# THESIS EVALUATION FORM

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and qualify as an undergraduate thesis, based on the result of the oral examination taken place on …/.../2011.

…….....

(Advisor)

…….. ……

(Committee Member) (Committee Member)

Prof. Dr. Gülay TOHUMOĞLU

(Chairman)

**ACKNOWLEDGEMENTS**

First of all i am glad to my advisor Assist. Prof. Dr. Gülden KÖKTÜRK encouraging me for this project, then i am glad to the committee members Assist. Prof. Dr. Hakkı Tarkan YALAZAN and Haldun SARNEL for spending time and guiding, and then i am glad to my dear friends İrem ÇAMAŞIRCIOĞLU and Kutlu TACER for helping me in the all steps of my project.

**FATİH KARTAL**

# ABSTRACT

Turkish songs processing with the speech recognition method is the goal of this project. Speech recognition was a curiosity whether it is possible or not to be realized but nowadays it is used in wide range applications from safety systems to funny games. By speech recognition success, it is desired to construct a "karaoke" system. Project included both hardware and software realization of the system but SONY ® did not reply to my e-mails which are about this project so it was impossible to realize hardware part. This project includes software realization.

Karaoke is a form of interactive entertainment or video game in which amateur singers sing along with recorded music using a microphone and public address system. The music is typically a well-known pop song. Lyrics are displayed on a video screen, along with a moving symbol or changing color to guide the singer. Karaoke is a very popular game between young people. The entertainment industry is very important and the karaoke charms the boss with its portion in the market. In this project, a karaoke system is constructed and as different the singer gets a point at the end of the song. The score is given according to the performance of singer and this performance is evaluated by speech recognition algorithm.

There are several developed techniques in digital communication for Automatic Speech Recognition (ASR). These techniques are developed by using different features of human voice. The speech is a continuous time signal and it is necessary to digitalize it with an appropriate sampling rate in order to be processed by a digital processor. There are lots of factors must be taken into account to develop ASR systems such as vocabulary size, if the system speaker dependent or speaker independent, if the ASR is required to recognize isolated words or continuous speech, and so forth.

In this project, the goal is to realize an ASR system which is speaker independent and is required to recognize isolated words. Word numbers is an important factor and the system performance decreases when the vocabulary size increases. And also the sickness, possible stress of the human, and noise interference affect the system performance.

Lots of factors affect performance of the system so it is hard to construct a perfect speech recognition system. It is important to train all factors and analyze the behavior of system against these factors. The algorithm should give the optimum result under all of these undesired conditions.

# ÖZET

Bu projenin amacı Türkçe şarkıların ses tanıma yöntemi ile işlenmesidir. Ses tanıma geçmişte mümkün olup olmadığı merak edilen bir konuydu ancak günümüzde güvenlik sistemlerinden eğlenceli oyunlara kadar birçok uygulamada kullanılan bir yöntem olmuştur. Ses tanımanın gerçekleştirilmesi yardımıyla bir "karaoke" sistemi gerçekleştirilmek istenmiştir. Projenin amacı bu sistemi hem donanımsal hem de yazılımsal olarak gerçekleştirmekti ancak SONY ®, PS3 oyun konsolunda şarkıların nasıl seçildiği konusunda göndermiş olduğum e-postalara cevap vermediği için bu proje sadece yazılımsal olarak gerçekleştirilmiştir.

"Karaoke" amatör şarkıcıların ekranda gördükleri şarkı sözlerini söyleyerek eğlendikleri bir oyundur. Şarkı genelde yaygın olarak bilinen pop müzik türünden seçilir. Şarkı sözleri ekranda gösterilir ve şarkının zamanlamasına uygun olarak kelimeler değişerek şarkıcıya şarkıyı nasıl söylemesi gerektiği konusunda rehberlik eder. Eğlence sektörünün piyasada büyük bir payı olduğu günümüzde, gençler arasında çok yaygın olan bu oyun patronlarında dikkatini çekmektedir. Bu projede "karaoke" oyununa ek olarak oyunun sonunda şarkıcının performansına göre puan verilmektedir. Bu performansın başarısına projede yapmış olduğum ses tanıma sistemi karar vermektedir.

İnsan sesinin farklı özellikleri kullanılarak otomatik ses tanıma için birçok farklı teknik geliştirilmiştir. Ses sürekli bir sinyaldir ve dijital bir işlemci tarafından işlenebilmesi için uygun örnekleme frekansıyla dijitale çevrilmelidir. Otomatik ses tanıma sistemlerinde kelime sayısı, sistemin konuşmacıya bağlı olup olmadığı, ne tür konuşmaların anlaşılmak istendiği gibi dikkat edilmesi gereken birçok özellik vardır.

Bu projede amaç konuşmacıdan bağımsız ve yalıtılmış kelimeleri tanıyan otomatik ses tanıma sistemini gerçekleştirmektir. Kelime sayısı sistem performansını etkileyen en önemli faktörlerden bir tanesidir ve kelime sayısıyla başarı oranı ters orantılıdır. Ayrıca hastalık, insanın stresli olması, mikrofona olan uzaklık ve ortamdaki gürültüde sistemin performansını olumsuz etkileyen faktörlerdendir.

Sistemin performansını etkileyen birçok faktör olduğu için hatasız bir ses tanıma sistemi gerçekleştirmek oldukça zordur. Bu yüzden sistem bütün bu faktörlerle ayrı ayrı denenmeli ve herbirine verdiği tepki incelenerek bunlar giderilmeye çalışılmalıdır. Önemli olan tüm istenmeyen koşullar altında en iyi performansı verebilecek olan sistemi gerçekleştirmektir.

# TABLE OF CONTENTS

**Page**

## List of Figures

## List of Tables