

# DIKTEAPI - PUBLICITY DOCUMENT

## INTRODUCTION

---

In computer courses or Introduction to Computers lectures, the first two words mentioned while teaching “Input” section of three main components of computers (Input – Process - Output) are *keyboard* and *mouse*. Both are nearly the only ways used for input or interaction purposes. The invention of *mouse* by Douglas Engelbart in 1968 was a revolution in human-computer interaction concept. First personal computer was developed by IBM in 1981, and they became parts of our daily lives in less than two decades.

Today, using computers is not seen as a skill but a necessity like being literate (we already got used to the *computer literacy* term). Therefore, human computer interaction became more crucial than ever. Human beings communicate by talking, and it is the most natural way of interaction as well. So, why don’t we talk to computers? It would be a revolution that would provide everyone a chance to use computers.

The technology named as Speech Recognition today is a result of the above question. In fact, the question was asked before the invention of computers, and researchers, scientists sharing the dream of “interacting with the machine by speaking, *humanizing* the machine by giving it voice and controlling a machine with voice commands” have been working on the subject for more than half a century.

## HISTORY

---

This dream started to come true with the first electronic voice synthesizer VODER (VOIce CoDER) developed by AT&T Bell Laboratories in 1936. Voder was an electronic speech analyzer and synthesizer equipped with a keyboard and foot pedals. While the first studies were aiming to make the machine speak, they started to focus on human-machine interaction concept with the introduction of modern computers, and *speech recognition* technology was born. In 1969, John Pierce from Bell Laboratories was stating that a several decades had to pass to make automated speech recognition concept come true since it needed artificial intelligence. Pierce was right. Speech Recognition Technology was fed by several

concepts, like Artificial Intelligence, Machine Learning, Mathematics, Statistics, Cognition and Linguistics. Significant developments in Artificial Intelligence and Cognition titles were also backed up with developments in electronics and hardware components (microphones, CPUs, sound cards etc.) as well.

Today – 35 years after Pierce’s statement – computers can recognize human voice, identify specified commands, dictate sentences composed of words in their vocabularies (with hundreds or thousands of words) and partly understand the meanings of speeches. Making them understand daily conversations completely (Natural Language Understanding) is not a dream anymore.

## **APPLICATIONS OF SPEECH RECOGNITION**

---

Speech Recognition Technologies are being widely used from voice call utilities in mobile phones to complex intelligent systems trying to “understand” the speech (natural language processing) using very large vocabularies. Most common applications can be listed as:

- Recognizing word(s) and sentences,
- Recognizing commands
- Data entry
- Hands-free operations
- Custom reporting

## **WHY SPEECH RECOGNITION?**

---

In addition to make the dreams (mentioned in previous sections) come true, the advantages of Speech Recognition Technologies / Systems can be summarized as:

- Using the most common (an natural) way of communication
- Maximizing data entry performance
- Providing eyes and hands-free work environment
- Performance
- Reliability
- Flexibility
- Innovation

Promoting the usage of technology and making it more attractive, contributing to overall labor productivity by helping to increase the effective involvement of the handicapped, providing an innovative and more efficient way to human computer interaction are other economic and social benefits that can be added to the list above.

## **TURKISH AND SPEECH RECOGNITION**

---

Speech recognition systems for Turkish language have been studied academically and via research projects for years, rather than for commercial purposes. There are only a couple of commercial Turkish Speech Recognition systems. There are several reasons for having so few systems, such as the special linguistic features of Turkish and lack of interest on the subject. The linguistic difference of Turkish from other western languages (e.g. English) causes problems in Natural Language Processing (“understanding” what the person says), dictating the words spoken and various other speech based applications. It makes the direct adaptation of Speech Recognition Systems prepared for western languages nearly impossible. As a result, there are few or no products like Tibbi Dikte (Medical Dictation) of Yöndata.

## **APPLICATION DEVELOPMENT AND SPEECH RECOGNITION**

---

One of the major obstacles to the diffusion of Speech Recognition technologies is the monopolistic approaches of the companies developing the technology itself. Such companies should avoid being the only application developer using Speech Recognition. The most efficient method to remove this obstacle is to pack the technology and launch it as a module letting the market integrate it via API, component and other methods.

## **APPLICATION DEVELOPMENT WITH DIKTEAPI**

---

DikteAPI is based on Speech Recognition Engine (KTM<sup>®</sup>) of Yöndata and can be used with common application development tools. It is easy to add speech recognition skills to applications developed in Microsoft Visual Studio .NET (Visual Basic, C++, C#, J#), Sybase PowerBuilder, Borland Delphi and other well-known platforms. DikteAPI box includes *DikteAPI.dll*, *DikteAPI component* for Delphi, *DikteAPI user object* for PowerBuilder and ActiveX Controls (*DikteAPIOCX.ocx*, *DikteAPIDisplayOCX.ocx*) for other development platforms.

You can visit [www.dikte.com.tr](http://www.dikte.com.tr) for product updates, new releases and promotions.

**Hotline:** Please feel free to send e-mail to [dikteapi@dikte.com.tr](mailto:dikteapi@dikte.com.tr) for your questions and problems.