand any free and naturally spoken speech, as opposed to a conventional menu based system, where the user has to speak words or phrases in isolation as and when prompted by the system. The ability of the system to handle unconstrained speech is through the construction of a specifically designed Finite State Grammar (FSG). A snapshot of the actual grammar used in our system is shown in Table I. This captures the structure (or grammar) of the queries that the system can handle. Also the grammar captured in Table I can handle almost all the commonly phraseable queries that can be constructed by the farmer and posed to the system. Some of the actual queries that the system is able to handle are shown in Figure 2.

Notice that the structured grammar (Table I) also allows the system to handle queries that are incorrect grammatically but make semantic sense like *<loc> mei <com> ka bhav* or *<com> ka bhav <com> mei*. These *transactional grammar* [6] queries are handled due to the appropriate use of optionality for phrases (denoted by [] in Table I). The functional accuracy of the spoken interface is considerably higher than the raw recognition accuracies of the Mandi and commodity fields. The difference in the accuracies is attributable to the structure of the speech recognition grammar, and to the fact that the recognizer always generates a command defined as acceptable by that grammar. The flexibility of the grammar and the ability to paraphrase commands introduce redundancy. So unless the recognizer makes major errors in recognition of Mandi and commodity fields, which it does frequently in the implemented system, the generated command will usually produce the desired action.

- aaj <loc> ki mandi mei <com> ka bhav bataiye
- aaj <loc> ki mandi mei <com> ka bhav kya hai
- aaj <loc> ki mandi mei <com> ka bhav bataiye
- aaj <loc> ke market mei <com> ka bhav kya chal raha hai
- aaj <loc> ke market mei <com> ka bhav kya hai
- aaj <loc> ke market mei <com> ka bhav bataiye
- aaj <loc> mei <com> ka bhav kya chal raha hai
- aaj <loc> mei <com> ka bhav kya hai
- aaj <loc> mei <com> ka bhav bataiye
- mujhe <loc> ki mandi mei <com> ka bhav jaanna hai
- mujhe <loc> ke market mei <com> ka bhav jaanna hai
- mujhe <loc> mei <com> ka bhav jaanna hai
- mujhe <loc> mei <com> ka bhav kya hai yeh jaanna hai
- mujhe yeh jaanna hai ke <loc> mei <com> ka bhav kya hai

Fig. 2. Some queries generated from Table I

IV. *Mandi Bhav Jankari* PERSONALIZATION

In a large scale agro advisory system like mKRISHI [9] the system stores farmer produce related information in a database. Along with this information it is possible to store the list of Mandi's near the actual location of the farmer produce. This information can be exploited by *Mandi Bhav Jankari* system to not only enhance the system in terms of speech recognition accuracy but also to give personalized information to the farmer.

A. Improvement in Recognition Accuracy

The speech recognition accuracy of the free speech queries depends on the number of market locations and number[2] of

[2]perplexity of words to be recognized

Mandi Bhav Jankari: *Namaskar Gyan Singhji. Mandi Bhav Jankari seva mei aapka swagat hain. Is system ke dwara aap aaj ke Mandi Bhav jaan sakte hai. Kripya aapka mandi bhav prashn poochiye.*
Farmer: *Mujhe thane ke mandi mei gajar ka bhav jaanna hai.*
Mandi Bhav Jankari: *Thane mandi mei gajar ka bhav hai 1000 rupaiye prati quintal.*
Mandi Bhav Jankari: *Kya aap nazdiki mandi se apne liye fayidemand bhav jaanna chahte hai ?*
Farmer: *haan...*
Mandi Bhav Jankari: <give the price and mandi at which farmer can get better prices.>

English Translation below

Mandi Bhav Jankari*: Hello Gyan Singh. Welcome to Market Price Information system. You can know market price from this system. Kindly ask your query.*
Farmer: I want to know the price of carrot in thane market.
Mandi Bhav Jankari*: The price of carrot in thane market is 1000 rupees per quintal.*
Mandi Bhav Jankari*: Do you want to know profitable prices at nearby markets ?*
Farmer: yes
Mandi Bhav Jankari*: <give the price and mandi at which farmer can get better prices.>*

Fig. 3. Sample Farmer interaction with *Mandi Bhav Jankari*

commodity names. As the number of locations and commodities increases, there is a higher probability of speech misrecognition. This problem can be eliminated by personalizing *Mandi Bhav Jankari* to individual farmers. In this case the system would be aware of the farmer location which in turn would give information about the specific Mandi's that the farmer might ask for and the possible commodity list could come from the details regarding the crops grown by the farmer [9]. As a result of access to this information the Mandi and commodity names to be recognized by the system is reduced to a very small set, thus reducing the perplexity for speech recognition leading to increase in the speech recognition accuracies.

B. Enhance Farmer Use Experience

The system can use farmer's information to greet him with his name and give a personalized experience. Also the information like the Mandi's preferred by the farmer or the Mandi's located near to the farmer, can be used to proactively provide the farmer with better price information that could be profitable to that farmer. These aspects of personalization have been illustrated in Figure 3.

V. Conclusion

In this paper we addressed the problem of *Mandi Price Information* access for farmers and suggest enabling an easy and a natural speech interface to access *Mandi Price Information* by the rural farmer. We sketched and developed a Hindi natural language understanding speech interface by carefully developing the grammar without the need to take up the gigantic task of developing acoustic models for Hindi speech recognition. The system was built and tested with a grammar containing 116 markets (in 9 districts) and 54 commodities which were available at [7]. The highly structured Natural Language grammar gives a high performance of understanding of the natural spoken queries. The recognition of the market and commodity keywords in the queries has a poor recognition accuracy due to the perplexity of the number of Mandi's and commodities that need to be recognized using the default acoustic models. The problem of recognition of Mandi's and commodities is same as that would be with a menu based system. The Natural Language Hindi interface would enable the farmers in rural India to be aware of the latest mandi prices and hence fetch the best prices for their produce. The personalization of the application can further enhance the farmer experience by increasing the speech recognition accuracy by restricting the perplexity of the number of Mandi's and commodities that need to be recognized. In addition to this technique to improve recognition accuracy, acoustic models built specifically for Hindi language can achieve better recognition of Mandi and commodity names.

References

[1] Agricultural marketing in india. article. [Online]. Available: http://www.world-agriculture.com/agricultural_marketing/agricultural-marketing.php

[2] B. Bowonder, V. Gupta, and A. Singh, "Developing a rural market e-hub. the case study of e-choupal experience of itc." 2002. [Online]. Available: http://planningcommission.nic.in/reports/sereport/ser/stdy_ict/4_e-choupal%20.pdf

[3] D. Zarraga. (2009, Jul.) Indias agricultural revolution goes mobile with mandi bhav. article. [Online]. Available: http://www.mobilebehavior.com/2009/07/08/indias-agricultural-revolution-goes-mobile-with-mandi-bhav/

[4] Mandi information asr system demo. [Online]. Available: http://speech.iiit.ac.in/index.php/demos/asrmandidemo.html

[5] P. Madeline and N. Udhyakumar, "Speech interfaces for equitable access to information technology," *Information Technologies and International Development*, vol. 4, no. 1, pp. 69–86, Fall 2007.

[6] S. K. Kopparapu, "Natural language usable user interface," Keynote: 10th National Conference on Boarder Perspectives on Language, Thinking and Technology. VPM Polytechnic, Thane, Jul. 2010.

[7] Agmarknet homepage. [Online]. Available: http://agmarknet.nic.in/

[8] Cmusphinx toolkit. [Online]. Available: http://cmusphinx.sourceforge.net/wiki/tutorialoverview

[9] A. K. Pande *et al.*, "Mobile phone based agro-advisory system for agricultural challenges in rural india." in *In: IEEE Conference on Technology for Humanitarian Challenges (Aug 2009)*, Bangalore, India, Aug. 2009, pp. 290–294.