

**Türkçe Dil İşleme için Bürünsel Bilginin
Çıkarılması ve Kullanılması
(Extracting and Using Prosodic Information For
Turkish Spoken Language Processing)**

Proje No: 107E182

Yrd. Doç. Dr. Ümit GÜZ
Yrd. Doç. Dr. Hakan GÜRKAN
Sinan YİĞİT

NİSAN 2010
İSTANBUL

ÖNSÖZ

Bu projede genel olarak, konuşulan dili (Türkçe) anlamada, konuşulan dilin bürünsel/ezgisel (prosodic) ve sözcüksel (lexical) özelliklerinin ortaya çıkarılması ve bu özelliklerin konuşulan dilin bilgisayarla otomatik olarak işlenmesinde kullanılması amaçlanmıştır. Bu daha özel olarak, otomatik konuşma tanıyıcısının (Automatic Speech Recognizer) çıkışına ilişkin cümle bölütleme işlevini içermektedir.

Otomatik konuşma tanıma sistemlerinden çıkan yazılı metnin özellikle noktalama (punctuation), büyük küçük harf farklılıkları ve vurgu, tonlama, perde, durak gibi konuşmaya ilişkin temel bazı parametrelerden yoksun olması veya bu özellikleri kaybetmiş olması, özellikle anlamda farklılıklara yol açmaktadır. Bu çıktının zenginleştirilmesi (enrichment) başka bir deyiş ile bu özelliklerin tekrar geriye kazandırılması, bu metinlerin hem insanlar tarafından okunmasını ve doğru algılanmasını hem de makineler tarafından işlenmesini kolaylaştıracaktır. Bu projedeki amaç, bu zenginleştirme ve geri kazandırma işleminin dilin bürünsel özelliklerinden yararlanarak yapılmasıdır.

Projenin diğer bir sonucu ise biçimbilgisel açıklama (morphological disambiguation) ve kelime anlam açıklama (word sense disambiguation) gibi diğer amaçlar için kullanılacak kelime ve anlambirim (morpheme) düzeyinde bürünsel bir veritabanının elde edilmesidir.

Bu projenin çatısı altında, aynı zamanda bürünsel özelliklerin biçimbilgisel bilgi (morphological information) ile etkileşimi de incelenmektedir.

Cümle bölütlemenin rolü otomatik konuşma tanıma birimi tarafından sağlanan kelime akışındaki cümle sınırlarını, daha ileri düzeyde bölütler bazında çalışan uygulamalar için belirlemektir. Bu, çeşitli ayırma, makine çevirimi, ve soru yanıtlama gibi dil işleme işlevleri için oldukça yararlıdır. Cümle bölütlemeyi iki durumlu sınıflandırma işlevi olarak formüle etmekteyiz. Ardarda gelen iki kelime arasındaki her bir konumda sistem, konumun iki cümle arasında sınır olup olmadığına veya aynı cümleye ilişkin iki komşu kelime olduğuna karar vermektedir.

Yeni yöntem, sözcüksel bilgiyi kullanmak için Saklı Dil Modellerini bürünsel, konuşmacı değişim ve sözdizimsel (syntactic) özellikleri olduğu gibi sözcüksel özelliklerle de entegre olan maksimum entropi ve boosting sınıflandırıcılar ile birleştirmektedir. Boosting tabanlı sınıflandırıcı (kelimeleri ve tüm bürünsel özellikleri giriş olarak kullanan) yalnız başına diğer tüm sınıflandırma yapılarından daha iyi performans sergilemektedir. Saklı dil modelleri ile birleştirildiğinde ise geliştirim daha da iyi olmaktadır.

107E182 Proje No'lu, Türkçe Dil İşleme için Bürünsel Bilginin Çıkarılması ve Kullanılması (Extracting and Using Prosodic Information for Turkish Spoken Language Processing) adlı bu proje, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK), Araştırma Destek Programları Başkanlığı (ARDEB), Elektrik, Elektronik ve Enformatik Araştırma Grubu (EEEAG) tarafından, 3501 Ulusal Genç Araştırmacı Kariyer Geliştirme Programı (Kariyer Programı) çerçevesinde Kariyer Projesi olarak desteklenmiştir.

İÇİNDEKİLER

KAPAK	
ÖNSÖZ	1
İÇİNDEKİLER	2
TABLO LİSTESİ	4
ŞEKİL LİSTESİ	5
ÖZET	6
ABSTRACT	7
1. GİRİŞ	8
1.1. Projenin Tanıtımı ve Amaçları	8
1.2. Önceki Çalışmalar	9
2. GENEL BİLGİLER	12
2.1. Temel Tanımlamalar	12
2.1.1. Bürün (prosody) ve Bürünsel (prosodic) Özellikler	12
2.1.2. Cümle Bölütleme (Sentence Segmentation)	12
2.1.3. Özellik Dizaynı (Feature Design)	14
3. GEREÇ VE YÖNTEM	15
3.1. Cümle Bölütleme için Önerilen Yöntem	15
3.2. Yönteme İlişkin Temel Adımlar	17
3.3. Yöntemin Gerçeklenmesine İlişkin Detaylar	18
3.3.1. Türkçe Konuşma Verilerinin Hazırlanması	18
3.3.2. Türkçe'nin Biçimbilgisel Analizi (Morphological Analysis) ve Biçimbilgisel Özelliklerin (Morphological Features) Çıkarılması	19
3.3.3. Türkçe Konuşma Verilerinin Sözcüksel Özelliklerinin (Lexical features) Çıkarılması	21
3.3.4. Türkçe Konuşma Verilerinin Bürünsel Özelliklerinin (Prosodic features) Çıkarılması	22
3.3.4.1. Algemy (SRI-International'ın bürünsel özellikleri hesaplama yazılımı) ve Temel Bürünsel Özellikler	23
3.3.4.2. Praat ile Bürünsel Özelliklerin Çıkarılması	35
3.3.5. Cümle Bölütlemesi için Üretken (Generative) ve Ayırdedici (Discriminative) Yöntemler	54
3.3.5.1. fHELM (Factored Hidden Event Language Models)	56
3.3.5.2. Ayırdedici Sınıflandırma Modelleri	58
3.3.5.3. Hibrid Üretken ve Ayırdedici Modelleme	59

4. BULGULAR	60
4.1. Veri Setleri	60
4.2. Deęerlendirme Ölçütleri (Evaluation Methods)	61
4.3. Sözcüksel ve Biçimbilgisel Özellikler ile Yapılan Deneyler	61
4.4. Bürünsel Özellikler ile Yapılan Deneyler	63
5. TARTIŞMA VE SONUÇ	64
TEŞEKKÜR	66
REFERANSLAR	66

TABLO LİSTESİ

	<u>Sayfa No</u>
Tablo 1. Bürünsel Özelliklerin (Süre, F0 ve Enerji) Hesaplanmasında Kullanılan Temel Elemanlar.....	35
Tablo 2. Farklı Düzeylerdeki Biçimbilgisel Özelliklerin Kelime-İfade Belirsizlik İstatistikleri: Herbir kelime için biçimbilgisel analiz edici tarafından kelime başına ortalama ayırmsallaştırılma sayıları ve tek bir ayırmsallaştırılmaya sahip kelimelerin (kelime-ifade belirsizliği olmayan kelimeler) oranları.....	61
Tablo 3. Sözcüksel (L), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özelliklerin Boosting, CRF ve HELM/fHELM ile Kullanılması ile Elde Edilen F-Measure Sonuçları.....	62
Tablo 4. Sözcüksel (L), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özelliklerin Boosting, CRF ve HELM/fHELM ile Kullanılması ile Elde Edilen NIST Hata Ölçütü Sonuçları.....	62
Tablo 5. Sözcüksel (L) ve Biçimbilgisel (M) Özellikleri Kullanan Boosting ve CRF in fHELM ile Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları.....	63
Tablo 6. Sadece Bürünsel (P) Özellikleri Kullanan Boosting ile Sözcüksel (L) ve Biçimbilgisel (M) Bilgiyi Kullanan fHELM in Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları.....	64
Tablo 7. Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Elde Edilen F-Measure Sonuçları.....	64
Tablo 8. Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Elde Edilen NIST Hata Oranları.....	64
Tablo 9. fHELM in Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları.....	64

ŞEKİL LİSTESİ

		<u>Sayfa No</u>
Şekil 1.	Cümle bölütleme blok şeması.....	14
Şekil 2.	Herbir kelime sınırı için özellik (duraklar vb.) çıkarım bölgeleri.....	14
Şekil 3.	Bigram model.....	15
Şekil 4.	Algemy Ana Menü.....	23
Şekil 5.	Biçimlendirilmiş pitch (stylized pitch).....	24
Şekil 6.	Biçimlendirilmiş enerji (stylized energy).....	24
Şekil 7.	Pitch ve enerji izleri.....	25
Şekil 8.	Cümle Bölütleme için HELM in kavramsal gösterimi.....	56
Şekil 9.	Okların olasılıkların kestirimi için kullanılan faktörleri, W ve M nin sırası ile kelime ve biçimbilgisel faktörleri gösterdiği FLM grafiği için bir örnek.....	57
Şekil 10.	Okların olasılıkların kestirimi için kullanılan faktörleri, W ve M ve Y nin sırası ile kelime, biçimbilgisel faktörler ve kelime sınırlarını gösterdiği HELM için oluşturulmuş bir FLM örneği.....	58
Şekil 11.	Çeşitli Eğitim Veri Seti Büyüklükleri ile HELM ve fHELM için F-Measure Öğrenme Eğrileri.....	63

ÖZET

Türkçe Dil İşleme için Bürünsel Bilginin Çıkarılması ve Kullanılması

Otomatik konuşma tanıma sistemlerinden çıkan yazılı metnin özellikle noktalama, büyük küçük harf farklılıkları ve vurgu, tonlama, perde, durak gibi konuşmaya ilişkin temel bazı parametrelerden yoksun olması veya bu özellikleri kaybetmiş olması, özellikle anlamda farklılıklara yol açmaktadır. Bu çıktının zenginleştirilmesi başka bir deyiş ile bu özelliklerin tekrar geriye kazandırılması, bu metinlerin hem insanlar tarafından okunmasını ve doğru algılanmasını hem de makineler tarafından işlenmesini kolaylaştıracaktır. Bu projedeki amaç, bu zenginleştirme ve geri kazandırma işleminin dilin bürünsel özelliklerinden yararlanarak yapılmasıdır.

Bu projede, Türkçe dil işleme için sözcüksel özelliklere ilaveten bürünsel bilginin çıkarılması ve kullanılması daha özel olarak ise Türkçe konuşma üzerinde cümle bölütleme için bürünsel özelliklerin kullanılması araştırılmaktadır. Projenin diğer bir sonucu ise biçimbilgisel açıklama ve kelime anlam açıklama gibi diğer amaçlar için kullanılacak kelime ve anlambirim düzeyinde bürünsel bir veritabanının elde edilmesidir.

Türkçe sondan eklemeli veya bitişken bir dildir. Bu nedenle daha ileri analizlerden önce kelimelerin kök biçimleri ve soneklerinin biçimbilgisel olarak incelenmesi yararlı olmaktadır. Bu projenin çatısı altında, aynı zamanda bürünsel özelliklerin biçimbilgisel bilgi ile etkileşimi de incelenmektedir.

Cümle bölütlemenin rolü otomatik konuşma tanıma birimi tarafından sağlanan kelime akışındaki cümle sınırlarını, daha ileri düzeyde bölütler bazında çalışan uygulamalar için belirlemektir. Bu, çeşitli ayırma, makine çevirimi, ve soru yanıtama gibi dil işleme işlevleri için oldukça yararlıdır. Cümle bölütlemeyi iki durumlu sınıflandırma işlevi olarak formüle etmekteyiz. Ardarda gelen iki kelime arasındaki her bir konumda sistem, konumun iki cümle arasında sınır olup olmadığına veya aynı cümleye ilişkin iki komşu kelime olduğuna karar vermektedir. Cümle bölütleme işlemi, Saklı Dil Modelleri (HELM) nin sınıf ayırıcı yöntemler ile birleştirilmesiyle gerçekleştirilmektedir. HELM kelime dizilerini dikkate almakta ve karar ağaçları gibi sınıf ayırıcı yöntemlerin çıkışı, duraksama süreleri gibi bürünsel özellikleri baz almaktadır. Yeni yöntem, sözcüksel bilgiyi kullanmak için Saklı Dil Modellerini bürünsel, konuşmacı değişim ve sözdizimsel özellikleri olduğu gibi sözcüksel özelliklerle de entegre olan maksimum entropi ve boosting sınıflandırıcılar ile birleştirmektedir. Boosting tabanlı sınıflandırıcı yalnız başına diğer tüm sınıflandırma yapılarından daha iyi performans sergilemektedir. Saklı dil modelleri ile birleştirildiğinde ise geliştirim daha da iyi olmaktadır.

Anahtar Kelimeler: Bürünsel Bilgi, Dil İşleme, Cümle Bölütleme, Konu Bölütleme.

ABSTRACT

Extracting and Using Prosodic Information for Turkish Spoken Language Processing

The text which the output of the Automatic Speech Recognition (ASR) system lacks especially punctuation, differences in the capitalization and the parameters related to the speaking such as stress, tone, pitch, pause cause some differences in the meaning. Enrichment of this output or another words to gain this features back to the output will provide either reading and understanding of the humans or processing of the machines easily. The aim of this project is doing this enrichment and the process of gaining back by using the prosodic features of the spoken language.

In this proposal, we would like to examine the extraction and use of prosodic information in addition to lexical features for spoken language processing of Turkish. Specifically, we would like to research the use of prosodic features for sentence segmentation of Turkish speech. Another outcome of the project is to obtain a database of prosodic features at the word and morpheme level, which can be used for other purposes such as morphological disambiguation or word sense disambiguation.

Turkish is an agglutinative language. Thus, the text should be analyzed morphologically in order to determine the root forms and the suffixes of the words before further analysis. In the framework of this project, we also would like to examine the interaction of prosodic features with morphological information.

The role of sentence segmentation is to detect sentence boundaries in the stream of words provided by the ASR module for further downstream processing. This is helpful for various language processing tasks, such as parsing, machine translation and question answering. We formulate sentence segmentation as a binary classification task. For each position between two consecutive words the system must decide if the position marks a boundary between two sentences or if the two neighboring words belong to the same sentence. The sentence segmentation process is established by combining the Hidden Event Language Models (HELMS) with discriminative classification methods. The HELM takes into account the sequence of words and the output discriminative classification methods such as decision tree that is based on prosodic features such as pause durations. The new approach combines the HELMS for exploiting lexical information, with maximum entropy and boosting classifiers that tightly integrate lexical, as well as prosodic, speaker change and syntactic features. The boosting-based classifier alone performs better than all the other classification schemes. When combined with a hidden event language model the improvement is even more pronounced.

Keywords: Prosodic Information, Spoken Language Processing, Sentence Segmentation, Topic Segmentation.

1. GİRİŞ

1.1. Projenin Tanıtımı ve Amaçları:

Bu projede genel olarak, anadilimiz Türkçe'nin bürünsel/ezgisel özelliklerinin (prosodic features) analiz edilerek çıkarılması (feature extraction) ve bu özelliklerin daha ileri düzeyde dil işleme ve uygulamaları konularında araştırmalara temel teşkil eden cümle bölütlemesi alanında kullanılması amaçlanmıştır.

Konuşma işaretlerinin otomatik konuşma sistemleri tarafından basit kelime dizilerine dönüştürülmesi sonucu konuşma ile ilgili bürünsel özellikler (zamanlama ve perde (pitch) yapılarına ilişkin bilgiler, örneğin; duraklar, vurgular, duygunun aksettirilmesi) kaybolmaktadır. Tüm bu özelliklerden yoksun olan metnin gerek insanlar tarafından okunması, gerekse makineler tarafından işlenmesi oldukça zordur. Tüm bu özelliklerin yeniden kazandırılması (enrichment), insanların okudukları metni daha doğal bir biçimde algılaması ve makinelerin daha doğru bir biçimde işlem yapmasını sağlamaktadır.

Özellikle konuşma işaretleri incelendiğinde, içerisinde varolan bürünsel bilginin konuşmanın doğal akışı sırasında oldukça sık kullanıldığı ve konuşmayı tek düzelikten arındırarak zenginleştirdiği görülmektedir. Yine konuşmanın doğal akışı sırasında bu bürünsel bilginin yada özelliklerin, bir cümlenin bitirilip yeni bir cümleye başlandığı, yada bir konudan başka bir konuya geçildiğinde de kendisini gösterdiği görülmektedir. Bu projenin en önemli amaçlarından biri, bürünsel özelliklerin yeniden kazandırılmasının yanısıra bu özelliklerin yukarıda belirtilen cümle bölütlemesi için de kullanılmasını sağlamaktır.

Bu projenin en önemli özelliklerinden biri metin tabanlı çalışmalardan farklı olarak konuşma işaretinin bizzat kendisinin kullanılması, başka bir deyiş ile konuşma tabanlı olmasıdır. Bu nedenle çok fazla miktarda ve içerik olarak çok çeşitli alanlardan (domain) konuşma verileri elde edilerek incelenmiş ve bu konuşma verileri kullanılarak Türkçe'nin bürünsel özellikleri çıkarılmıştır. Daha sonra Türkçe'nin bu bürünsel özellikleri, özellikle bilginin veri tabanlarından otomatik olarak taranması, bilgi çıkarımı (information extraction) bilginin geri kazanımı (information retrieval), metin özetleme (text summarization) ve makine ile dilden dile çeviri (machine translation) gibi uygulamaların temel adımı olan cümle bölütleme için kullanılmaktadır.

Böylece gerek insanların yukarıda belirtilen uygulamaları kendi başlarına yapabilmek için gerekli olan emek, zaman ve maliyetleri azaltmak, gerekse makineler tarafından yapılacak daha ileri işlemler ve uygulamalar için çok daha az ve sınıflandırılmış verileri kullanmaları mümkün olabilecektir. Bilgi, bilgiye erişim, bilginin kullanımı ve iletişim alanındaki hızlı gelişim, özellikle iletişimden gelen doğası ve iletişim uygulamaları ile olan sıkı ilişkisi itibari ile konuşma teknolojilerine ve uygulamalarına yönelik araştırmaların hızla gelişimine yolaçmaktadır. Bu alanlarda özellikle İngilizce başta olmak üzere farklı dillerde uygulamaların geliştirilmesi, gerek yukarıda belirtilen nedenlerden dolayı bu alanlarda yapılması gereken araştırmaların bir gereksinim yada zorunluluk haline gelmesi, gerekse uluslararası alanda yaşanan rekabet ve söz sahibi olma nedenlerinden dolayı oldukça önem taşımaktadır. Ayrıca, bu ve benzer uygulamaların Türkçe için yapılması ise ayrı bir önem arz etmektedir.

Yukarıda belirtildiği üzere bu proje genel kapsamı itibari ile, Türkçe'nin kendine özgü bürünsel özelliklerinin çıkarılması ve bu özelliklerin daha ileri dil işleme uygulamalarında yapılması gereken ilk adım olan cümle bölütleme uygulamalarında kullanılmasını kapsamaktadır.

Projenin ilk aşamasında Türkçe'nin bürünsel açıdan yapılacak analizleri sonucu bürünsel özelliklerin çıkarılmasının yanısıra Türkçe'ye ilişkin başka analizler (sözcüksel (lexical), sözdizimsel (syntactic), biçimbilgisel (morphological)) de yapılmış ve bulunan özellikler ile bürünsel özelliklerin ilişkisi irdelenerek özellikle sınıflandırma aşamalarında cümle bölütleme uygulamalarında performansın artırılması sağlanmıştır. Böylece özellikle Türkçe için geliştirilmiş ve daha ileri dil işleme uygulamalarında kullanılabilir bir altyapı hazırlanmış olmaktadır. Ayrıca, pek çok farklı kaynaktan elde edilen veriler ile Türkçe konuşmaya ilişkin veritabanları oluşturulacak, uygulanacak yöntemlerin gerektirdiği analizlerin yapılması sonucunda bürünsel ve diğer özellikleri ve bunların birbirleri ile ilişkilerini içeren veritabanları oluşturulmuştur. Yöntemlerin tümüne ilişkin çeşitli başarımlar ölçümleri baz alınarak değerlendirmeler yapılmıştır. Böylece özellikle Türkçe konuşma verileri için cümle bölütleme uygulamalarına yönelik olarak başarımda etken olan bürünsel özelliklerin ortaya çıkarılması sağlanmıştır.

Bu projede uygulanacak yöntemler, Türkçe dil ailesinin içerisinde bulunduğu ve yapı olarak Türkçe ile büyük benzerlikler gösteren diğer diller üzerinde yapılacak çalışmalara da referans olabilecektir.

1.2. Önceki Çalışmalar

Türkçe'nin biçimbilgisel ve dilbilimsel yapısına ilişkin belli başlı çalışmalar mevcut olmakla birlikte, özel olarak Türkçe'nin bürünsel özelliklerinin çıkarılması ve kullanılması konusunda yapılan çalışmalar ise oldukça sınırlıdır.

Türkçe'nin biçimbilgisel yapısı ile ilgili olarak Türkçe kelime yapılarının iki tabanlı biçimbilgisel açıklanması, Türkçe metinler için derlem etiketleyicileri, Türkçe'de cümlelerin ayrılmasında kullanılan yapılar gibi metin tabanlı çalışmalar gerçekleştirilmiştir (Oflazer ve diğerleri, 1994; Oflazer ve diğerleri, 1995).

Türkçe ile ilgili dilbilimsel olarak bazı bürünsel özelliklerin incelendiği yayınlar bulunmaktadır (Saylı ve diğerleri, 2003; Oskay ve diğerleri, 2001, Levi, 2002; Levi, 2001; Levi, 2005).

Literatürde, özellikle dünya üzerinde birçok insanın konuştuğu ortak bir dil olması nedeniyle İngilizce ve yine sayı olarak çok fazla insanın konuştuğu Arapça ve Çince dillerinde cümle bölütleme ve konu bölütleme ile bürünsel özelliklerin kullanıldığı bazı çalışmalar mevcuttur.

Konuşma akışının (speech stream) kendisinin ele alındığı cümle bölütleme konusundaki ilk çalışmalar konuşma işaretinin kendisinde varolan ancak yazıya uyarlanamayan özellikle vurgu, durak gibi özelliklerin kullanılmasına yöneliktir. Hirschberg ve Grosz vurgu ve söylem (discourse) yapısı arasındaki ilişki konusunda çalışmış ve vurgu özellikleri ile etiketlenen söylem yapıları arasında bir bağ olduğunu ortaya koymuştur (Hirschberg ve diğerleri, 1992).

Hirschberg ve Nakatani, söylemin, metnin ve konuşmanın birlikte kullanılması ile yapılan bölütlemesinin sadece metin kullanılarak yapılan bölütlemeden daha iyi olduğunu göstermişlerdir (Hirschberg ve diğerleri, 1996).

Özellikle konuşma işaretlerinde cümle sınırlarının tespiti (ve benzer biçimde noktalama işaretleme) otomatik konuşma tanıyıcısı çıkışının zenginleştirilmesi olarak ele alınmış ve bu konuda çalışmalar yapılmıştır (Shriberg ve diğerleri, 2000; Gotoh ve diğerleri, 2000; Liu, Shriberg ve diğerleri, 2005; Roark ve diğerleri, 2006).

Önceki yaklaşımlarda bu işlev için hem text yada metin hem de bürünsel özelliklerin yararlandığı farklı sınıflandırıcılar (örneğin, saklı Markov modelleri (HMM), maximum entropy gibi) kullanılarak değerlendirmeler yapılmıştır. Bu konuda yapılan diğer bir proje ise Amerikan Savunma Bakanlığı'na

bağlı çalışan İleri Savunma Araştırmaları Projeleri Ajansı (DARPA-Defense Advanced Research Projects Agency) EARS (Effective, Affordable, Reusable Speech-to-Text) programı için geliştirilmiştir. Bu projede, konuşmanın, otomatik bir biçimde üretilen cümle sınırları, konuşmacının konuşmasına devam ederken durakladığı ve bir sonraki kelime veya kalıbı düşünürken arada çıkardığı “uhh, ah” gibi sesleri ve doldurulmuş kelimeler gibi yapısal bilgileri de içeren zengin bir biçimde yazıya dönüştürülmesi konusunda özel çabalar harcanmıştır.

Diğer bir çalışmada hem telefon konuşma verileri üzerinde hem de radyo haber yayınlarına ilişkin konuşma verileri üzerinde farklı modelleme yaklaşımları (saklı Markov modelleri, maximum entropy ve conditional random fields) ve çeşitli bürünsel ve metinsel özellikler değerlendirilmiştir (Liu, Shriberg ve diğerleri, 2005). Bir yeniden sıralama (reranking) tekniği (Roark ve diğerleri, 2006), referans olarak kullandığı (Liu, Shriberg ve diğerleri, 2005) tarafından verilen yöntemden daha yüksek bir cümle sınırı belirleme performansı göstermiştir.

Cümle bölütleme yöntemleri telefon konuşma verilerine uygulandıktan ve olumlu sonuçlar alındıktan sonra benzer yöntemler bu defa içerisinde birçok kişinin yer aldığı toplantı konuşma derlemleri için de uygulanmıştır (Ang ve diğerleri, 2005; Zimmermann ve diğerleri, 2005).

Cümle sınırlarının bulunduğu ve çoğunlukla metin tabanlı benzer yaklaşımlar saklı Markov modellerinin kullanılması ile Çekçe (Kolar ve diğerleri, 2004) ye ve maximum entropy sınıflandırıcıları kullanarak da Çince’ye uyarlanmışlardır (Zong ve diğerleri, 2003; Liu ve diğerleri, 2003).

Son zamanlarda yapılan çalışmalarda support vector machine, boosting ve maximum entropy classifiers ile entropy tabanlı sınıflandırıcıların birleştirilmeleri ve bunların çıkışlarının da Saklı dil modelleri ile bütünleştirilmeleri ile İngilizce, Arapça ve Çince için oldukça belirgin gelişimler sağlanmıştır.

Konu bölütleme ile ilgili önceki çalışmalar, genellikle sözcüksel özelliklerin kullanımında yoğunlaşmıştır. Buna örnek olarak kelime benzerliklerinin kullanıldığı çalışmalar (Kozima, 1993), ipucu kalıplarının kullanıldığı çalışmalar (Passonneau, 1997), sözcüksel pencerelerin kosinüs benzerliklerinin kullanıldığı çalışmalar (Hearst, 1997) ve adaptif dil modellemenin kullanıldığı çalışmalar (Beeferman ve diğerleri, 1999) verilebilir.

Radyo haber yayınlarına ilişkin konuşma verilerinin konu bölütlemesi alanında başka bir deyiş ile akustik özelliklerin kullanıldığı konu bölütlemesi konusunda çalışmalar yapılmıştır (Shriberg ve diğerleri, 2000; Tur ve diğerleri, 2001).

Konuşmanın konu ünitelerine otomatik olarak bölütlenmesi için hem bürünsel hem de sözcüksel özelliklerin kullanıldığı olasılıksal bir model sunulmuştur. Bu yaklaşım, saklı Markov modelleri, istatistiksel dil modellerini ve bürün-tabanlı karar ağaç yapılarını birleştiren bir yaklaşımdır. Sözcüksel bilgi, konuşma tanıyıcısından, bürünsel özellikler ise otomatik bir biçimde konuşma dalga biçimlerinden elde edilmektedir. Bu çalışmanın sonucunda, bürünsel modelin tek başına kelime tabanlı bölütlemeye üstün olduğu ve hem bürünsel hem de kelime tabanlı bilgi kaynaklarının birleştirilmesi ile çok daha iyi bir model oluşturulabileceği gösterilmiştir (Tur ve diğerleri, 2001).

Bürünsel özelliklerin kelime bilgisi ile birleştirilerek bilgi çıkarımı (Hakkani-Tür ve diğerleri, 2005) ve konu bölütlemesinde kullanıldığı çalışmalar (Stolcke ve diğerleri, 1999) da mevcuttur.

(Guz ve diğerleri, 2007) Co-training ile bürünsel ve sözcüksel bilgiyi kullanarak cümle bölütleme alanında uygulamışlardır.

Bu projede önerilen yöntemler, Türkçe konuşma verilerine ilişkin bürünsel özelliklerin ortaya çıkarılmasını ve cümle bölütleme uygulamalarında bürünsel özelliklerin kullanılmasını sağlamaya dönüktür. Bununla birlikte yapılan analizlerden elde edilen diğer özellikler (biçimbilgisel, sözcüksel, sözdizimsel vb.) ile bu özelliklerin bürünsel özellikler ile ilişkisi de incelenmiştir. Ayrıca Türkçe dilinin yapısından kaynaklanan nedenlerden dolayı özellikle karar ağaç yapıları yerine boosting, support vector machines, conditional random fields gibi farklı sınıflandırıcılar ile çalışılmıştır.

Bu projede önerilen yöntem ile Türkçe konuşma işaretlerinin pekçok yönden analiz edilmesi sağlanmış ve önemli bulgulara ulaşılmıştır.

Özellikle metin tabanlı yaklaşımlardan farklı olarak Türkçe konuşma işaretlerinin (örneğin radyo haber yayınlarına ilişkin konuşma verilerinin) kendisinin kullanılması ile, sorgulanabilen uygulamaların ve sistemlerin geliştirilmesi (konuyu özetleme veya belirli özelliklere göre sınıflandırma) cümle bölütleme adımının gerçekleştirilmesi ile yapılmaktadır. Bu yapılırken diğer yöntemlerden farklı olarak dilin bürünsel özelliklerinin kullanılması sağlanmaktadır.

Literatürde özellikle cümle ve konu bölütleme alanında yapılan araştırmalarda çoğunlukla dile ilişkin sözcüksel bilginin kullanılmasına yönelik yöntemlerin geliştirildiği görülmektedir. Bu projede ise cümle bölütleme için öncelikle bürünsel bilginin çıkarılması ve kullanılması sağlanmıştır.

Özellikle Türkçe konuşma işaretlerinin kullanılarak bürünsel özelliklerin çıkarılması ve bu özelliklerin cümle bölütleme gibi dil işleme alanında pek çok uygulamanın ilk adımına uygulanmış olması projenin diğer bir özgün yönünü oluşturmaktadır.

Ayrıca belirtilen Türkçe konuşma veya audio işaretlerin oldukça başarılı bir dilden bağımsız otomatik konuşma tanıyıcısından geçirilmesi, (SRI Decipher) önerilen yöntemin performansını ve güvenilirliğini arttıran bir özelliktir.

Bu işlemler SRI Decipher konuşma tanıyıcısı üzerinde gerçekleştirildiği gibi bağımsız açık kaynak kodlu yazılımlar (Hidden Markov Toolkit (HTK)) üzerinde de gerçekleştirildiğinden gerek telif hakkı, gerekse daha sonradan yapılacak çalışmalarda oluşabilecek yazılım bağımlılığı sözkonusu olmamaktadır. Ayrıca belirtilen açık kaynak kodlu yazılımların kullanılması, proje araştırmacıları ve başka araştırmacılar tarafından daha sonradan sistem üzerinde yapılabilecek geliştirmeler ve farklı ileri uygulamalar (konu bölütleme vb.) için de uygun ve ortak kullanıma açık bir taban oluşturmaktadır.

Bu projenin bir kısmında özellikle Türkçe'nin biçimbilgisel yapısının analizi konusunda daha önce gerçekleştirilen araçlar da kullanılmaktadır.

Bu projede, Türkçe'nin yapısından kaynaklanan nedenler ile ilgili olarak büyük miktarlarda özellik setleri ile çalışıldığından karar ağaç yapıları (decision trees) yerine boosting ve support vector machine sınıflandırıcılar tercih edilmiştir. Sınıflandırma işleminde özellik olarak kelime n-gramlar da kullanılmıştır.

Ayrıca farklı bir yaklaşım olarak da cümle bölütleme işlemlerinde herbir cümle veya konu sınırı için sınır olma ve sınır olmama durumlarına göre ayrı ayrı sınıflandırmanın yapılması ve bu yapılırken de bağlamsal (contextual) özelliklerin kullanılmasıdır. Bu amaçla boosting ve support vector machine sınıflandırıcıları kullanılmıştır.

Türkçe için gerçekleştirilen bu proje ile, bürünsel özelliklerin çıkarılması ve kullanılması sonucu elde edilen cümle bölütleme yaklaşımlarının daha sonra tarafımızdan ve diğer araştırmacılar tarafından yapılacak daha ileri dil işleme uygulamaları için bir temel teşkil etmesi amaçlanmaktadır.

2. GENEL BİLGİLER

2.1. Temel Tanımlamalar

2.1.1. Bürün (prosody) ve bürünsel (prosodic) özellikler

Bürün genel olarak bir dilin vurgu, perde, tonlama, duraksama, vb. özelliklerini içerir. Bu özellikler, anlam farkı yaratmaktan çok, niyete veya işleve işaret eder. Soru sorulurken son hecelerin yüksek bir tonlamayla ifade edilmesi, keskin bir vurgunun kızgınlık, yüksek bir tınının kaygı ifade edebilmesi buna birkaç örnektir.

Konuşma işaretlerinin otomatik konuşma sistemleri tarafından basit kelime dizilerine dönüştürülmesi sonucu konuşma ile ilgili zamanlama ve perde yapılarına ilişkin bilgiler veya özellikler kaybolur. Bu yapılar (ve bu yapılarla bağlı kelimelerden bağımsız diğer özellikler) bürün olarak adlandırılır. Tüm dillerde bürün, yapısal, anlamsal ve işlevsel bilgiyi taşımaktadır. Bürünsel özellikler veya ipuçları (cues) dilin söylem yapısı ile ilişkili olduklarından özellikle çeşitli bilgi çıkarılması (information extraction) işlevlerinde önemli bir rol oynamaktadır. Dilbiliminde, okuma ve anında yapılan monologların analizinde ve buna bağlı işlevlerde cümleler ve paragraflar gibi bilgi birimlerinin bürünsel olarak işaretlenmedikleri gösterilmiştir. İngilizce’de ve ona benzer veya bağlı dillerde bürünsel belirteçler olarak duraklar, perde aralığındaki ve genliğindeki değişim, genel perde eğimi, ezgi ve sınır ton dağılımı ve konuşma oranındaki değişim gibi özellikler sayılabilir. Örneğin cümle veya paragraf sınırları ile konu sınırlarının her ikisi de bazı uzun durakların, tonda meydana gelen bir düşmenin ve perdenin yeniden düzenlenmesinin olduğu durumlarda işaretlenir. Bunlar sınır belirleme veya işaretleme için önemli göstergelerdir. Ayrıca, bürünsel özellikler doğası gereği göreceli bir biçimde kelimenin kimliğinden etkilenmediğinden, başka bir ifade ile kelimelerden bağımsız olduklarından otomatik konuşma sistemi çıkışı baz alan sözcüksel bilgi çıkarım yöntemlerinin gürbüzlüğü (robustness) geliştirmektedir.

2.1.2. Cümle Bölütleme (Sentence Segmentation)

Aşağıda yeralan örnekte de görüleceği üzere, insanlar tarafından gerçekleştirilen konuşma sinyalinden metine (text) çevirme işlemi-çevriyazı (transcription) sonunda elde edilen sonuç ile otomatik konuşma sisteminin/konuşma tanıyıcısının (ASR-Automatic Speech Recognition system/speech recognizer) girişinden uygulanan konuşma sinyalinin metine dönüştürülmesi sonucunda elde edilen çıkış arasında bazı farklılıklar oluşmaktadır. Bilindiği üzere günümüzde kullanılmakta olan otomatik konuşma sistemlerinin çıkışından elde edilen kelime dizileri biçimindeki metinler, noktalama işaretleri, büyük küçük harf ayrımı, paragraf başlangıcı ve bitişi, konuşmacı değişikliğini gösterir belirteçler gibi pek çok bilgiden eksik bir biçimde elde edilirler. Bütün bu eksiklikler aynı zamanda cümlelerin bölütlenmemiş olduğu gerçeğine de işaret eder.

İnsan çevriyazımı (human transcription):

but uh i'm i i i think that you know i mean we always uh i mean i've i've had a a lot of good experiences with uh with many many people especially where they've had uh extended family and i and an- i i kind of see that that you know perhaps you know we may need to like get close to the family environment and and get down to the values of you know i mean uh it's money seems to be too big of an issue wi- with with with with with what's going on today

Otomatik Konuşma Tanıma sistemi çıkışı (ASR output/machine transcription):

but um that that {uh i'm i i} i think that you know we {i mean} we always uh i mean i've i've had it there {a} a lot of good experiences with the {uh} with many many people especially with have {where they've} had extended family night and i and {an- i} i kind of see that that you know perhaps you know we may need to like you're {get} close to the family environment and in {and} get down to the values of you know i mean no and {uh it's} money seems to be too big of an issue we would {wi- with with with} with with really was we would what's going on today

Çoğu zaman insanlar tarafından el ile yapılan çevriyazıların işlenmesinin ve okunmasının dahi zor olabileceği durumlar düşünüldüğünde makineler tarafından gerçekleştirilen bu eksikliğin öneminin oldukça büyük olduğu görülecektir. Özellikle insanlar tarafından konuşma işaretlerinin aslına uygun ve yukarıda belirtilen tüm eksiklikleri karşılayacak bir biçimde metine çevrilmesi, emek yoğun, dikkat gerektiren ve oldukça pahalı bir işlemdir. Bu nedenle bu işlevlerin makineler tarafından etkin bir biçimde yapılmasına yönelik çalışmalar hız kazanmıştır. Bu bağlamda, özellikle cümle bölütleme konusu bu alanda yapılan pek çok ileri araştırmaya (makine ile dilden dile çeviri, bilgi çıkarımı, konu belirleme ve özetleme vb.) temel oluşturması veya bu araştırmaların ilk adımını oluşturması açısından oldukça önemli bir yere sahiptir.

Konuşma işaretlerinden, cümle bölütlemesi başka bir ifade ile konuşmanın cümlelere ayrılması, standart otomatik konuşma tanıyıcılarından çıkan yada bu tanıyıcılar tarafından üretilen ve üzerinde bir işlem yapılmamış/işlenmemiş kelime akışının (word stream) zenginleştirilmesi (enrichment) amacıyla yapılan bir işlemdir. Cümle bölütleme, cümle sınırlarının önceden belirlenmiş olmasını gerektiren yada varolduğunu varsayan çözümleme (parsing), makine çevirimi (machine translation), bilgi çıkarımı (information extraction) gibi konuşma işleme ile ilgili daha ileri işleme adımlarında özel bir öneme sahiptir. Burada adı belirtilen tüm ileri işleme adımları ve daha pek çok uygulamada cümle bölütlemesi ilk yapılacak işlemdir.

Cümle bölütlemesi için farklı bilgi kaynakları dikkate alınmaktadır. En önemli bilgi kaynakları, otomatik konuşma tanıma biriminden gelen kelimeler dizisi ve komşu kelimeler arasındaki duraksama süreleridir.

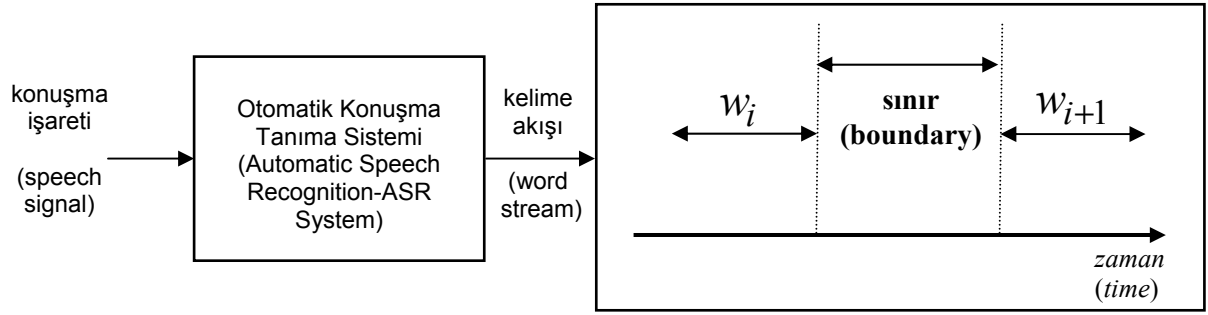
Cümle bölütleme problemi, verilen bir özellik seti üzerinden sınıfın sonsal olasılık (posterior probability) değerinin kestirimine göre herbir kelime sınırının bir sınıf etiketi (cümle sınırı olan veya cümle sınırı olmayan) ile ilişkilendirildiği bir ikili sınıflandırma problemi olarak ele alınabilir.

Cümle bölütleme için, sözcüksel ve duraksal özelliklerin birleşimi yada sözcüksel, duraksal veya diğer bürünsel özellikler kullanılabilir. Sınıflandırma için, boosting, maximum entropy ve Vektör destek makineleri (SVM-support vector machine) sınıflandırıcıları gibi farklı sınıflandırıcıları kullanmak mümkündür. Burada ayrıca belirtmek gerekmektedir ki, cümle bölütlemesi konu bölütlemesi için gerekli olan ilk adımdır.

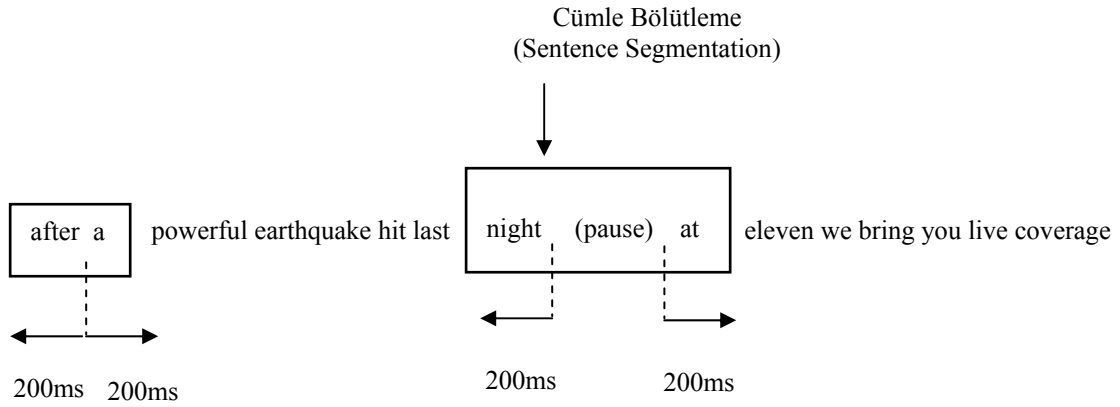
Yukarıda değinildiği üzere, cümle bölütlemesi “cümle sınırı” (s) ve “cümle sınırı değil” (n) kararlarının sınıf olarak atandığı bir ikili sınır sınıflandırma problemi olarak ele alınabilir. Verilen bir kelime akışı veya dizisi ($\{w_1, \dots, w_N\}$) için amaç, sınırlar ($\{e_1, \dots, e_N\}, e \in \{s, n\}$) için sınıfların kestirimidir. Burada ($s_i, i = 1, \dots, N$), w_i ve w_{i+1} arasındaki sınırdır. Genellikle bu, sonsal olasılığın $P(s_i = k | o_i)$, $k \in \{s, n\}$ kestirimi için sınıflandırıcının eğitilmesi ile yapılır. Burada o_i , ler s_i kelime sınırı için özellik gözlemleridir (feature observations).

İdeal durumda, sınıflandırıcının kararı en yüksek olasılık değerine $P(s_i = k | o_i)$ sahip sınıftır. Bununla birlikte, cümle bölütleme işlevinde cümle sınırı için olasılık $P(s_i = s | o_i)$, bir eşik değeri (threshold) ile karşılaştırılır. Bu eşik değerinin üzerindeki olasılık değerlerinde karar cümle sınırı olduğu yönünde, aksi durumda ise cümle sınırı olmadığı yönünde verilmektedir.

Daha sonradan da görüleceği üzere, farklı sınıflandırıcılar ve farklı değerlendirme ölçütleri için en iyi (optimal) eşik değeri farklıdır.



Şekil 1. Cümle bölütleme blok şeması



Şekil 2. Herbir kelime sınırı için özellik (duraklar vb.) çıkarım bölgeleri

2.1.3. Özellik Dizaynı (Feature Design): Bürünsel ve sözcüksel özellikler sınıflandırıcının cümle sınırlarını temsil etmek için kullandığı özelliklerdir. Sözcüksel özellikler ilgilenilen sınırı izleyen kelime ile iki önceki kelimenin n-gram larından oluşmaktadır. Bürünsel özellikler ilgilenilen kelime sınırında yer alan iki kelime arasındaki durak süreleri ile konuşmacının sesine ilişkin pitch ve enerji ile ilgili değişik ölçümleri içermektedir. Dizayn edilen bu özellikler öyle hesaplanmaktadır ki ilgilenilen kelime sınırından önceki ve sonraki pitch ve enerji değerlerinin farkını veya karşılaştırmasını yapacak biçimde değer alırlar. Ölçülen bu değerler ilgilenilen kelime sınırındaki kelime bazında olduğu kadar çerçeve bazında da hesaplanmaktadır. Bu ölçümler bu aralıklardaki minimum, maksimum ve ortalama değerleri kapsamaktadır. Bazı özellikler aynı zamanda konuşmacıya göre normalize de edilmektedir.

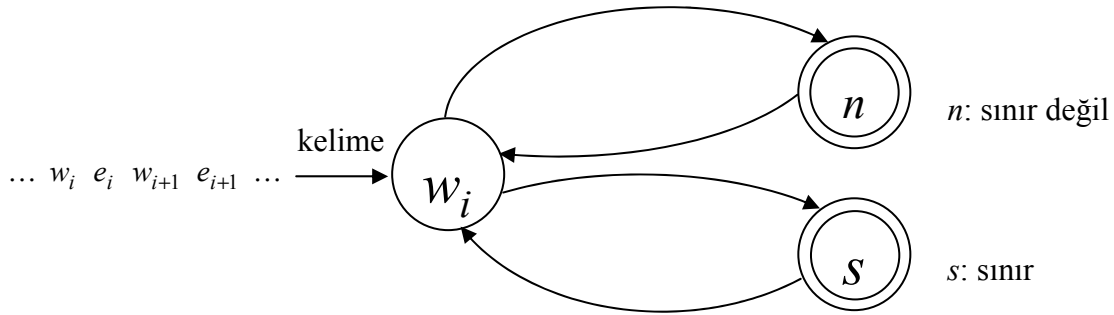
3. GEREÇ VE YÖNTEM

3.1. Cümle Bölütleme için Önerilen Yöntem

Otomatik konuşma tanıma sisteminin çıkışı, herhangi bir uygun noktalama ve harflerin büyük küçük olarak ayrıldığı, genellikle ham yada işlenmemiş olarak ifade edilebilecek kelime dizisinden oluşmaktadır. Cümle bölütlemenin bu noktadaki amacı text yada metindeki cümle sınırlarının tespit edilmesidir. Bu sınırların belirlenmesinin yararı, insanlar tarafından okunabilirliğin artırılması olduğu kadar makineler tarafından da kabul edilebilir büyüklük birimlerinde olan verilerin işlenebilmesini sağlamaktır. Diğer bir yararı ise, dilden dile yapılan otomatik makine çevriminde yeniden sıralama işlemini geliştirmesidir. Bu anlamda bakıldığında daha ileri düzeyde yapılacak tüm işlemlerde cümle bölütleme işleminin ilk adım olduğu söylenebilir.

Bu projede cümle bölütleme problemini bir ikili sınır sınıflandırma problemi olarak ele alınmıştır. Bu nedenle, verilen bir kelime dizisi W ve ona bağlı özelliklerin kullanılması ile her bir kelime sınırının türü bulunmaya çalışılmıştır. Cümle bölütleme durumunda sınır türleri cümle sınırı olma ve cümle sınırı olmama biçiminde olmak üzere iki türdür. Bizim temel yaklaşımımız, herbir sınır için ayrı ayrı sınıflandırmanın yapılması ve bu yapılırken de bağlamsal (contextual) özelliklerin kullanılmasıdır. Bu amaçla boosting ve support vector machine sınıflandırıcıları kullanılmıştır.

Bir diğer yaklaşım, dizi sınıflandırma yöntemi olup saklı dil modellerini (HELMS-Hidden Event Language Models) kullanmaktır. Bu nedenle, kelime dizisinin kelimeler (w_i) ve sınırlardan (e_i) oluşan bir dizi ($\dots, w_{i-1}, e_{i-1}, w_i, e_i, w_{i+1}, e_{i+1}, \dots$) olması gerektiği anlaşılmalıdır. Sınır türleri, saklı olaylardır (hidden events) (E_i) ve her bir sınır ile ilişkili özellikler F_i dir. Burada sınırların Şekil 3. deki bigram durumunda olduğu gibi bir Markov modeli (yada sonlu durum makinesi (finite-state machine)) tarafından üretildiği veya tanındığı varsayılmaktadır. Bir durum kelimeleri üretir ve herbir saklı olay için bir durumumuz olup bu durumda ikinci durum cümle sınırı olma ve cümle sınırı olmama ya ilişkindir. Geçiş olasılıkları eğitime datasının saklı olay formatındaki kelimeye dönüştürülmesi ile kestirilirler. Daha sonra verilen diğer kelime dizisi için enyüksek birleşik olasılık (maximum joint probability) değerine sahip kelimeler ve sınır türlerinden saklı olay dizisi bulunmaktadır.



Cümle bölütlemenin amacı verilen bir kelime dizisi ve bürünsel özellik dizisi için enyüksek olasılığa sahip sınır türleri dizisini (E) bulmaktır.

$$\hat{E} = \arg \max_E P(E | F, W) \quad (1)$$

$$= \arg \max_E \frac{P(E, F, W)}{P(F, W)} \quad (2)$$

$$= \arg \max_E P(E, F, W) \quad (3)$$

Yukarıdaki denklemde argmax dikkate alındığında farklı E değerleri için $P(F | W)$ sabit olup, gözardı edilebilir.

$$= \arg \max_E P(E, W) P(F | E, W) \quad (4)$$

Yukarıdaki denklemde $P(E, W)$ olasılıkları eğitim metninin bir kelime olay dizisine $w_1, e_1, w_2, e_2, \dots$ dönüştürülmesi ve n-gram dil modeli olasılıklarının hesaplanması ile belirlenebilir.

$P(F | E, W)$, aşağıda açıklandığı gibi $P_B = (E_i | F_i, W_i)$ olasılık değerlerini üreten ayırdedici sınıflandırıcılar kullanılarak hesaplanabilmektedir.

$$P(F_i | E_i, W_i) = \frac{P_B(E_i | F_i, W_i) \cdot P(F_i | W_i)}{P(E_i | W_i)} \quad (5)$$

Bu denklemde $P(F_i | W_i)$ farklı E dizileri için sabit olup gözardı edilebilir.

$$P(F_i | W_i) = \text{Sabit} \quad (6)$$

Yukarıdaki ifadelerde belirtildiği gibi $P(E_i | W_i)$ değerlerine $P(E_i)$ ile yaklaşılabilir. Bu projede aynı zamanda bu yaklaşımın elimine edilmesi de ele alınmıştır. Bu nedenle, ayırdedici sınıflandırıcı çıkışlarındaki sonsal olasılık değerleri normalize edilmiş olabilirliğe dönüştürülebilir ve Şekil 3 de gösterildiği gibi s ve n durumları için durum geçiş olasılıkları olarak kullanılabilirler.

Bu formülasyon (Tur ve diğerleri, 2001) çalışması ile benzerlikler göstermekle birlikte adı geçen çalışmada durum geçiş olasılıklarının hesaplanmasında karar ağaç (decision trees) yapıları kullanılmıştır. Burada yapılan çalışmada ise büyük miktarlarda özellik setleri ile çalışıldığından ayırdedici boosting ve support vector machine sınıflandırıcılar kullanılmış ve diğer faydaları yanında sınıflandırma boyunca özellik olarak kelime n-gramlar da kullanılmıştır.

Bu yaklaşım bürünsel özellikler gibi diğer özellik setleri F ile kolayca birleştirebilmekte ve verilen kelimeler ve özellikler için enyüksek olasılığı veren olayların olasılıkları kestirilebilmektedir.

Daha sonra verilen olaylar ve kelimeler için özelliklerin olasılıklarına ihtiyaç duyulacaktır $P(E_i | F_i, W_i)$. Bu olasılıklar boosting veya benzeri sınıflandırıcılar tarafından elde edilmişlerdir.

Sınıflandırma işlemi boyunca, sözcüksel, bürünsel ve sözdizimsel özellikler ile bunların n-gram özelliklerde kombinasyonları kullanılmıştır. Sözcüksel özellikler kelime n-gramlardır. Bürünsel özellikler heceler, süreleri ve durakları, perde ve enerji ile ilgili özellikleri içermektedir ve bunların karşılaştırılmasında çeşitli yollar mevcuttur. Şekil 1 de gösterildiği gibi zaman ekseninde kayan bir kelime dizisi ele alındığında, bu dizi içerisinde sınırdan önceki ve sınırdan sonraki kelimeler için, bu iki kelime arasındaki durak, perde ve enerji ile ilgili değerler ve bunlar arasındaki karşılaştırmalar hesaplanabilir.

3.2. Yönteme İlişkin Temel Adımlar

Aşağıda projenin gerçekleştirilmesinde izlenen temel adımlar sıralanmaktadır.

1. Veri Toplama ve Etiketleme (Data Collection and Annotation): Projenin başlangıç dönemlerinde Boğaziçi Üniversitesi'nden Türkçe yayın haberlerine ilişkin verilerin elde edilmesi konusu araştırılmış, bunun yanısıra diğer kaynaklardan veri kaydedilmesi veya toplanmasına ilişkin çalışmalar yapılmıştır. Bu veriler yazıya uyarlanarak cümle sınırları dikkate alınarak etiketlenmişlerdir.

2. Zorlamalı Hizalamalar (Forced Alignments): İnsan eforu ile yada elle yazıya çevrilmiş kısımların audio dosya ile hizalanması yada üstüste getirilmesi SRI-International'ın Decipher otomatik konuşma tanıyıcısı tarafından yapılmıştır. Sesbirimlerin (phone) ve kelimelerin başlangıç ve bitiş süreleri tanıma araçları tarafından üretilmişlerdir. Bu bilgiler bürünsel bilginin çıkarılmasında kullanılmıştır. Tanıyıcıyı zorlamalı hizalama modunda çalıştırmak için, Türkçe için bir söyleyiş (pronunciation) sözlüğünden ve Türkçe işitsel (acoustic) model den yararlanılmıştır.

3. Biçimbilgisel Analiz (Morphological Analysis): Kelimelerin biçimbilgisel analizini yapabilmek için Sayın Prof. Dr. Kemal Oflazer'in araçları kullanılmıştır. Bu araçlar aynı zamanda biçimbirim/anlambirim (morpheme) ve tonlamaya ilişkin (inflectional) grup sınırlarını da işaretleyebilmektedir. Bunlar özellikle biçimbirim tabanlı özelliklerin çıkarılmasında yararlı olmuşlardır. Tanıyıcı çıkışındaki sesbirimlerin başlangıç ve bitiş zamanlarından biçimbirimlerin başlangıç ve bitiş zamanları hesaplanmıştır. Bazı kelimeler belirsizlikten (ambiguity) dolayı birden fazla biçimbirimsel analize sahiptirler. Örneğin "bakan" kelimesi devlet kademesindeki "bakan" anlamı ile yorumlanabileceği gibi bakmak fiilini gerçekleştiren özne (bak+an) olarak da yorumlanabilecektir. Özelliklerin hesaplanması sırasında belirsizliğin korunmasına devam edilmiş ve tüm çoğul versiyonlar için bürünsel özellikler kestirilmiştir. Bu özelliklere ilişkin veritabanı, daha sonra örneğin; diğer yararlı bilgilerin de birleştirilebileceği bürünsel özelliklerin kullanılması ile biçimbirimsel açıklama (morphological disambiguation) için kullanılabilir. Bu özelliklere ilişkin veritabanı, daha sonra örneğin; diğer yararlı bilgilerin de birleştirilebileceği bürünsel özelliklerin kullanılması ile biçimbirimsel açıklama (morphological disambiguation) için kullanılabilir.

4. Sözcüksel Özelliklerin Hesaplanması (Computation of Lexical Features): Herbir sınır için özellik olarak kelime unigram lar (kelime sınırından önceki ve sonraki kelime) ve bigram lar (sınırı çevreleyen kelimeler) çıkarılmıştır. Ayrıca, kelimelerin biçimbirimsel analizi ile ilgili işlevlerden diğer özelliklerin çıkarılması sağlanmıştır. Örneğin, bir kelimenin nihai kategori bilgisi fiil ise bu bilgi cümle bölütlemesi için yararlı olabilmektedir.

5. Bürünsel Özelliklerin Hesaplanması (Computation of Prosodic Features): Zorlanmış hizalamalar, audio dosyalar, Algemy ve PRAAT yazılımı ile kelime ve biçimbirim tabanlı perde (pitch) ve enerji ilişkili bürünsel özelliklerin çıkarılması gerçekleştirilmiştir.

6. Bölütleme Yöntemi (Segmentation Method): Boosting ve support vector machine sınıflandırma yöntemleri Saklı Dil Modelleri (HELMS-Hidden Event Language Models) ile birleştirilmişlerdir.

7. Değerlendirme Yöntemleri (Evaluation Methods): Cümle bölütleme işlemi için üç değerlendirme ölçütü kullanılmıştır. Bunlar; Sınıflandırma hata oranı (classification error rate), NIST-SU hata oranı (NIST-SU error rate) ve F-Ölçüm (F-Measure) değerlendirme ölçütleridir. Sınıflandırma hata oranı, tüm sınırlar içerisinde hatalı sınıflandırılmış sınırların yüzdesi olarak tanımlanabilir. NIST-SU hata oranı, hatalı sınıflandırılmış sınırların cümle sınırlarının toplam sayısına oranı olarak tanımlanmaktadır. F-Ölçüm ise sınıflandırıcılar tarafından cümle sınırı olarak varsayılmış olma bilgisinin insanlar tarafından etiketlenmiş bilgiye olan doğruluk ve duyarlılık (precision) ölçülerinin harmonik ortalamasıdır.

Ayrıca, yöntemin performansı, şans ve referans system performansları ile de karşılaştırılmıştır. Cümle bölütleme için şans performansı tüm sınırlar çoğunluk sınıfı olarak işaretlendiğindeki performanstır. Referans sistemi, sadece sözcüksel ve biçimbirimsel analiz ile ilgili özellikleri kullanan bir cümle bölütleme sistemi olarak ele alınmıştır.

3.3. Yöntemin Gerçeklenmesine ilişkin Detaylar

3.3.1. Türkçe Konuşma Verilerinin Hazırlanması

Projenin ilk 6 aylık döneminde proje çalışma takviminde planlandığı gibi öncelikle proje için belirlenen teknik (donanım ve yazılım) altyapısının kurulması işlemleri gerçekleştirilmiştir. Bu bağlamda öncelikle bilgisayar sisteminin satın alınması ve kurulması, yazılımların yüklenmesi ve ilk etapta yapılacak işler sırası ile gözden geçirilmiştir. Ayrıca projede çalışacak burslu öğrencinin tespit edilebilmesine yönelik çalışmalar yapılmıştır.

Çalışma takviminde öngörülen veri toplama (data collection) ve etiketleme (annotation) konusunda Boğaziçi Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, BUSİM Speech Processing Group¹ olanaklarından yararlanılmış ve özellikle öğretim üyesi sayın Yrd. Doç. Dr. Murat Saraçlar ve araştırma ekibinde bulunan sayın Sıddıka Parlak dan destek alınmıştır. Söz konusu veriler Boğaziçi Üniversitesi'nde bulunan BUSİM Speech Processing Group tarafından kaydedilen ve belirli projelerde üzerinde çalışılmakta olunan, Amerika'nın Sesi (Voice of America) Türkçe Yayın Bölümü² tarafından hazırlanan Türkçe radyo yayını haberlerine ilişkin konuşma verilerinden oluşmaktadır. Bu veriler belirli günlerde kaydedilmiş herbiri 30 dakikalık ve wav formatında kaydedilmiş çok sayıda konuşma dosyasından oluşmaktadır. Sözü edilen konuşma verilerine ilişkin aşağıda detayları ile verilen değişik formatlarda dosyalar hazırlanmıştır.

a) Amerika'nın Sesi Türkçe Yayın Bölümü'nün günlük sabah ve akşam saatlerinde yapmış olduğu yayınlardan kaydedilen konuşma verileri audio (wav) formatında 32kHz örnekleme frekansında ve 16bit linear PCM de örneklenmişlerdir.

b) Audio verilerin STM (segment time marks) başka bir deyiş ile herbir zaman aralığına (frame/segment) ilişkin başlangıç ve bitiş zamanları ile ona karşı gelen yazılı metni (transcription) içeren versiyonudur. Bu formatta hazırlanmış dosyalar aynı zamanda kaydedilen dosyanın adını, konuşmacıya ilişkin bilgileri (speaker-id, native-non native, male-female), konuşma kaydının arka plan bilgisi vb. bilgileri de içermektedir. Ayrıca bu dosyalar kişiler tarafından mümkün olabildiği ölçüde noktalama işaretlerine uygun olarak metne dönüştürülmüşlerdir. Referans transkripsiyon dosyaları olarak da adlandırılırlar. Bu dosyalar otomatik konuşma tanıyıcısına ilişkin çıkış dosyası ile karşılaştırılarak otomatik konuşma sisteminin performansının ölçülmesinde de kullanılmaktadır.

c) Audio verilere ilişkin kelime (word-level) ve phoneme (phone-level) tabanlı CTM (conversation time mark) dosyaları hazırlanmıştır. Bu dosyalar dosyanın adı, kanal sayısı ve kelime tabanlı ise her bir kelimenin başlangıç süresini, kelimenin başlangıç ve bitiş süreleri arasındaki toplam süre ile sözkonusu kelimenin metin bilgisini, phone tabanlı ise herbir kelimeyi oluşturan phone ların başlangıç ve toplam süreleri ile sözkonusu phone ların metin bilgisini içermektedir. Belirtilen bu verilere ilişkin dosyaların elde edilmesinin yanısıra, otomatik konuşma sistemine ilişkin çıkış (asr-output) verilerinin elde edilebilmesi de sağlanmıştır.

¹ <http://www.busim.ee.boun.edu.tr/>

² <http://www.voanews.com/turkish/>

Elde edilen belirli sayıdaki konuşma verisine ilişkin dosyalar üzerinde çalışılarak bunların mümkün olduğunca birbirleri ile eşleştirilmeleri ve bazı hataların giderilmesine çalışılmıştır. Bu sayede birbirleri ile eşleştirilmiş 1) wav, 2) stm, 3) kelime tabanlı ctm ve 4) phone tabanlı ctm formatında hazırlanmış tüm dosyaları mevcut olan toplam 42 adet farklı kayıta ilişkin dosyalar elde edilmiştir. Bu dosyalardan wav dosyalarının herbiri 30 dakikalık ve tek kanal olarak kayıt edilmiş haber içerikli konuşma kayıtlarını içermekte olup, verilerin toplamı 21 saatlik bir veriye karşı gelmektedir. Bir perl kodu ile özellikle stm dosyalarındaki cümlelere ilişkin kelimeler tek tek elde edilerek ctm dosyalarındaki kelimeler ile hizalanmışlardır. Tespit edilen hatalar Boğaziçi Üniversitesi BUSİM Speech Processing Group ile paylaşılmış ve bazı dosyaların yeniden elde edilmesi sağlanmıştır. Projeye ilişkin tüm işlemler bu dosyalar üzerinde gerçekleştirilmiştir.

Öncelikle referans dosyaların (stm) kullanılması ile herbir dosyada yeralan herbir kelime için ilgili kelimenin cümle sınırı olup olmadığını (sentence boundary/non-sentence boundary) belirleyen bir perl kodu yazılmıştır. Bu işlemde, stm dosyalarında yeralan kelimelerin sonunda nokta gibi cümle sonunu belirten noktalama işaretlerinin bulunmaması durumunda, kelime sınırı, cümle sınırı olmadığını gösteren n etiketi ile (n: non-sentence boundary), bulunması durumunda ise cümle sınırı olduğunu gösteren s (s: sentence boundary) etiketi ile etiketlenmişlerdir.

Yapılan bu işlemler sonucunda belirlenen 42 adet dosyada toplam 127,915 adet kelime olduğu tespit edilmiştir. Sözcüksel, biçimbilgisel ve bürünsel özellik çıkarımı (feature extraction) işlemlerinden sonra, elde edilen dosyaların büyük bir kısmı training set, bir kısmı development set bir kısmı ise test set olarak ayrılmaktadırlar. Özellikle cümle bölütleme sistemimizin performansı bu development ve test set üzerinde ölçülmektedir.

3.3.2. Türkçe'nin Biçimbilgisel Analizi (Morphological Analysis) ve Biçimbilgisel Özelliklerin (Morphological Features) Çıkarılması:

Türkçe cümle öge sıralamasının serbest olarak kullanılabilirdiği (free-constituent language) bir dildir. Bununla birlikte cümle öğelerinin (part of speech) tipik sırası özellikle haber tarzında konuşma türlerinde özne-nesne-yüklem (SOV-subject-object-verb) biçimindedir.

Türkçe'de "Çocuk yemek yedi" cümlesini ele aldığımızda İngilizce'deki karşılığı "The child ate the meal" olacaktır. Doğru bir biçimde yapılan biçimbilgisel analiz;

çocuk: Noun+A3sg+Pnon+Nom (the child)
yemek: Noun+ A3sg+Pnon+Nom (the meal)
yedi: Verb+Pos(+dH)+Past+A3sg (ate)

biçiminde olacaktır.

Türkçe, çekimsel (inflectional) ve türetsel (derivational) sondan eklemelerin (suffixation) yapılabildiği bitişken bir biçimbilgisine (agglutinative morphology) sahiptir (Oflazer, 1994). Bir kökten çok sayıda yeni kelime formu türetilebilir (Hankamer, 1989). Aşağıda verilen örnekte olduğu gibi, herhangi bir yüklem, isim veya sıfatın sonuna sadece 3 adet biçimbirim (morpheme) in sondan eklenmesi ile yeni kelime formları türetilebilmektedir (Hakkani-Tür ve diğerleri, 2002).

Türkçe'de biçimbilgisel bilgi genel olarak aşağıda verilen biçimde gösterilebilir.

Root + IG₁ + ^DB + IG₂ + ^DB + ... + ^DB + IG_n

Bu gösterimde ((Ofłazer, 1999) den uyarlanarak) çekimsel gruplar (IG-Inflectional Groups) türetsel sınırları göstermekte ve türetsel sınırlar (DB-Derivational Boundaries) ^DB ile işaretlenmektedirler. Bir kelimenin kök ve türetsel elemanları farklı IG ler ile temsil edilebilmektedir. Herbir IG_i uygun bir çekimsel özellik dizisini göstermektedir. Bu çekimsel özelliklerin bazıları aşağıdaki gibi listelenebilirler.

- +Adj: adjective (sıfat)
- +Noun: noun (isim)
- +Verb: verb (yüklem)
- +A3sg: 3rd person singular agreement (3. tekil şahıs uyumu)
- +P1sg: 1st person singular possessive agreement (1.tekil şahıs iyelik (mülkiyet) uyumu)
- +Pnon: no possessive agreement (iyelik uyumu bulunmaması)
- +Nom: nominative case (yalın durum)
- +Past: past tense (geçmiş zaman)
- +Fut: future tense (gelecek zaman)
- +FutPart: future participle (gelecek zaman ortağı)

Bir örnek olarak “yapabileceğim” kelimesini analiz edelim. Yapabileceğim kelimesi; (yap) + (abil) + (ecek) + (im) biçiminde biçimbirimlerine ayrılabilir.

Bu kelimeye ilişkin potansiyel 3 farklı biçimbilgisel analiz yapılabilir. Bunlar;

- 1) (yap) yap+Verb+Pos(+yAbil)^DB+Verb+Able(+yAcAk)+Fut(+yHm)+A1sg
[I will be able to do it]
- 2) (yap) yap+Verb+Pos(+yAbil)^DB+Verb+Able(+yAcAk)^DB+Adj+FutPart(+Hm)+P1sg
[The (thing that) I will be able to do]
- 3) (yap)
yap+Verb+Pos(+yAbil)^DB+Verb+Able(+yAcAk)^DB+Noun+FutPart+A3sg(+Hm)+P1sg+
Nom
[The one I will be able to do]

Bu örnekte, kök bir yüklem olmakla birlikte herbir analizin en sonunda yer alan çekimsel gruplar (IG) sırası ile yüklem (Verb), sıfat (Adj) ve isimdir (Noun).

Türkçe’de yer alan kelimeler yapı itibari ile biçimbilgisel olarak ayrılacak olası çok sayıda birimlerden (cümlelerin öğeleri) (POS-Part of Speech Tags) oluşabildiğinden istatistiksel modeller için oldukça ilginç bir problem ortaya koymaktadır. Bu özellik, Türkçe’nin oldukça üretken ve türetsel bir biçimbilgisel yapıya sahip olmasından ileri gelmektedir.

Yaptığımız biçimbilgisel analizler Sayın Prof. Dr. Kemal Ofłazer’in (Ofłazer, 1994) geliştirmiş olduğu biçimbilgisel analiz yapabilen araçlar kullanılarak gerçekleştirilmiştir. Dizayn ettiğimiz biçimbilgisel özellikler (morphological features) ise bu analizlerden sonra dizayn edilmişlerdir. Biçimbilgisel özellikler dizayn edilirken, sözkonusu araçların herbir kelime için verdiği olası tüm biçimbilgisel olarak ayrılmış kısımlar kullanılmıştır (Bkz. Yukarıda verilen örnek).

Geliştirdiğimiz en önemli biçimbilgisel özelliklerden bir tanesi, modellerimizi basitleştirmek için kullandığımız ve herbir kelimenin en sonunda yer alan çekimsel gruplardır (IG ler). Bu gruplar cümledeki enson kategoriyi işaret etmektedir.

Kelime-ifade belirsizliğini (word-sense ambiguity) çözümleneksizin her bir kelimeye ilişkin en son çekimsel grup ile birlikte biçimbilgisel olası cümle öğeleri de (POS-Part of Speech Tags) elde edilmiştir.

Cümle öğelerinin etiketlerinin (Part of Speech Tags) özellik değeri olarak, eğer kelime birden fazla ayrıştırılmış kısım (parse) içeriyorsa bilinmeyen (unknown) olarak işaretlenmektedir. Ayrıca, bir kelimenin herhangi bir biçimbilgisel ayrışımının final kategorisine göre yüklem olup olmadığını denetleyen bir başka iki durumlu biçimbilgisel özellik (binary feature) daha çıkarılmıştır. Böylece bu sayede Türkçe'nin SOV (subject, object, verb) doğasının avantajının kullanılması öngörülmektedir. Bu özelliklerden başka her bir kelimenin son üç harfini içeren bir başka biçimbilgisel özellik daha çıkarılmıştır (pseudo morphological features). İngilizce'deki "ed" son ekine benzer biçimde Türkçe de de belirli son ekler yüklem kategorisine işaret edebilmektedir. Bu durum, özellikle cümle bölütleme işleminde, cümle sınırlarının belirlenmesinde önemli bir ipucu olarak yarar sağlamaktadır.

Özellikle yüklem bilgisi Türkçe'nin biçimbilgisel analizinin getirdiği dilbilimsel açıdan en önemli özelliktir. Her ne kadar Türkçe, kelime sırasının serbestçe kullanılabilirdiği bir dil olsa da, en sık kullanılan öğe sıralaması SOV biçiminde olmaktadır. Özellikle radyo haber yayınlarında ve Türkçe'nin formal kullanımında öğe sıralamasının daima SOV biçiminde olduğu görülmektedir. Bu nedenle, özellikle, biçimbilgisel analizlerden biri yüklem formunda ise bu durum cümle sonunu işaret eden en önemli belirteçlerden biri olmaktadır. Dolayısı ile bu ipucu yada özellik, cümle sınırlarının belirlenmesi, başka bir ifade ile cümle bölütlemesi için oldukça yararlı bir bilgiyi taşımaktadır.

3.3.3. Türkçe Konuşma Verilerinin Sözcüksel Özelliklerinin (Lexical features) Çıkarılması

Word n -grams: Verilen bir kelime dizisinden maksimum n adet kelimedenden oluşan alt diziler yada kelime kombinasyonlarıdır. Örneğin "word 3-grams" word trigrams olarak adlandırılır ve tekli, ikili, üçlü kelime gruplarını içerir. Kullandığımız sözcüksel özellikler her bir kelime sınırı için oluşturulmuş 6 adet kelime n -gram özellikleridir. Bunlardan 3 adet unigram, 2 adet bigram ve 1 adet trigram oluşturulmuştur. n -gramlar cümle sınırı olup olmadığı ile ilgilenilen kelime (current word), bir sonraki kelime (next word) ve bir önceki kelime (previous word) nin yanyana getirilmiş kombinasyonları ile oluşturulmaktadır (Cuendet ve diğerleri, 2007). Perl dilinde yazılan bir kod ile Türkçe konuşma verilerine ilişkin tüm kelimelerin sözcüksel özellikleri aşağıda verilen formatta çıkarılmıştır.

6 adet sözcüksel özellik aşağıda verilmektedir.

Unigram lar : {bir önceki (previous)}, {şimdiki (current)}, {bir sonraki (next)}

Bigram lar : {şimdiki, bir sonraki}, {bir önceki, şimdiki}

Trigram lar : {bir önceki, şimdiki, bir sonraki}

Sözcüksel özellikler aşağıda verilen formatta hazırlanmışlardır.

s, n : cümle sınırı, cümle sınırı değil (sentence : s, nonsentence boundary : n)

w : güncel yada ilgilenilen kelime (current word : w_i)

wn : bir sonraki kelime (next word : w_{i+1})

wp : bir önceki kelime (previous word : w_{i-1})

wwn : güncel kelime ve bir sonraki kelime (current word and next word : w_i, w_{i+1})

wwp : güncel kelime ve bir önceki kelime (current word and previous word : w_i, w_{i-1})

$wpwn$: bir önceki kelime, güncel kelime, bir sonraki kelime (previous word, current word, next word: w_{i-1}, w_i, w_{i+1})

Burada w_i ilgilenilen cümle sınırından önceki son kelimeyi işaret eder.

Sözcüksel özellikler (lexical features) (word *n*-grams)

s,n.
w: text.
wn: text.
wp: text.
wwn: text.
wpp: text.
wpwwn: text.

veya gerçek formatta;

w,wn,wp,wwn,wpp,wpwwn, boundary (s veya n).

Türkçe konuşma verileri için oluşturulmuş sözcüksel (lexical) özelliklere bir örnek.

Örnek Cümle: Burası Amerikanın sesi Türkçe yayın bölümü Washington.
(burasI amerikanIn sesi tUrkCe yayIn bOIUmU \$washington.)

burasI,-,- burasI amerikanIn,amerikanIn,burasI amerikanIn, - burasI, **n**.
amerikanIn,burasI,burasI amerikanIn sesi,sesi,amerikanIn sesi, burasI amerikanIn, **n**.
sesi,amerikanIn,amerikanIn sesi tUrkCe,tUrkCe,sesi tUrkCe, amerikanIn sesi, **n**.
tUrkCe,sesi,sesi tUrkCe yayIn,yayIn,tUrkCe yayIn, sesi tUrkCe, **n**.
yayIn,tUrkCe,tUrkCe yayIn bOIUmU,bOIUmU,yayIn bOIUmU, tUrkCe yayIn, **n**.
bOIUmU,yayIn,yayIn bOIUmU \$washington,\$washington,bOIUmU \$washington, yayIn bOIUmU, **n**.
\$washington,bOIUmU,bOIUmU \$washington bugUn,bugUn,\$washington bugUn, bOIUmU
\$washington, **s**.

3.3.4. Türkçe Konuşma Verilerinin Bürünsel Özelliklerinin (Prosodic features) Çıkarılması

Bürün (prosody) genel olarak bir dilin vurgu, perde, tonlama, duraksama, vb. özelliklerini içerir. Bu özellikler, anlam farkı yaratmaktan çok, niyete veya işleve işaret eder. Soru sorulurken son hecelerin yüksek bir tonlamayla ifade edilmesi, keskin bir vurgunun kızgınlık, yüksek bir tınının kaygı ifade edebilmesi buna birkaç örnektir.

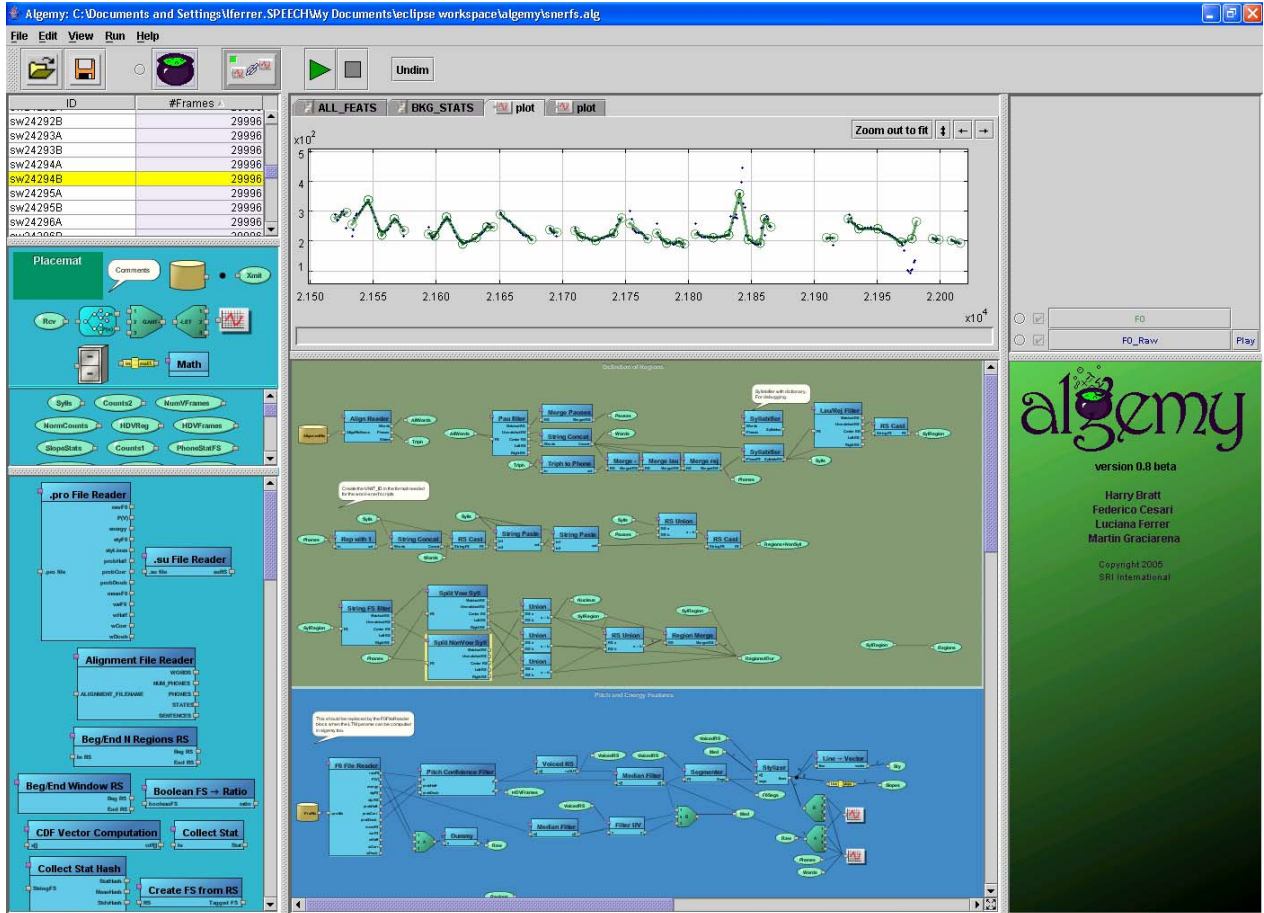
Konuşma işaretlerinin otomatik konuşma sistemleri tarafından basit kelime dizilerine dönüştürülmesi sonucu konuşma ile ilgili zamanlama ve perde yapılarına ilişkin bilgiler veya özellikler kaybolur. Bu yapılar (ve bu yapılarla bağlı kelimelerden bağımsız diğer özellikler) bürün olarak adlandırılır. Tüm dillerde bürün, yapısal, anlamsal ve işlevsel bilgiyi taşımaktadır. Bürünsel özellikler veya ipuçları dilin söylem yapısı ile ilişkili olduklarından özellikle çeşitli bilgi çıkarılması (information extraction) işlevlerinde önemli bir rol oynamaktadır. Dilbiliminde, okuma ve anında yapılan monologların analizinde ve buna bağlı işlevlerde cümleler ve paragraflar gibi bilgi birimlerinin bürünsel olarak işaretlenmedikleri gösterilmiştir. İngilizce'de ve ona benzer veya bağlı dillerde bürünsel belirteçler olarak duraklar, perde aralığındaki ve genliğindeki değişim, genel perde eğimi, ezgi ve sınır ton dağılımı ve konuşma oranındaki değişim gibi özellikler sayılabilir. Örneğin cümle veya paragraf sınırları ile konu sınırlarının her ikisi de bazı uzun durakların, tonda meydana gelen bir düşmenin ve perdenin yeniden düzenlenmesinin olduğu durumlarda işaretlenir. Bunlar sınır belirleme veya işaretleme için önemli göstergelerdir. Ayrıca, bürünsel özellikler doğası gereği göreceli bir biçimde kelimenin kimliğinden etkilenmediğinden, başka bir ifade ile kelimelerden bağımsız olduklarından otomatik konuşma sistemi çıkışını baz alan sözcüksel bilgi çıkarım yöntemlerinin gürbzlüğünü (robustness) geliştirmektedir. Bürünsel özellikler ICSI+ nin cümle bölütleme sisteminden transfer edilmiştir (Zimmermann ve diğerleri, 2006).

Kullanılan bürünsel özelliklerden bazılarının ilişkin özet bilgiler aşağıda verilmiştir.

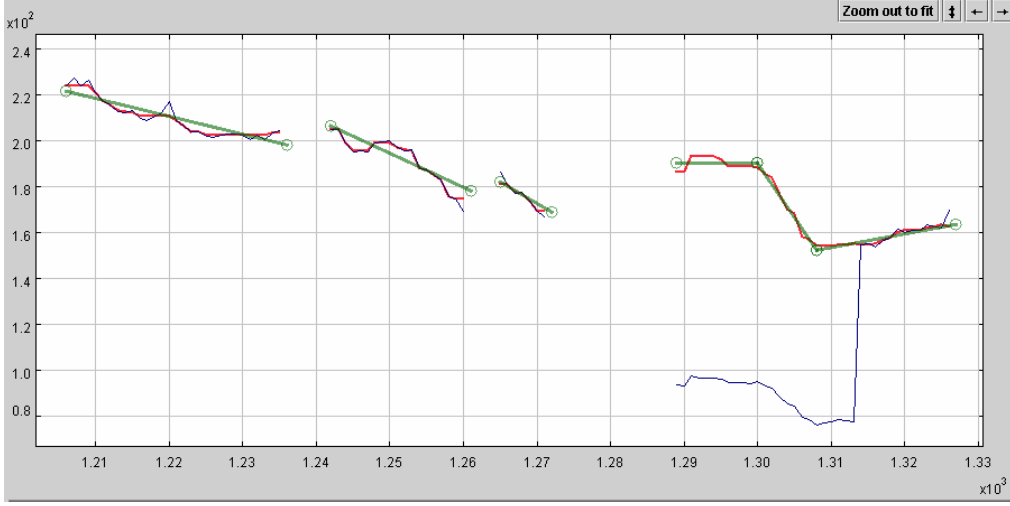
3.3.4.1. AlgeMy (SRI-International’ın bürünsel özellikleri hesaplama yazılımı) ve Temel Bürünsel Özellikler

Bürünsel özelliklerin tamamı SRI-International tarafından geliştirilen ve AlgeMy adı verilen Graphical User Interface yapılarındaki bir yazılım ile gerçekleştirilmiştir. Yazılıma ilişkin sınıflandırılmış bürünsel özellikler aşağıda verildiği gibidir.

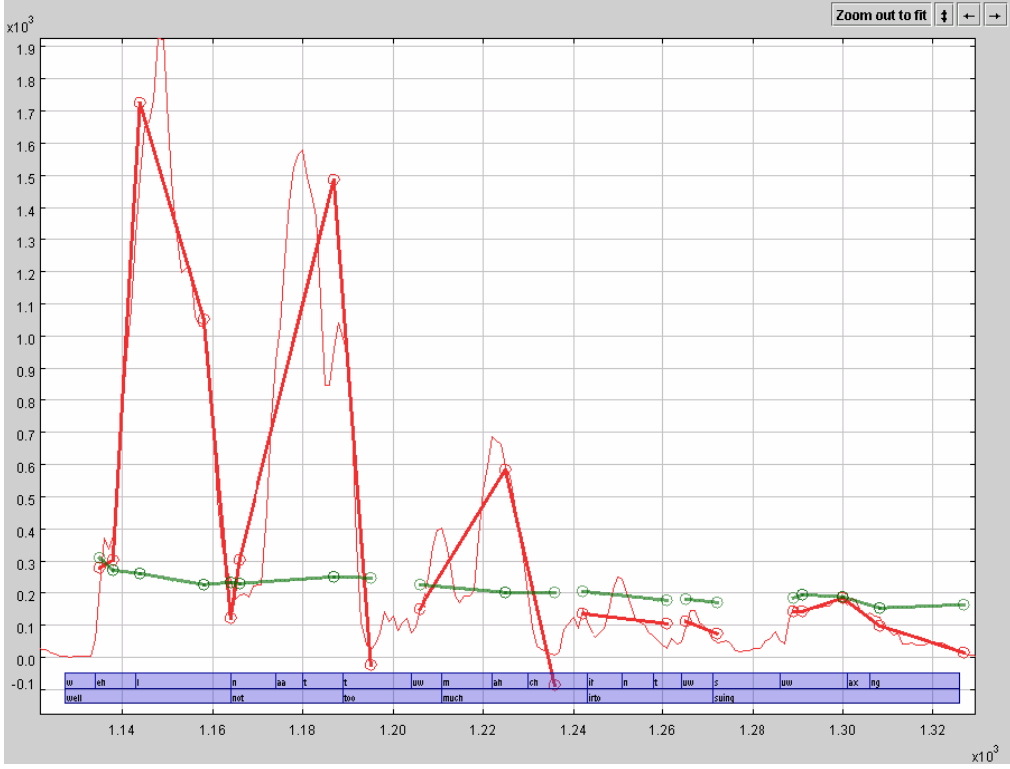
- Durak Özellikleri (pause features):
Aralıktan önceki ve sonraki durak uzunlukları
- Süre Özellikleri (duration features):
Aralığın toplam süresi
Aralıktaki en yüksek sesli harf (vowel) süresi
- Pitch özellikleri:
Aralıktaki biçimlendirilmiş pitch in ortalama/enyüksek/endüşük değerleri
Aralıktaki ilk/son pitch eğimi. Eğimin uzunluğu.
Aralıktaki enyüksek pozitif/negatif pitch eğiminin değeri
- Enerji özellikleri (energy features)
Biçimlendirilmiş enerjinin ilk/son eğim değerleri
Aralıktaki enyüksek pozitif/negatif enerji eğiminin değeri



Şekil 4. AlgeMy Ana Menü

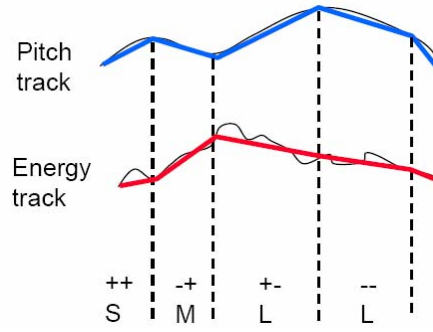


Şekil 5. Biçimlendirilmiş pitch (stylized pitch)



Şekil 6. Biçimlendirilmiş enerji (stylized energy)

Özelliklerin belirlenmesinde en basit yaklaşım pitch ve enerji dağılımlarının (pitch and energy distributions) modellenmesidir. Özellikler herbir frame için, pitch in logaritması $\log(\text{pitch})$, enerjinin logaritması $\log(\text{energy})$ ve sesli frame ler için delta özellikleridir. Bu özellikler UBM-GMM (Universal Background Model-Gaussian Mixture Models) ile modellenmektedir. Daha az basit olan bir yaklaşım da, pitch enerji ve işaret sürelerinin modellenmesidir. Bu yaklaşımda pitch ve sayısal olarak süreleri ile etiketlenmiş pitch ve enerji izlerinin eğimlerini açıklayan bir semboller dizisi yaratmaktır. Bu özellikler ayrık saklı Markov modelleri ile (DHMMs-Discrete Hidden Markov Models) ile modellenmektedirler.



Şekil 7. Pitch ve enerji izleri

Temel bazı bürünsel özellikler (Base prosodic features):

PAUSE_DUR: Kelime sınırındaki durağın süresidir. Eğer bu sınır dalga biçimleri arasındaki bir sınır ise, toplam durma süresi, sınırdan önceki dalga biçiminde yer alan son kelimenin sonundan (aynı konuşmacıya ilişkin konuşmalar bazında) sınırdan sonraki dalga biçiminde yer alan ilk kelimenin başlangıcına kadar olan süredir. Eğer önceki (sonraki) dalga biçimi tümüyle sessiz bir dalga biçimi ise o anki dalga biçimi için ilk (son) durma süresi öncekinin ilk çerçevesinden (bir sonrakinin son çerçevesine kadar) alınır.

PATTERN BOUNDARY: PATTERN_NEXT_WORD ile birleştirilen PATTERN_WORD ün son “f” veya “r” sinin PATTERN_NEXT_WORD ün ilk “f” veya “r” si arasındaki sınır. (PATTERN_WORD: Bu özellik; bir alçalan eğim, bir sessiz bölge ve bir sınırdan önceki kelimedeki yükselen eğimi temsil eden “f”, “uv” ve “r” dizisi ile tanımlanır. Bu dizi min_frame_length den daha az ise stilize edilmiş f0 dosyasındaki f ler, “r” ler veya “uv” ler atlanır. Daha uzun diziler sadece bir “f”, “r” ve “uv” ile temsil edilir. Farklı eğimli “f” veya “r”, dizileri “ff” veya “rr” ile temsil ettirilir.)

SLOPE_DIFF: Son non_zero (min_frame_length den uzun) kelime eğimi ile bir sonraki kelimenin ilk non_zero (min_frame_length den uzun) eğimi arasındaki farktır. Eğer min_frame_length zamanlarından daha çok oluşan non_zero eğime sahip olmayan kelimeler var ise bu özellik “X” değerine sahip olur.

PAU_DUR_PREV: Bir önceki sınıra ilişkin değerler olup, PAU_DUR açıklamasının aynısı geçerlidir.
rp: gerçek durak (real pause)

f0 (fundamental frequency) temelli bazı özellikler (f0 derived features):

Önceki ve sonraki kelimelerin maksimum ve minimum değerlerinin oranının logaritmasına ilişkin özellikler:

$$F0K_WRD_DIFF_HIHI_N = \log(\text{MAX_PWLFIT_F0} / \text{MAX_PWLFIT_F0_NEXT})$$

$$F0K_WRD_DIFF_HILO_N = \log(\text{MAX_PWLFIT_F0} / \text{MIN_PWLFIT_F0_NEXT})$$

$$F0K_WRD_DIFF_LOLO_N = \log(\text{MIN_PWLFIT_F0} / \text{MIN_PWLFIT_F0_NEXT})$$

$$F0K_WRD_DIFF_LOHI_N = \log(\text{MIN_PWLFIT_F0} / \text{MAX_PWLFIT_F0_NEXT})$$

$$F0K_WRD_DIFF_MNMN_N = \log(\text{MEAN_PWLFIT_F0} / \text{MEAN_PWLFIT_F0_NEXT})$$

Bu ifadelerdeki MAX_PWLFIT_F0, MAX_PWLFIT_F0_NEXT, MIN_PWLFIT_F0, MIN_PWLFIT_F0_NEXT, MEAN_PWLFIT_F0 ve MEAN_PWLFIT_F0_NEXT değerlerinin tümü f0 (fundamental frequency) değerleridir.

Üç kelimelerdeki pitch değerlerinin oranının logaritmasına ilişkin bazı özellikler:

$$F0K_WRD_DIFF_BEGBEG = \log(\text{FIRST_PWLFIT_F0} / \text{FIRST_PWLFIT_F0_NEXT})$$

$$F0K_WRD_DIFF_ENDBEG = \log(\text{LAST_PWLFIT_F0} / \text{FIRST_PWLFIT_F0_NEXT})$$

$$F0K_INWRD_DIFF = \log(\text{FIRST_PWLFIT_F0} / \text{LAST_PWLFIT_F0})$$

Bu ifadelerdeki FIRST_PWLFIT_F0, LAST_PWLFIT_F0, FIRST_PWLFIT_F0_NEXT değerleri f0 (fundamental frequency) değerleridir.

Eğimdeki normalizasyonlara ilişkin özellikler:

SLOPE_DIFF_N: SLOPE_DIFF / SPKR_FEAT_SD_SLOPE

LAST_SLOPE_N: LAST_SLOPE / LAST_PWLFIT_F0

Bu ifadelerdeki SLOPE_DIFF, LAST_SLOPE ve LAST_PWLFIT_F0 f0 özellikleridir.

SPKR_FEAT_SD_SLOPE ise konuşmacı (speaker) temelli özelliklerden biridir.

Bürünsel özellikler (prosodic features) den bazıları

s,n.

PAUSE_DUR: continuous.

PATTERN_BOUNDARY: X, f+f, f+r, r+f, r+r.

ENERGY_PATTERN_BOUNDARY: X, f+f, f+r, r+f, r+r.

SLOPE_DIFF: continuous.

ENERGY_SLOPE_DIFF: continuous.

SLOPE_LAST: continuous.

ENERGY_SLOPE_LAST: continuous.

SLOPE_LAST_N: continuous.

ENERGY_SLOPE_LAST_N: continuous.

WRD_F0K_DIFF_HIHI_N: continuous.

WRD_ENERGY_DIFF_HIHI_N: continuous.

WRD_F0K_DIFF_HILO_N: continuous.

WRD_ENERGY_DIFF_HILO_N: continuous.

WRD_F0K_DIFF_LOLO_N: continuous.

WRD_ENERGY_DIFF_LOLO_N: continuous.

WRD_F0K_DIFF_LOHI_N: continuous.

WRD_ENERGY_DIFF_LOHI_N: continuous.

F0K_WIN_DIFF_HIHI_N: continuous.

ENERGY_WIN_DIFF_HIHI_N: continuous.

F0K_WIN_DIFF_HILO_N: continuous.

ENERGY_WIN_DIFF_HILO_N: continuous.

F0K_WIN_DIFF_LOLO_N: continuous.

ENERGY_WIN_DIFF_LOLO_N: continuous.

F0K_WIN_DIFF_LOHI_N: continuous.

ENERGY_WIN_DIFF_LOHI_N: continuous.

WRD_F0K_DIFF_MNMN_N: continuous.

WRD_ENERGY_DIFF_MNMN_N: continuous.

F0K_WRD_DIFF_BEGBEG: continuous.

ENERGY_WRD_DIFF_BEGBEG: continuous.

F0K_WRD_DIFF_ENDBEG: continuous.

ENERGY_WRD_DIFF_ENDBEG: continuous.

F0K_INWRD_DIFF: continuous.

ENERGY_INWRD_DIFF: continuous.

PAU_DUR_PREV: continuous.

.....

.....

rp: continuous.

Veya gerçek formatta;

PAUSE_DUR,PATTERN_BOUNDARY, ENERGY_PATTERN_BOUNDARY, SLOPE_DIFF,
ENERGY_SLOPE_DIFF, SLOPE_LAST, ENERGY_SLOPE_LAST, SLOPE_LAST_N,
ENERGY_SLOPE_LAST_N, WRD_F0K_DIFF_HIHI_N, WRD_ENERGY_DIFF_HIHI_N,
WRD_F0K_DIFF_HILO_N, WRD_ENERGY_DIFF_HILO_N, WRD_F0K_DIFF_LOLO_N,
WRD_ENERGY_DIFF_LOLO_N, WRD_F0K_DIFF_LOHI_N, WRD_ENERGY_DIFF_LOHI_N,
F0K_WIN_DIFF_HIHI_N, ENERGY_WIN_DIFF_HIHI_N, F0K_WIN_DIFF_HILO_N,
ENERGY_WIN_DIFF_HILO_N, F0K_WIN_DIFF_LOLO_N, ENERGY_WIN_DIFF_LOLO_N,

F0K_WIN_DIFF_LOHI_N, ENERGY_WIN_DIFF_LOHI_N, WRD_F0K_DIFF_MNMN_N,
WRD_ENERGY_DIFF_MNMN_N, F0K_WRD_DIFF_BEGBEG, ENERGY_WRD_DIFF_BEGBEG,
F0K_WRD_DIFF_ENDBEG, ENERGY_WRD_DIFF_ENDBEG, F0K_INWRD_DIFF,
ENERGY_INWRD_DIFF, PAU_DUR_PREV,, rp, boundary (s veya n).

Bürünsel (prosodic) ve Sözcüksel (lexical) özellikler birarada

s,n.

PAUSE_DUR: continuous.

PATTERN_BOUNDARY: X, f+f, f+r, r+f, r+r.

ENERGY_PATTERN_BOUNDARY: X, f+f, f+r, r+f, r+r.

SLOPE_DIFF: continuous.

ENERGY_SLOPE_DIFF: continuous.

SLOPE_LAST: continuous.

ENERGY_SLOPE_LAST: continuous.

SLOPE_LAST_N: continuous.

ENERGY_SLOPE_LAST_N: continuous.

WRD_F0K_DIFF_HIHI_N: continuous.

WRD_ENERGY_DIFF_HIHI_N: continuous.

WRD_F0K_DIFF_HILO_N: continuous.

WRD_ENERGY_DIFF_HILO_N: continuous.

WRD_F0K_DIFF_LOLO_N: continuous.

WRD_ENERGY_DIFF_LOLO_N: continuous.

WRD_F0K_DIFF_LOHI_N: continuous.

WRD_ENERGY_DIFF_LOHI_N: continuous.

F0K_WIN_DIFF_HIHI_N: continuous.

ENERGY_WIN_DIFF_HIHI_N: continuous.

F0K_WIN_DIFF_HILO_N: continuous.

ENERGY_WIN_DIFF_HILO_N: continuous.

F0K_WIN_DIFF_LOLO_N: continuous.

ENERGY_WIN_DIFF_LOLO_N: continuous.

F0K_WIN_DIFF_LOHI_N: continuous.

ENERGY_WIN_DIFF_LOHI_N: continuous.

WRD_F0K_DIFF_MNMN_N: continuous.

WRD_ENERGY_DIFF_MNMN_N: continuous.

F0K_WRD_DIFF_BEGBEG: continuous.

ENERGY_WRD_DIFF_BEGBEG: continuous.

F0K_WRD_DIFF_ENDBEG: continuous.

ENERGY_WRD_DIFF_ENDBEG: continuous.

F0K_INWRD_DIFF: continuous.

ENERGY_INWRD_DIFF: continuous.

PAU_DUR_PREV: continuous.

.....

.....

rp: continuous.

w: text.

wn: text.

wp: text.

wwn: text.

wwp: text.

wpwwn: text.

Veya gerçek formatta;

PAUSE_DUR, PATTERN_BOUNDARY, ENERGY_PATTERN_BOUNDARY,
SLOPE_DIFF, ENERGY_SLOPE_DIFF, SLOPE_LAST, ENERGY_SLOPE_LAST, SLOPE_LAST_N,
ENERGY_SLOPE_LAST_N, WRD_F0K_DIFF_HIHI_N, WRD_ENERGY_DIFF_HIHI_N,

WRD_F0K_DIFF_HILO_N, WRD_ENERGY_DIFF_HILO_N, WRD_F0K_DIFF_LOLO_N,
WRD_ENERGY_DIFF_LOLO_N,
WRD_F0K_DIFF_LOHI_N, WRD_ENERGY_DIFF_LOHI_N, F0K_WIN_DIFF_HIHI_N,
ENERGY_WIN_DIFF_HIHI_N, F0K_WIN_DIFF_HILO_N, ENERGY_WIN_DIFF_HILO_N,
F0K_WIN_DIFF_LOLO_N,
ENERGY_WIN_DIFF_LOLO_N, F0K_WIN_DIFF_LOHI_N, ENERGY_WIN_DIFF_LOHI_N,
WRD_F0K_DIFF_MNMN_N, WRD_ENERGY_DIFF_MNMN_N, F0K_WRD_DIFF_BEGBEG,
ENERGY_WRD_DIFF_BEGBEG, F0K_WRD_DIFF_ENDBEG, ENERGY_WRD_DIFF_ENDBEG,
F0K_INWRD_DIFF, ENERGY_INWRD_DIFF, PAU_DUR_PREV,, rp, w, wn, wp, wwn, wwp, wpwwn,
boundary (s veya n).

Türkçe konuşma verileri için oluşturulmuş bürünsel (prosodic) özelliklere bir örnek

Örnek Cümle: Burası Amerikanın sesi Türkçe yayın bölümü Washington.
(burası amerikanın sesi türkçe yayın bölümü Washington.)

Özellik isimleri (prosodic.names dosyası)

n,s.

ENERGY_DIFF_ENDBEG:continuous.
ENERGY_DIFF_FIRSTNEXT_KBASELN:continuous.
ENERGY_DIFF_LAST_KBASELN:continuous.
ENERGY_FIRSTNEXT_SLOPE:continuous.
ENERGY_FIRSTNEXT_SLOPE_N:continuous.
ENERGY_LAST_SLOPE:continuous.
ENERGY_LAST_SLOPE_N:continuous.
ENERGY_LR_FIRSTNEXT_KBASELN:continuous.
ENERGY_LR_LAST_KBASELN:continuous.
ENERGY_PATTERN_BOUNDARY:text.
ENERGY_PREVWIN20_NEXTWIN20_DIFF_BEGBEG:continuous.
ENERGY_PREVWIN20_NEXTWIN20_DIFF_MINNEXT_KBASELN:continuous.
ENERGY_PREVWIN20_NEXTWIN20_DIFF_MIN_KBASELN:continuous.
ENERGY_PREVWIN20_NEXTWIN20_HIHI_N:continuous.
ENERGY_PREVWIN20_NEXTWIN20_HIHI_NG:continuous.
ENERGY_PREVWIN20_NEXTWIN20_HILO_N:continuous.
ENERGY_PREVWIN20_NEXTWIN20_HILO_NG:continuous.
ENERGY_PREVWIN20_NEXTWIN20_INWRD_DIFF:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LOHI_N:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LOHI_NG:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LOLO_N:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LOLO_NG:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LR_MINNEXT_KBASELN:continuous.
ENERGY_PREVWIN20_NEXTWIN20_LR_MIN_KBASELN:continuous.
ENERGY_PREVWIN50_NEXTWIN50_DIFF_BEGBEG:continuous.
ENERGY_PREVWIN50_NEXTWIN50_DIFF_MINNEXT_KBASELN:continuous.
ENERGY_PREVWIN50_NEXTWIN50_DIFF_MIN_KBASELN:continuous.
ENERGY_PREVWIN50_NEXTWIN50_HIHI_N:continuous.
ENERGY_PREVWIN50_NEXTWIN50_HIHI_NG:continuous.
ENERGY_PREVWIN50_NEXTWIN50_HILO_N:continuous.
ENERGY_PREVWIN50_NEXTWIN50_HILO_NG:continuous.
ENERGY_PREVWIN50_NEXTWIN50_INWRD_DIFF:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LOHI_N:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LOHI_NG:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LOLO_N:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LOLO_NG:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LR_MINNEXT_KBASELN:continuous.
ENERGY_PREVWIN50_NEXTWIN50_LR_MIN_KBASELN:continuous.
ENERGY_PREVWRD1_NEXTWRD1_DIFF_BEGBEG:continuous.

ENERGY_PREVWRD1_NEXTWRD1_DIFF_MINNEXT_KBASELN:continuous.
ENERGY_PREVWRD1_NEXTWRD1_DIFF_MIN_KBASELN:continuous.
ENERGY_PREVWRD1_NEXTWRD1_HIHI_N:continuous.
ENERGY_PREVWRD1_NEXTWRD1_HIHI_NG:continuous.
ENERGY_PREVWRD1_NEXTWRD1_HILO_N:continuous.
ENERGY_PREVWRD1_NEXTWRD1_HILO_NG:continuous.
ENERGY_PREVWRD1_NEXTWRD1_INWRD_DIFF:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LOHI_N:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LOHI_NG:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LOLO_N:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LOLO_NG:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LR_MINNEXT_KBASELN:continuous.
ENERGY_PREVWRD1_NEXTWRD1_LR_MIN_KBASELN:continuous.
ENERGY_PREVWRD2_NEXTWRD2_DIFF_BEGBEG:continuous.
ENERGY_PREVWRD2_NEXTWRD2_DIFF_MINNEXT_KBASELN:continuous.
ENERGY_PREVWRD2_NEXTWRD2_DIFF_MIN_KBASELN:continuous.
ENERGY_PREVWRD2_NEXTWRD2_HIHI_N:continuous.
ENERGY_PREVWRD2_NEXTWRD2_HIHI_NG:continuous.
ENERGY_PREVWRD2_NEXTWRD2_HILO_N:continuous.
ENERGY_PREVWRD2_NEXTWRD2_HILO_NG:continuous.
ENERGY_PREVWRD2_NEXTWRD2_INWRD_DIFF:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LOHI_N:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LOHI_NG:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LOLO_N:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LOLO_NG:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LR_MINNEXT_KBASELN:continuous.
ENERGY_PREVWRD2_NEXTWRD2_LR_MIN_KBASELN:continuous.
ENERGY_SLOPE_DIFF:continuous.
F0K_DIFF_ENDBEG:continuous.
F0K_DIFF_FIRSTNEXT_KBASELN:continuous.
F0K_DIFF_LAST_KBASELN:continuous.
F0K_FIRSTNEXT_SLOPE:continuous.
F0K_FIRSTNEXT_SLOPE_N:continuous.
F0K_LAST_SLOPE:continuous.
F0K_LAST_SLOPE_N:continuous.
F0K_LR_FIRSTNEXT_KBASELN:continuous.
F0K_LR_LAST_KBASELN:continuous.
F0K_PATTERN_BOUNDARY:text.
F0K_PREVWIN20_NEXTWIN20_DIFF_BEGBEG:continuous.
F0K_PREVWIN20_NEXTWIN20_DIFF_MINNEXT_KBASELN:continuous.
F0K_PREVWIN20_NEXTWIN20_DIFF_MIN_KBASELN:continuous.
F0K_PREVWIN20_NEXTWIN20_HIHI_N:continuous.
F0K_PREVWIN20_NEXTWIN20_HIHI_NG:continuous.
F0K_PREVWIN20_NEXTWIN20_HILO_N:continuous.
F0K_PREVWIN20_NEXTWIN20_HILO_NG:continuous.
F0K_PREVWIN20_NEXTWIN20_INWRD_DIFF:continuous.
F0K_PREVWIN20_NEXTWIN20_LOHI_N:continuous.
F0K_PREVWIN20_NEXTWIN20_LOHI_NG:continuous.
F0K_PREVWIN20_NEXTWIN20_LOLO_N:continuous.
F0K_PREVWIN20_NEXTWIN20_LOLO_NG:continuous.
F0K_PREVWIN20_NEXTWIN20_LR_MINNEXT_KBASELN:continuous.
F0K_PREVWIN20_NEXTWIN20_LR_MIN_KBASELN:continuous.
F0K_PREVWIN50_NEXTWIN50_DIFF_BEGBEG:continuous.
F0K_PREVWIN50_NEXTWIN50_DIFF_MINNEXT_KBASELN:continuous.
F0K_PREVWIN50_NEXTWIN50_DIFF_MIN_KBASELN:continuous.
F0K_PREVWIN50_NEXTWIN50_HIHI_N:continuous.
F0K_PREVWIN50_NEXTWIN50_HIHI_NG:continuous.
F0K_PREVWIN50_NEXTWIN50_HILO_N:continuous.

F0K_PREVWIN50_NEXTWIN50_HILO_NG:continuous.
 F0K_PREVWIN50_NEXTWIN50_INWRD_DIFF:continuous.
 F0K_PREVWIN50_NEXTWIN50_LOHI_N:continuous.
 F0K_PREVWIN50_NEXTWIN50_LOHI_NG:continuous.
 F0K_PREVWIN50_NEXTWIN50_LOLO_N:continuous.
 F0K_PREVWIN50_NEXTWIN50_LOLO_NG:continuous.
 F0K_PREVWIN50_NEXTWIN50_LR_MINNEXT_KBASELN:continuous.
 F0K_PREVWIN50_NEXTWIN50_LR_MIN_KBASELN:continuous.
 F0K_PREVWRD1_NEXTWRD1_DIFF_BEGBEG:continuous.
 F0K_PREVWRD1_NEXTWRD1_DIFF_MINNEXT_KBASELN:continuous.
 F0K_PREVWRD1_NEXTWRD1_DIFF_MIN_KBASELN:continuous.
 F0K_PREVWRD1_NEXTWRD1_HIHI_N:continuous.
 F0K_PREVWRD1_NEXTWRD1_HIHI_NG:continuous.
 F0K_PREVWRD1_NEXTWRD1_HILO_N:continuous.
 F0K_PREVWRD1_NEXTWRD1_HILO_NG:continuous.
 F0K_PREVWRD1_NEXTWRD1_INWRD_DIFF:continuous.
 F0K_PREVWRD1_NEXTWRD1_LOHI_N:continuous.
 F0K_PREVWRD1_NEXTWRD1_LOHI_NG:continuous.
 F0K_PREVWRD1_NEXTWRD1_LOLO_N:continuous.
 F0K_PREVWRD1_NEXTWRD1_LOLO_NG:continuous.
 F0K_PREVWRD1_NEXTWRD1_LR_MINNEXT_KBASELN:continuous.
 F0K_PREVWRD1_NEXTWRD1_LR_MIN_KBASELN:continuous.
 F0K_PREVWRD2_NEXTWRD2_DIFF_BEGBEG:continuous.
 F0K_PREVWRD2_NEXTWRD2_DIFF_MINNEXT_KBASELN:continuous.
 F0K_PREVWRD2_NEXTWRD2_DIFF_MIN_KBASELN:continuous.
 F0K_PREVWRD2_NEXTWRD2_HIHI_N:continuous.
 F0K_PREVWRD2_NEXTWRD2_HIHI_NG:continuous.
 F0K_PREVWRD2_NEXTWRD2_HILO_N:continuous.
 F0K_PREVWRD2_NEXTWRD2_HILO_NG:continuous.
 F0K_PREVWRD2_NEXTWRD2_INWRD_DIFF:continuous.
 F0K_PREVWRD2_NEXTWRD2_LOHI_N:continuous.
 F0K_PREVWRD2_NEXTWRD2_LOHI_NG:continuous.
 F0K_PREVWRD2_NEXTWRD2_LOLO_N:continuous.
 F0K_PREVWRD2_NEXTWRD2_LOLO_NG:continuous.
 F0K_PREVWRD2_NEXTWRD2_LR_MINNEXT_KBASELN:continuous.
 F0K_PREVWRD2_NEXTWRD2_LR_MIN_KBASELN:continuous.
 F0K_SLOPE_DIFF:continuous.
 PAU_DUR:continuous.
 PAU_DUR_2PREV:continuous.
 PAU_DUR_NEXT:continuous.
 PAU_DUR_PREV:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_DIFF_MEAN_KBASELN:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_DIFF_MNMN_N:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_DIFF_MNMN_NG:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_LR_MEAN_KBASELN:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_ZRANGE_MEANNEXT_KBASELN:continuous.
 PREVWIN20_ENERGY_NEXTWIN20_ZRANGE_MEAN_KBASELN:continuous.
 PREVWIN20_F0K_NEXTWIN20_DIFF_MEAN_KBASELN:continuous.
 PREVWIN20_F0K_NEXTWIN20_DIFF_MNMN_N:continuous.
 PREVWIN20_F0K_NEXTWIN20_DIFF_MNMN_NG:continuous.
 PREVWIN20_F0K_NEXTWIN20_LR_MEAN_KBASELN:continuous.
 PREVWIN20_F0K_NEXTWIN20_ZRANGE_MEANNEXT_KBASELN:continuous.
 PREVWIN20_F0K_NEXTWIN20_ZRANGE_MEAN_KBASELN:continuous.
 PREVWIN50_ENERGY_NEXTWIN50_DIFF_MEAN_KBASELN:continuous.
 PREVWIN50_ENERGY_NEXTWIN50_DIFF_MNMN_N:continuous.
 PREVWIN50_ENERGY_NEXTWIN50_DIFF_MNMN_NG:continuous.
 PREVWIN50_ENERGY_NEXTWIN50_LR_MEAN_KBASELN:continuous.
 PREVWIN50_ENERGY_NEXTWIN50_ZRANGE_MEANNEXT_KBASELN:continuous.

PREVWIN50_ENERGY_NEXTWIN50_ZRANGE_MEAN_KBASELN:continuous.
PREVWIN50_F0K_NEXTWIN50_DIFF_MEAN_KBASELN:continuous.
PREVWIN50_F0K_NEXTWIN50_DIFF_MNMN_N:continuous.
PREVWIN50_F0K_NEXTWIN50_DIFF_MNMN_NG:continuous.
PREVWIN50_F0K_NEXTWIN50_LR_MEAN_KBASELN:continuous.
PREVWIN50_F0K_NEXTWIN50_ZRANGE_MEANNEXT_KBASELN:continuous.
PREVWIN50_F0K_NEXTWIN50_ZRANGE_MEAN_KBASELN:continuous.
PREVWRD1_ENERGY_NEXTWRD1_DIFF_MEAN_KBASELN:continuous.
PREVWRD1_ENERGY_NEXTWRD1_DIFF_MNMN_N:continuous.
PREVWRD1_ENERGY_NEXTWRD1_DIFF_MNMN_NG:continuous.
PREVWRD1_ENERGY_NEXTWRD1_LR_MEAN_KBASELN:continuous.
PREVWRD1_ENERGY_NEXTWRD1_ZRANGE_MEANNEXT_KBASELN:continuous.
PREVWRD1_ENERGY_NEXTWRD1_ZRANGE_MEAN_KBASELN:continuous.
PREVWRD1_F0K_NEXTWRD1_DIFF_MEAN_KBASELN:continuous.
PREVWRD1_F0K_NEXTWRD1_DIFF_MNMN_N:continuous.
PREVWRD1_F0K_NEXTWRD1_DIFF_MNMN_NG:continuous.
PREVWRD1_F0K_NEXTWRD1_LR_MEAN_KBASELN:continuous.
PREVWRD1_F0K_NEXTWRD1_ZRANGE_MEANNEXT_KBASELN:continuous.
PREVWRD1_F0K_NEXTWRD1_ZRANGE_MEAN_KBASELN:continuous.
PREVWRD2_ENERGY_NEXTWRD2_DIFF_MEAN_KBASELN:continuous.
PREVWRD2_ENERGY_NEXTWRD2_DIFF_MNMN_N:continuous.
PREVWRD2_ENERGY_NEXTWRD2_DIFF_MNMN_NG:continuous.
PREVWRD2_ENERGY_NEXTWRD2_LR_MEAN_KBASELN:continuous.
PREVWRD2_ENERGY_NEXTWRD2_ZRANGE_MEANNEXT_KBASELN:continuous.
PREVWRD2_ENERGY_NEXTWRD2_ZRANGE_MEAN_KBASELN:continuous.
PREVWRD2_F0K_NEXTWRD2_DIFF_MEAN_KBASELN:continuous.
PREVWRD2_F0K_NEXTWRD2_DIFF_MNMN_N:continuous.
PREVWRD2_F0K_NEXTWRD2_DIFF_MNMN_NG:continuous.
PREVWRD2_F0K_NEXTWRD2_LR_MEAN_KBASELN:continuous.
PREVWRD2_F0K_NEXTWRD2_ZRANGE_MEANNEXT_KBASELN:continuous.
PREVWRD2_F0K_NEXTWRD2_ZRANGE_MEAN_KBASELN:continuous.
TURN_CNT:continuous.
TURN_F:text.
TURN_TIME:continuous.
TURN_TIME_N:continuous.
VOW_DUR_N_LAST:continuous.
VOW_DUR_N_MAX:continuous.
VOW_DUR_Z_LAST:continuous.
VOW_DUR_Z_MAX:continuous.
WRD_End:ignore.
WRD_Start:ignore.
pause:continuous.

Örnek cümleye ilişkin hesaplanmış bürünsel (prosodic) özellikler

Örnek Cümle: Burası Amerikanın sesi Türkçe yayın bölümü Washington.

(burasI amerikanIn sesi tUrkCe yayIn bOIUmU \$washington.)

burasI:

-0.00892,4293.74783,4204.35477,-381.34411,-0.08304,-461.02584, -0.09269, 1.63197, 1.62305, f+f,
0.03280, 1738.17708, 1399.83040, 0.06466, 0.00991, 0.44699, 0.01107, 0.04172, -0.47053, 0.00848,
-0.08820, 0.00947, 1.25353, 1.16533, 1.17330, 73.10857, 192.29744, 0.06466, 0.00991, 1.46702,
0.01610, -1.16438, -1.17756, 0.00659, 0.22479, 0.01071, 0.23350, 0.45830, -1.17330, 73.19507,
192.38394, 0.06466, 0.00992, 1.46702, 0.01611, -1.16438, -1.17756, 0.00660, 0.22479, 0.01072,
0.23387, 0.45867, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 79.68173, 0.01295, 55.78308, 60.57901, -4.79593,
-0.03028, -0.41836, -0.00256, 0.18857, 0.20152, f+f,-0.01901, 22.07138, 45.98395, 0.01295, 0.00979,

0.11660, 0.01028, -0.03196, -0.02748, 0.00961, 0.07616, 0.01009, 0.08456, 0.16072, -0.01188, 20.20615, 45.98395, 0.01295, 0.00979, 0.12314, 0.01031, -0.02483, -0.02748, 0.00961, 0.08270, 0.01012, 0.07801, 0.16072, -0.01188, 20.29266, 46.07045, 0.01295, 0.00980, 0.12314, 0.01032, -0.02483, -0.02748, 0.00962, 0.08270, 0.01013, 0.07838, 0.16108, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, -4.37757, 0.0, ?, 0.0, ?, 2744.20810, -0.05197, 0.00959, 1.44288, 30.25072, 26.72609, 48.89759, 0.05698, 0.01000, 0.16915, 0.29469, 0.47621, 2520.02882, 0.04375, 0.00986, 1.40726, 22.09469, 24.54278, 49.48926, 0.06331, 0.01003, 0.17084, 0.28092, 0.48198, 2520.11532, 0.04498, 0.00987, 1.40763, 22.04873, 24.56432, 49.57576, 0.06329, 0.01003, 0.17120, 0.28208, 0.48322, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 0, false, 33.0, 0.02815, 0.58263, 1.09371, -0.54445, 0.15298, 1844, 1812, 0.00, n.

amerikanIn:

0.04588, 1722.98636, 1926.42154, -86.12283, -0.04728, -86.12283, -0.04514, 1.25028, 1.29616, f+f, -0.40052, 978.33858, 623.21327, 0.10421, 0.01003, 0.52549, 0.01142, -0.44641, -0.59424, 0.00806, -0.17296, 0.00918, 1.02235, 0.84939, 0.38595, 136.77133, 73.10857, 0.19224, 0.01028, 1.26813, 0.01493, 0.34006, -1.21012, 0.00632, -0.13422, 0.00918, 0.36773, 0.23350, 0.38169, 136.85783, 73.19507, 0.19224, 0.01028, 1.26813, 0.01494, 0.33580, -1.21012, 0.00633, -0.13422, 0.00919, 0.36809, 0.23387, -0.79161, 136.77133, 73.10857, 0.11076, 0.01003, 1.33280, 0.01519, -0.83749, -1.35625, 0.00607, -0.13422, 0.00918, 0.36773, 0.23350, 0.0, 1.67734E-4, 29.21933, 29.27025, -0.05091, -3.86300E-4, -0.05091, -3.86150E-4, 0.10883, 0.10900, f+f, 0.00267, 28.98007, 29.18375, -0.04606, 0.00953, 0.00317, 0.00975, 0.00250, -0.04857, 0.00952, 6.71326E-4, 0.00974, 0.10796, 0.10864, 0.06637, 28.98007, 20.20615, 0.01763, 0.00981, 0.06688, 0.01004, 0.06621, -0.07919, 0.00938, -0.02995, 0.00960, 0.10796, 0.07801, 0.07973, 29.06657, 20.29266, 0.03099, 0.00988, 0.08023, 0.01011, 0.07956, -0.07919, 0.00939, -0.02995, 0.00960, 0.10833, 0.07838, 0.06785, 28.98007, 20.20615, 0.02031, 0.00982, 0.09319, 0.01016, 0.06768, -0.10282, 0.00928, -0.02995, 0.00960, 0.10796, 0.07801, 0.0, 0.0, ?, 0.0, 0.0, 2124.79121, 0.11220, 0.01007, 1.33633, 15.75414, 20.69353, 29.56564, -0.01928, 0.00965, 0.10989, 0.34642, 0.28794, 2221.31431, 0.21661, 0.01040, 1.35475, 12.74495, 21.63358, 28.18254, -0.02735, 0.00961, 0.10533, 0.35732, 0.27447, 2262.03507, 0.19721, 0.01035, 1.36264, 13.63630, 22.04873, 28.94012, -0.02137, 0.00964, 0.10791, 0.34678, 0.28208, 2363.33724, 0.15488, 0.01020, 1.38051, 15.81232, 23.01675, 34.38102, -0.02072, 0.00964, 0.12542, 0.40007, 0.33483, 0, false, 84.0, 0.07167, 0.58263, 0.73213, -0.54445, -0.54260, 1895, 1845, 0.00, n.

sesi:

-0.01563, 145.63447, 136.85783, -363.95993, ?, 140.71638, ?, 0.38373, 0.36809, f+f, 0.79075, 145.54797, 136.77133, 0.05777, 0.00990, 1.06026, 0.01405, 0.80639, -1.01812, 0.00682, -0.01563, 0.00967, 0.38336, 0.36773, 1.07153, 145.54797, 136.77133, -0.04191, 0.00962, 1.16448, 0.01447, 1.08717, -1.22203, 0.00643, -0.01563, 0.00967, 0.38336, 0.36773, 0.86655, 145.63447, 136.85783, -0.14613, 0.00935, 1.06026, 0.01406, 0.88218, -1.22203, 0.00643, -0.01563, 0.00968, 0.38373, 0.36809, 1.24824, 145.54797, 73.10857, 0.04610, 0.00986, 1.25250, 0.01483, 1.26387, -1.35625, 0.00607, -0.14986, 0.00912, 0.38336, 0.23350, -223.24354, -0.02562, 43.21121, 34.85852, -2.01842, ?, -2.00331, ?, 0.15265, 0.12703, f+f, 0.00492, 35.05103, 34.77202, 0.00492, 0.00976, 0.02966, 0.00987, 0.03054, -0.02562, 0.00962, -8.80665E-4, 0.00973, 0.12754, 0.12666, -0.04113, 34.65797, 28.98007, -0.02363, 0.00963, 0.03090, 0.00987, -0.01551, -0.07287, 0.00941, -0.01833, 0.00965, 0.12630, 0.10796, -0.04381, 35.13753, 29.06657, -0.02363, 0.00964, 0.02966, 0.00988, -0.01819, -0.07287, 0.00942, -0.01957, 0.00965, 0.12791, 0.10833, 0.03591, 20.87535, 20.20615, 0.00736, 0.00977, 0.10783, 0.01024, 0.06154, -0.10282, 0.00928, -0.00235, 0.00972, 0.08037, 0.07801, -0.01511, 0.0, 0.0, 0.0, 0.0, 1460.88596, 0.17564, 0.01030, 1.18263, 9.16223, 14.22770, 39.78030, 0.00211, 0.00974, 0.14220, 0.38068, 0.38742, 1728.17472, -0.02279, 0.00967, 1.25117, 17.79173, 16.83084, 31.84453, -0.03389, 0.00958, 0.11731, 0.41649, 0.31013, 1398.98237, -0.10045, 0.00944, 1.16543, 17.44534, 13.63630, 35.57726, -0.02649, 0.00962, 0.12929, 0.43151, 0.34678, 1916.72749, 0.00488, 0.00975, 1.29374, 18.44706, 18.66718, 30.15501, -0.01113, 0.00968, 0.11182, 0.32726, 0.29368, 0, false, 118.0, 0.10068, 0.52254, 0.87856, -0.72084, -0.24598, 1929, 1896, 0.00, n.

tUrkCe:

-8.34687E-4, 2619.31522, 2614.08890, -113.36779, -0.04847, -113.36779, -0.04623, 1.42375, 1.42292, f+f, 0.03023, 1422.84993, 1247.57637, 0.16637, 0.01021, 0.41782, 0.01101, 0.03106, -0.30445, 0.00887, -0.05300, 0.00957, 1.17193, 1.11893, -0.34934, 439.88899, 136.77133, 0.16521, 0.01020, 0.86679, 0.01282, -0.34850, -1.05682, 0.00674, -0.35523, 0.00847, 0.72297, 0.36773, -1.04002, 1422.93643, 145.63447, 0.16521, 0.01021, 0.41782, 0.01102, -1.03918, -1.04118, 0.00679, -0.78857, 0.00733, 1.17230, 0.38373, -0.17347, 439.88899, 136.77133, 0.16521, 0.01020, 0.86679, 0.01282, -0.17263, -1.05682, 0.00674, -0.35523, 0.00847, 0.72297, 0.36773, 0.0, 0.00352, 38.19012, 39.33868, -1.14855, -0.00815, -1.14855, -0.00809, 0.13743, 0.14096, f+f, 0.04377, 20.87535, 39.25218, 0.04377, 0.00993, 0.10046, 0.01020, 0.04024, 0.00352, 0.00975, 0.06021, 0.01001, 0.08037, 0.14059, 0.01521, 10.34802, 35.05103, 0.04377, 0.00993, 0.13914, 0.01039, 0.01169, -0.00952, 0.00969, 0.08584, 0.01014, 0.04170, 0.12754, 0.01521, 20.96185, 35.13753, 0.04377, 0.00994, 0.10046, 0.01021, 0.01169, -0.00952, 0.00970, 0.04717, 0.00996, 0.08074, 0.12791, -0.02859, 10.34802, 28.98007, 0.01287, 0.00979, 0.13914, 0.01039, -0.03212, -0.06000, 0.00947, 0.06626, 0.01005, 0.04170, 0.10796, 0.0, 0.0, 0.0, 1.0, 0.0, 2883.13084, 0.15824, 0.01020, 1.46358, 19.19938, 28.07907, 46.14348, 0.05621, 0.00999, 0.16118, 0.27341, 0.44939, 1688.24421, -0.02834, 0.00965, 1.24159, 17.61818, 16.44196, 44.18359, 0.05574, 0.00999, 0.15542, 0.25799, 0.43030, 1789.76146, -0.04999, 0.00960, 1.26588, 19.69591, 17.44534, 44.27009, 0.04874, 0.00997, 0.15579, 0.27953, 0.43151, 1623.59385, -0.06182, 0.00955, 1.22562, 18.38428, 15.81232, 41.07901, 0.04134, 0.00992, 0.14615, 0.27291, 0.40007, 0, false, 164.0, 0.13993, 2.19641, 2.19641, 2.42360, 2.42360, 1975, 1930, 0.00, n.

yayIn:

0.01214, 1666.67801, 1716.84195, -109.06880, -0.07255, -109.06880, -0.06336, 1.23667, 1.24882, f+f, 0.01287, 439.88899, 1656.36165, 0.00159, 0.00974, 0.70158, 0.01223, 7.28331E-4, -0.18916, 0.00920, 0.51082, 0.01155, 0.72297, 1.23379, 0.30362, 439.88899, 1422.84993, 0.07910, 0.00996, 0.81696, 0.01264, 0.29147, -0.28888, 0.00892, 0.44896, 0.01133, 0.72297, 1.17193, 0.18707, 439.97549, 1422.93643, 0.00159, 0.00975, 0.70158, 0.01224, 0.17493, -0.25101, 0.00903, 0.44896, 0.01134, 0.72333, 1.17230, -0.85294, 378.13393, 145.54797, -0.04838, 0.00960, 0.91927, 0.01307, -0.86508, -1.25477, 0.00639, -0.28712, 0.00869, 0.67049, 0.38336, 0.0, 0.01550, 29.66598, 34.47385, -2.40393, -0.01817, 0.61418, 0.00448, 0.11030, 0.12581, r+f, -0.02313, 10.34802, 22.71790, 0.00981, 0.00978, 0.08374, 0.01013, -0.03864, -0.02883, 0.00960, 0.04510, 0.00995, 0.04170, 0.08680, 0.06768, 10.34802, 20.87535, -0.01172, 0.00968, 0.13592, 0.01038, 0.05218, -0.10897, 0.00925, 0.03867, 0.00992, 0.04170, 0.08037, 0.02712, 10.43452, 20.96185, -0.03090, 0.00960, 0.09536, 0.01020, 0.01161, -0.08759, 0.00935, 0.03867, 0.00993, 0.04206, 0.08074, 0.04234, 10.34802, 20.87535, -0.07385, 0.00942, 0.13914, 0.01039, 0.02683, -0.17432, 0.00898, 0.03867, 0.00992, 0.04170, 0.08037, -3.01811, 1.0, 0.0, 0.0, 0.0, 2043.89004, 0.09942, 0.01003, 1.32026, 15.62783, 19.90562, 28.55262, 0.02585, 0.00985, 0.10655, 0.20421, 0.27807, 2260.17545, 0.07884, 0.00997, 1.36195, 18.19149, 22.01205, 32.66616, -0.00135, 0.00973, 0.11996, 0.32225, 0.31813, 2020.65324, 0.06583, 0.00994, 1.31588, 16.78477, 19.69591, 28.67801, 0.00482, 0.00976, 0.10705, 0.26539, 0.27953, 1894.12601, -0.01869, 0.00968, 1.28885, 19.30240, 18.44706, 33.60325, -0.02276, 0.00963, 0.12295, 0.39869, 0.32726, 0, false, 202.0, 0.17235, 0.58263, 0.58263, -0.54445, -0.54445, 2013, 1976, 0.01, n.

bOlUmU:

0.01483, 2437.08450, 2525.33646, -108.92568, -0.04113, 225.95986, 0.08195, 1.39366, 1.40849, r+f, -0.33331, 1409.26790, 1076.17281, 0.01802, 0.00979, 0.25490, 0.01051, -0.34815, -0.34495, 0.00875, -0.10807, 0.00940, 1.16805, 1.05997, -0.14411, 378.13393, 439.88899, -0.21359, 0.00916, 0.75406, 0.01247, -0.15894, -0.91517, 0.00729, 0.05247, 0.00992, 0.67049, 0.72297, -0.15698, 378.22043, 439.97549, -0.21519, 0.00917, 0.75246, 0.01248, -0.17181, -0.91517, 0.00730, 0.05247, 0.00993, 0.67086, 0.72333, 0.03009, 255.72878, 439.88899, -0.21359, 0.00916, 0.88165, 0.01310,

0.01525, -0.91517, 0.00729, 0.18007, 0.01042, 0.54289, 0.72297, -334.88555, -0.00544, 50.48123, 48.57425, 1.90698, 0.01245, 1.90698, 0.01261, 0.17378, 0.16834, r+r, -0.12445, 24.41739, 12.25500, -0.02137, 0.00964, 0.07532, 0.01008, -0.11900, -0.14038, 0.00911, -0.04368, 0.00953, 0.09265, 0.04896, -0.08661, 24.41739, 10.34802, -0.02137, 0.00964, 0.07532, 0.01008, -0.08116, -0.14765, 0.00908, -0.05095, 0.00950, 0.09265, 0.04170, -0.06347, 24.50389, 10.43452, -0.08672, 0.00937, 0.07532, 0.01009, -0.05803, -0.21300, 0.00883, -0.05095, 0.00951, 0.09301, 0.04206, -0.03635, 6.70725, 10.34802, -0.08672, 0.00936, 0.14049, 0.01041, -0.03090, -0.21300, 0.00882, 0.01421, 0.00980, 0.02748, 0.04170, 0.0, 0.0, 0.0, 0.0, 1.0, 2035.65096, 0.00120, 0.00974, 1.31859, 19.76777, 19.82538, 30.37138, -0.03688, 0.00957, 0.11253, 0.41066, 0.29578, 1849.16986, -0.02247, 0.00967, 1.27896, 19.01891, 18.00923, 27.80274, -0.04009, 0.00955, 0.10406, 0.39368, 0.27077, 1721.99186, -0.08005, 0.00951, 1.25004, 20.38485, 16.78477, 27.22792, -0.06886, 0.00943, 0.10222, 0.48284, 0.26539, 1887.67983, 0.01715, 0.00978, 1.28744, 17.63354, 18.38428, 28.02297, -0.06520, 0.00944, 0.10480, 0.47912, 0.27291, 0, false, 232.0, 0.19795, 0.53289, 0.86329, -0.84228, -0.31815, 2043, 2015, 0.00, n.

\$washington:

0.09481, 1101.18423, 1394.89188, -162.86340, -0.13117, -162.86340, -0.11596, 1.06943, 1.16424, f+f, -0.01634, 331.94392, 1056.62984, 0.15611, 0.01018, 0.93155, 0.01317, -0.11116, -0.34936, 0.00874, 0.42608, 0.01131, 0.62663, 1.05271, 0.16910, 247.22035, 378.13393, 0.16780, 0.01020, 1.10568, 0.01397, 0.07428, -0.79984, 0.00750, 0.13803, 0.01026, 0.53246, 0.67049, 0.32423, 255.81528, 378.22043, 0.16780, 0.01021, 1.09525, 0.01392, 0.22941, -0.79984, 0.00750, 0.12759, 0.01023, 0.54326, 0.67086, 0.16724, 247.22035, 378.13393, -0.08046, 0.00952, 1.10568, 0.01397, 0.07243, -1.04811, 0.00700, 0.13803, 0.01026, 0.53246, 0.67049, 0.0, 0.01608, 47.91119, 53.59053, -5.67934, -0.03773, -5.67934, -0.03636, 0.16643, 0.18251, f+f, -0.01685, 6.70725, 42.09436, 0.08572, 0.01012, 0.22722, 0.01082, -0.03294, -0.01977, 0.00965, 0.12172, 0.01032, 0.02748, 0.14920, -0.00839, 6.70725, 28.54181, 0.01403, 0.00979, 0.22722, 0.01082, -0.02448, -0.13414, 0.00915, 0.07903, 0.01011, 0.02748, 0.10652, 0.00735, 6.79375, 24.50389, 0.01403, 0.00980, 0.22722, 0.01083, -0.00873, -0.14801, 0.00910, 0.06517, 0.01005, 0.02784, 0.09301, -0.05612, 6.70725, 10.34802, 0.01403, 0.00979, 0.22722, 0.01082, -0.07221, -0.19896, 0.00887, 0.01421, 0.00980, 0.02748, 0.04170, 0.0, 0.0, 1.0, 6.0, 0.0, 2385.52552, 0.37067, 0.01093, 1.38440, 9.32109, 23.23284, 62.75154, 0.11259, 0.01025, 0.20713, 0.24318, 0.61114, 2107.50961, 0.26194, 0.01056, 1.33294, 10.77625, 20.52522, 54.39150, 0.01578, 0.00980, 0.18461, 0.47513, 0.52972, 2091.33319, 0.19823, 0.01036, 1.33010, 12.54774, 20.38485, 49.53611, 0.00189, 0.00975, 0.17109, 0.47637, 0.48284, 1981.95181, 0.08676, 0.01000, 1.30754, 15.62580, 19.30240, 40.93727, -0.01520, 0.00967, 0.14572, 0.44851, 0.39869, 0, false, 301.0, 0.25682, 0.98743, 0.98743, -0.02723, -0.02723, 2112, 2044, 0.00, s.

Biçimbilgisel, sözcüksel ve bürünsel özelliklerin çıkarılmasından sonra bu özelliklerin tüm kombinasyonları (örneğin; sözcüksel, bürünsel, biçimbilgisel, sözcüksel+bürünsel, sözcüksel+bürünsel+biçimbilgisel vb.) Boostexter, Conditional Random Fields (CRF) gibi farklı sınıflandırıcı yapıları ile özellikle cümle bölütleme alanında kullanılmışlardır. Her durumda kullanılan sistemlerin performans değerlendirmeleri F-measure ve NIST ölçütleri ile değerlendirilmiştir. Özellikle cümle bölütleme amacı ile geliştirilmiş bu özelliklerin farklı kombinasyonları kullanılarak, daha az etiketlenmiş (cümle bölütlemesi yapılmış) veri ile çok büyük miktarlarda etiketlenmemiş (cümle bölütlemesi yapılmamış) verinin yüksek doğruluk oranları ile etiketlenmesi (cümle bölütlemesi) hedeflenmiştir.

Bürünsel özelliklerin çıkarılmasında Algemy dışında Praat gibi bürünsel özellikleri hesaplamak için kullanılan açık kaynak kodlu başka bir yazılım üzerinde de denemeler yapılmıştır.

3.3.4.2. Praat ile Brnsel zelliklerin ıkarılması

Brnsel zellikleri ıkarmak amacı ile kullandığımız SRI nı Algemy yazılımının yanısıra açık kaynak kodlu bir yazılım olan Praat (Boersma ve diğeri, 2009) ile de denemeler yapılmıştır. Bu amaçla zellikle Purdue University de gerekleřtirilen Prosodic Feature Extraction Tool (Huang ve diğeri, 2006) dan yararlanılmıřtır. Bu bađlamda Algemy ile ıkarılan brnsel zelliklerin byk bir kısmı Praat yazılımı ile elde edilmiřtir. Esasen bu yazılım SRI yazılımından esinlenerek hazırlanmıřtır.

Praat ile brnsel zelliklerin ıkarılması iin wav formatında hazırlanmıř bir audio veri ile bu veriye iliřkin kelime ve phone hizalarına gereksinim vardır.

Giriř olarak audio veri ve zamanda hizalanmıř kelime ve phone lar verildiđinde yazılım ncelikle bazı temel eleman setlerini (rneđin; ham pitch, biimlendirilmif pitch, sesli/sessiz (voiced/unvoiced) bltlemesi, duraklar, F0 ve enerji ile iliřkili bileřenler) ıkarmaktadır. Daha sonra sre istatistikleri (rneđin; durak sresinin, phone sresinin ve son uyak sresinin ortalama ve varyans deđerleri), F0 a iliřkin istatistikler (rneđin; logaritmik F0 deđerlerinin ortalama deđer ve varyansı) ve enerji ile iliřkili istatistikler hesaplanmaktadır.

Sre, F0 ve enerji bilgisi ve istatistiklerin de yeralması ile her bir kelime sınırı iin brnsel zelliklerin ıkarılması olduka basit bir iřlem haline gelmektedir. Ařađıdaki izelgede elde edilen hangi temel elemanların hangi zelliklerin ıkarılmasında rol oynadıkları grlmektedir. rneđin; kelime hizalaması, sre, F0 ve Enerji zelliklerinin tmnn hesaplanmasında kullanılırken, sesli.sessiz bltlemesi sadece F0 zelliklerinin hesaplanmasında kullanılmaktadır.

Tablo1: Brnsel zelliklerin (Sre, F0 ve Enerji) Hesaplanmasında Kullanılan Temel Elemanlar

	Sre zellikleri	F0 zellikleri	Enerji zellikleri
Kelime	√	√	√
Phone	√	×	×
Sesli harf	√	×	×
Uyak	√	×	×
Sesli/Sessiz	×	√	×
Ham Pitch	×	√	×
Biimlendirilmif Pitch	×	√	×
Pitch Eđimi	×	√	×
Ham Enerji	×	×	√
Biimlendirilmif Enerji	×	×	√
Enerji Eđimi	×	×	√

Audio, Kelime ve Phone Hizalaması

Wav formatındaki ses dosyası ve zorlanmıř hizalamalar (forced alignment) ile elde edilmiř kelime ve phone sre bilgileri Praat yazılımı ile okunmaktadır. Bunun iin Praat nı *TextGrid* formatı yardımı ile her bir dosya iin ayrı bir katman (tier) kullanılmaktadır. Sessiz aralıklar boř bırakılmaldır. Ayrıca phone ların zamanlamaları kelime sınırının bařlangı ve bitiřinde kelime zamanlaması ile hizalanmıř olmalıdır. Trke iin bir Phone seti hazırlanmıř ve bu phone seti kullanılmıřtır.

Sesli phone seti: I1 (ı), O1 (), U1 (), a, e, i, o, u.

Sessiz phone seti: b, c, C1 (), d, f, g, G1 (đ), h, j, k, l, m, n, p, r, s, S1 (ř), t, v, y, z.

rnek olarak; “řok iddia ABD Haitiye iřgal mi ediyor (S1ok ida a b d haitiyi iS1gal mi ediyor)” cmlesi ele alınarak *kelime_textgrid* ve *phone_textgrid* dosyaları ařađıdaki gibi oluřturulmuřtur.

Kelime_textgrid dosyası:

```
File type = "ooTextFile"
Object class = "TextGrid"
xmin = 0
xmax = 32
tiers? <exists>
size = 1
item []:
  item [1]:
    class = "IntervalTier"
    name = ""
    xmin = 0
    xmax = 32
    intervals: size = 12
    intervals [1]:
      xmin = 0
      xmax = 0.160362
      text = ""
    intervals [2]:
      xmin = 0.160362
      xmax = 0.509192
      text = "S1ok"
    intervals [3]:
      xmin = 0.509192
      xmax = 0.967007
      text = "ida"
    intervals [4]:
      xmin = 0.967007
      xmax = 1.250423
      text = ""
    intervals [5]:
      xmin = 1.250423
      xmax = 1.416112
      text = "a"
    intervals [6]:
      xmin = 1.416112
      xmax = 1.642845
      text = "b"
    intervals [7]:
      xmin = 1.642845
      xmax = 1.869578
      text = "d"
    intervals [8]:
      xmin = 1.869578
      xmax = 2.384086
      text = "haitiyi"
    intervals [9]:
      xmin = 2.384086
      xmax = 2.789589
      text = "iS1gal"
    intervals [10]:
      xmin = 2.789589
      xmax = 2.891225
      text = "mi"
    intervals [11]:
      xmin = 2.891225
```

```
xmax = 3.452346
text = "ediyor"
intervals [12]:
  xmin = 3.452346
  xmax = 32
  text = ""
```

Phone_textgrid dosyası:

```
File type = "ooTextFile"
Object class = "TextGrid"

xmin = 0
xmax = 32
tiers? <exists>
size = 1
item []:
  item [1]:
    class = "IntervalTier"
    name = ""
    xmin = 0
    xmax = 32
    intervals: size = 32
    intervals [1]:
      xmin = 0
      xmax = 0.160362000000000
      text = ""
    intervals [2]:
      xmin = 0.160362000000000
      xmax = 0.347072000000000
      text = "S1"
    intervals [3]:
      xmin = 0.347072000000000
      xmax = 0.473646000000000
      text = "o"
    intervals [4]:
      xmin = 0.473646000000000
      xmax = 0.509192000000000
      text = "k"
    intervals [5]:
      xmin = 0.509192000000000
      xmax = 0.604040000000000
      text = "i"
    intervals [6]:
      xmin = 0.604040000000000
      xmax = 0.717745000000000
      text = "d"
    intervals [7]:
      xmin = 0.717745000000000
      xmax = 0.967007000000000
      text = "a"
    intervals [8]:
      xmin = 0.967007000000000
      xmax = 1.250423000000000
      text = ""
    intervals [9]:
```

```
xmin = 1.2504230000000000
xmax = 1.4161120000000000
text = "a"
intervals [10]:
  xmin = 1.4161120000000000
  xmax = 1.6428450000000000
  text = "b"
intervals [11]:
  xmin = 1.6428450000000000
  xmax = 1.8695780000000000
  text = "d"
intervals [12]:
  xmin = 1.8695780000000000
  xmax = 1.9341020000000000
  text = "h"
intervals [13]:
  xmin = 1.9341020000000000
  xmax = 2.0161700000000000
  text = "a"
intervals [14]:
  xmin = 2.0161700000000000
  xmax = 2.1359470000000000
  text = "i"
intervals [15]:
  xmin = 2.1359470000000000
  xmax = 2.2049230000000000
  text = "t"
intervals [16]:
  xmin = 2.2049230000000000
  xmax = 2.2818050000000000
  text = "i"
intervals [17]:
  xmin = 2.2818050000000000
  xmax = 2.3155870000000000
  text = "y"
intervals [18]:
  xmin = 2.3155870000000000
  xmax = 2.3840860000000000
  text = "i"
intervals [19]:
  xmin = 2.3840860000000000
  xmax = 2.4622940000000000
  text = "i"
intervals [20]:
  xmin = 2.4622940000000000
  xmax = 2.5795850000000000
  text = "S1"
intervals [21]:
  xmin = 2.5795850000000000
  xmax = 2.6546010000000000
  text = "g"
intervals [22]:
  xmin = 2.6546010000000000
  xmax = 2.7506940000000000
  text = "a"
intervals [23]:
  xmin = 2.7506940000000000
```

```
xmax = 2.7895890000000000
text = "l"
intervals [24]:
xmin = 2.7895890000000000
xmax = 2.8632090000000000
text = "m"
intervals [25]:
xmin = 2.8632090000000000
xmax = 2.8912250000000000
text = "i"
intervals [26]:
xmin = 2.8912250000000000
xmax = 3.0137840000000000
text = "e"
intervals [27]:
xmin = 3.0137840000000000
xmax = 3.0636790000000000
text = "d"
intervals [28]:
xmin = 3.0636790000000000
xmax = 3.0923540000000000
text = "i"
intervals [29]:
xmin = 3.0923540000000000
xmax = 3.1152950000000000
text = "y"
intervals [30]:
xmin = 3.1384870000000000
xmax = 3.2205040000000000
text = "o"
intervals [31]:
xmin = 3.2205040000000000
xmax = 3.4523460000000000
text = "r"
intervals [32]:
xmin = 3.4523460000000000
xmax = 32
text = ""
```

1) Hesaplanan Brnsel zellik Listesi

A) Temel zellikler

A1) Base zellikler:

WAV: Audio dosyanin dizini

SPK_ID: Geerli dalgaformu iin konumacıyı tanımlayan etiket

SPK_GEN: Konumacının cinsiyeti

A2) Sre zellikleri:

WORD: Sınırdan bir nceki kelime

WORD_START: Sınırdan nce gelen kelimenin balangı zamanı

WORD_END: Sınırdan nce gelen kelimenin biti zamanı

FWORD: Sınırın ardından gelen kelime

FWORD_START: Sınırın ardından gelen kelimenin balangı zamanı

FWORD_END: Sınırın ardından gelen kelimenin biti zamanı

PAUSE_START: Sınır etrafındaki durağın başlangıç zamanı

PAUSE_END: Sınır etrafındaki durağın bitiş zamanı

PAUSE_DUR: Sınır etrafındaki durağın süresi PAUSE_DUR = PAUSE_END-PAUSE_START.

WORD_PHONES: kelime içerisindeki phone lar ve süreleri

phone1:süre1 phone2:süre2 ...

FLAG: Bu özellik sınırdan önceki kelimenin güvenilir bir phone süresine sahip olup olmadığına işaret eder.

LAST_VOWEL: Bir sınırdan önceki son sesli harf

LAST_VOWEL_START: Bir sınırdan önceki son sesli harfin başlangıç zamanı

LAST_VOWEL_END: Bir sınırdan önceki son sesli harfin bitiş zamanı

LAST_VOWEL_DUR: Bir sınırdan önceki son sesli harfin süresi

LAST_RHYME_START: Bir sınırdan önceki son uyağın (kafiye) başlangıç zamanı

LAST_RHYME_END: Bir sınırdan önceki son uyağın (kafiye) bitiş zamanı

$$NORM_LAST_RHYME_DUR = \sum_{\text{kelimedeki her phone için}} \frac{\text{phone süresi} - \text{ortalama phone süresi}}{\text{phone süresinin standart sapması}}$$

PHONES_IN_LAST_RHYME: Son uyaktaki toplam phone sayısı

A3) F0 Özellikleri

Temel F0 özellikleri aşağıda verildiği gibidir.

MIN_F0: Bir sınırdan önceki kelime için minimum ham F0 değeri

MAX_F0: Bir sınırdan önceki kelime için maksimum ham F0 değeri

MEAN_F0: Bir sınırdan önceki kelime için ortalama ham F0 değeri

F0 ile ilgili olarak kullandığımız daha detaylı özelliklere burada değinilmeyecektir.

A4) Enerji Özellikleri

Temel enerji ilişkili özellikler aşağıda verilmektedir.

MIN_ENERGY

MAX_ENERGY

MEAN_ENERGY

MIN_ENERGY_NEXT

MAX_ENERGY_NEXT

MEAN_ENERGY_NEXT

MIN_ENERGY_WIN

MAX_ENERGY_WIN

MEAN_ENERGY_WIN

MIN_ENERGY_NEXT_WIN

MAX_ENERGY_NEXT_WIN

MEAN_ENERGY_NEXT_WIN

MIN_STYLFIT_ENERGY

MAX_STYLFIT_ENERGY

MEAN_STYLFIT_ENERGY

FIRST_STYLFIT_ENERGY

LAST_STYLFIT_ENERGY

MIN_STYLFIT_ENERGY_NEXT

MAX_STYLFIT_ENERGY_NEXT

MEAN_STYLFIT_ENERGY_NEXT

FIRST_STYLFIT_ENERGY_NEXT

LAST_STYLFIT_ENERGY_NEXT

MIN_STYLFIT_ENERGY_WIN

MAX_STYLFIT_ENERGY_WIN

MEAN_STYLFIT_ENERGY_WIN

FIRST_STYLFIT_ENERGY_WIN

LAST_STYLFIT_ENERGY_WIN
MIN_STYLFIT_ENERGY_NEXT_WIN
MAX_STYLFIT_ENERGY_NEXT_WIN
MEAN_STYLFIT_ENERGY_NEXT_WIN
FIRST_STYLFIT_ENERGY_NEXT_WIN
LAST_STYLFIT_ENERGY_NEXT_WIN
ENERGY_PATTERN_WORD
ENERGY_PATTERN_WORD_CALLAPSED
ENERGY_PATTERN_SLOPE
ENERGY_PATTERN_WORD_NEXT
ENERGY_PATTERN_WORD_CALLAPSED_NEXT
ENERGY_PATTERN_SLOPE_NEXT
ENERGY_PATTERN_WORD_WIN
ENERGY_PATTERN_WORD_CALLAPSED_WIN
ENERGY_PATTERN_SLOPE_WIN
ENERGY_PATTERN_WORD_NEXT_WIN
ENERGY_PATTERN_WORD_CALLAPSED_NEXT_WIN
ENERGY_PATTERN_SLOPE_NEXT_WIN
ENERGY_PATTERN_BOUNDARY
ENERGY_SLOPE_DIFF

B) Türetilmiş Özellikler

Türetilmiş özellikler, yukarıda verilen temel özelliklerin ve istatistiklerin kullanılması ile elde edilen özelliklerdir.

B1) Normalize Edilmiş Kelime Süreleri

WORD_DUR
WORD_AV_DUR
NORM_WORD_DUR

B2) Normalize Edilmiş Durak

PAU_DUR_N

B3) Normalize Edilmiş Sesli Harf Süresi

LAST_VOWEL_DUR_Z
LAST_VOWEL_DUR_N
LAST_VOWEL_DUR_ZSP
LAST_VOWEL_DUR_NSP

B4) Normalize Edilmiş Uyak Süresi

LAST_RHYME_DUR_PH
LAST_RHYME_DUR_PH_ND
LAST_RHYME_DUR_PH_NR
LAST_RHYME_NORM_DUR_PH
LAST_RHYME_NORM_DUR_PH_ND
LAST_RHYME_NORM_DUR_PH_NR
LAST_RHYME_DUR_WHOLE_ND
LAST_RHYME_WHOLE_DUR_NR
LAST_RHYME_WHOLE_DUR_Z

B5) Türetilmiş F0 Özellikleri

Konuşmacının F0 karakteristiklerine ilişkin türetilmiş özelliklerdir.

B6) Türetilmiş Enerji Özellikleri

ENERGY_WORD_DIFF_HIHI_N
ENERGY_WORD_DIFF_HILO_N
ENERGY_WORD_DIFF_LOLO_N
ENERGY_WORD_DIFF_LOHI_N
ENERGY_WORD_DIFF_MNMN_N
ENERGY_WORD_DIFF_HIHI_NG
ENERGY_WORD_DIFF_HILO_NG
ENERGY_WORD_DIFF_LOLO_NG
ENERGY_WORD_DIFF_LOHI_NG
ENERGY_WORD_DIFF_MNMN_NG
ENERGY_WIN_DIFF_HIHI_N
ENERGY_WIN_DIFF_HILO_N
ENERGY_WIN_DIFF_LOLO_N
ENERGY_WIN_DIFF_LOHI_N
ENERGY_WIN_DIFF_MNMN_NG
ENERGY_WIN_DIFF_HIHI_NG
ENERGY_WIN_DIFF_HILO_NG
ENERGY_WIN_DIFF_LOLO_NG
ENERGY_WIN_DIFF_LOHI_NG
ENERGY_WIN_DIFF_MNMN_NG
ENERGY_DIFF_LAST_KBASELN
ENERGY_DIFF_MEAN_KBASELN
ENERGY_DIFF_WINMIN_KBASELN
ENERGY_LR_LAST_KBASELN
ENERGY_LR_MEAN_KBASELN
ENERGY_LR_WINMIN_KBASELN
ENERGY_ZRANGE_MEAN_KBASELN
ENERGY_ZRANGE_MEAN_KTOPLN
ENERGY_ZRANGE_MEANNEXT_KBASELN
ENERGY_ZRANGE_MEANNEXT_KTOPLN
ENERGY_DIFF_MEANNEXT_KTOPLN
ENERGY_DIFF_MAXNEXT_KTOPLN
ENERGY_DIFF_WINMAXNEXT_KTOPLN
ENERGY_LR_MEANNEXT_KTOPLN
ENERGY_LR_MAXNEXT_KTOPLN
ENERGY_LR_WINMAXNEXT_KTOPLN
ENERGY_MAXK_MODE_N
ENERGY_MAXK_NEXT_MODE_N
ENERGY_MAXK_MODE_Z
ENERGY_MAXK_NEXT_MODE_Z
ENERGY_WORD_DIFF_BEGBEG
ENERGY_WORD_DIFF_ENDBEG
ENERGY_INWRD_DIFF
ENERGY_LAST_SLOPE
ENERGY_SLOPE_DIFF_N
ENERGY_LAST_SLOPE_N

B7) Ortalama Phone Süresi

AVG_PHONE_DUR_Z
MAX_PHONE_DUR_Z

AVG_PHONE_DUR_N
MAX_PHONE_DUR_N

B8) Konuşmacıya Özel Normalizasyon

AVG_PHONE_DUR_ZSP
MAX_PHONE_DUR_ZSP
AVG_PHONE_DUR_NSP
MAX_PHONE_DUR_NSP

Praat ile Çıkarılan Bürünsel Özelliklerin Tümü

WORDS\$
WAV\$
SPKR_ID\$
GEN\$
WORD_START
WORD_END
FWORDS\$
FWORD_START
FWORD_END
PAUSE_START
PAUSE_END
PAUSE_DUR
WORD_PHONESS\$
FLAGS\$
LAST_VOWEL\$
LAST_VOWEL_START
LAST_VOWEL_END
LAST_VOWEL_DUR
LAST_RHYME_START
LAST_RHYME_END
LAST_RHYME_DUR
NORM_LAST_RHYME_DUR
PHONES_IN_LAST_RHYME
MIN_F0
MIN_STYLFIT_F0
MAX_F0
MAX_STYLFIT_F0
MEAN_F0
MEAN_STYLFIT_F0
FIRST_STYLFIT_F0
LAST_STYLFIT_F0
PATTERN_WORDS\$
PATTERN_WORD_COLLAPSED\$
PATTERN_SLOPE\$
NO_PREVIOUS_SSF
NO_PREVIOUS_VF
NO_FRAMES_LS_WE
NO_SUCCESSOR_SSF
NO_SUCCESSOR_VF
NO_FRAMES_WS_FS
MIN_F0_WIN
MIN_STYLFIT_F0_WIN
MAX_F0_WIN
MAX_STYLFIT_F0_WIN
MEAN_F0_WIN
MEAN_STYLFIT_F0_WIN
FIRST_STYLFIT_F0_WIN

LAST_STYLFIT_F0_WIN
 PATTERN_WORD_WINS
 PATTERN_WORD_COLLAPSED_WINS
 PATTERN_SLOPE_WINS
 NO_PREVIOUS_SSF_WIN
 NO_PREVIOUS_VF_WIN
 NO_FRAMES_LS_WE_WIN
 NO_SUCESSOR_SSF_WIN
 NO_SUCESSOR_VF_WIN
 NO_FRAMES_WS_FS_WIN
 MIN_F0_NEXT
 MIN_STYLFIT_F0_NEXT
 MAX_F0_NEXT
 MAX_STYLFIT_F0_NEXT
 MEAN_F0_NEXT
 MEAN_STYLFIT_F0_NEXT
 FIRST_STYLFIT_F0_NEXT
 LAST_STYLFIT_F0_NEXT
 PATTERN_WORD_NEXT\$
 PATTERN_WORD_COLLAPSED_NEXT\$
 PATTERN_SLOPE_NEXT\$
 NO_PREVIOUS_SSF_NEXT
 NO_PREVIOUS_VF_NEXT
 NO_FRAMES_LS_WE_NEXT
 NO_SUCESSOR_SSF_NEXT
 NO_SUCESSOR_VF_NEXT
 NO_FRAMES_WS_FS_NEXT
 MIN_F0_NEXT_WIN
 MIN_STYLFIT_F0_NEXT_WIN
 MAX_F0_NEXT_WIN
 MAX_STYLFIT_F0_NEXT_WIN
 MEAN_F0_NEXT_WIN
 MEAN_STYLFIT_F0_NEXT_WIN
 FIRST_STYLFIT_F0_NEXT_WIN
 LAST_STYLFIT_F0_NEXT_WIN
 PATTERN_WORD_NEXT_WINS\$
 PATTERN_WORD_COLLAPSED_NEXT_WINS\$
 PATTERN_SLOPE_NEXT_WINS\$
 NO_PREVIOUS_SSF_NEXT_WIN
 NO_PREVIOUS_VF_NEXT_WIN
 NO_FRAMES_LS_WE_NEXT_WIN
 NO_SUCESSOR_SSF_NEXT_WIN
 NO_SUCESSOR_VF_NEXT_WIN
 NO_FRAMES_WS_FS_NEXT_WIN
 PATTERN_BOUNDARY\$
 SLOPE_DIFF
 MIN_ENERGY
 MIN_STYLFIT_ENERGY
 MAX_ENERGY
 MAX_STYLFIT_ENERGY
 MEAN_ENERGY
 MEAN_STYLFIT_ENERGY
 FIRST_STYLFIT_ENERGY
 LAST_STYLFIT_ENERGY
 ENERGY_PATTERN_WORDS\$
 ENERGY_PATTERN_WORD_COLLAPSED\$
 ENERGY_PATTERN_SLOPE\$
 MIN_ENERGY_WIN

MIN_STYLFIT_ENERGY_WIN
 MAX_ENERGY_WIN
 MAX_STYLFIT_ENERGY_WIN
 MEAN_ENERGY_WIN
 MEAN_STYLFIT_ENERGY_WIN
 FIRST_STYLFIT_ENERGY_WIN
 LAST_STYLFIT_ENERGY_WIN
 ENERGY_PATTERN_WORD_WINS
 ENERGY_PATTERN_WORD_COLLAPSED_WINS
 ENERGY_PATTERN_SLOPE_WINS
 MIN_ENERGY_NEXT
 MIN_STYLFIT_ENERGY_NEXT
 MAX_ENERGY_NEXT
 MAX_STYLFIT_ENERGY_NEXT
 MEAN_ENERGY_NEXT
 MEAN_STYLFIT_ENERGY_NEXT
 FIRST_STYLFIT_ENERGY_NEXT
 LAST_STYLFIT_ENERGY_NEXT
 ENERGY_PATTERN_WORD_NEXT\$
 ENERGY_PATTERN_WORD_COLLAPSED_NEXT\$
 ENERGY_PATTERN_SLOPE_NEXT\$
 MIN_ENERGY_NEXT_WIN
 MIN_STYLFIT_ENERGY_NEXT_WIN
 MAX_ENERGY_NEXT_WIN
 MAX_STYLFIT_ENERGY_NEXT_WIN
 MEAN_ENERGY_NEXT_WIN
 MEAN_STYLFIT_ENERGY_NEXT_WIN
 FIRST_STYLFIT_ENERGY_NEXT_WIN
 LAST_STYLFIT_ENERGY_NEXT_WIN
 ENERGY_PATTERN_WORD_NEXT_WINS
 ENERGY_PATTERN_WORD_COLLAPSED_NEXT_WINS
 ENERGY_PATTERN_SLOPE_NEXT_WINS
 ENERGY_PATTERN_BOUNDARY\$
 ENERGY_SLOPE_DIFF
 WORD_DUR
 WORD_AV_DUR
 NORM_WORD_DUR
 PAUSE_DUR_N
 LAST_VOWEL_DUR_Z
 LAST_VOWEL_DUR_N
 LAST_VOWEL_DUR_ZSP
 LAST_VOWEL_DUR_NSP
 LAST_RHYME_DUR_PH
 LAST_RHYME_DUR_PH_ND
 LAST_RHYME_DUR_PH_NR
 LAST_RHYME_NORM_DUR_PH
 LAST_RHYME_NORM_DUR_PH_ND
 LAST_RHYME_NORM_DUR_PH_NR
 LAST_RHYME_DUR_WHOLE_ND
 LAST_RHYME_DUR_WHOLE_NR
 LAST_RHYME_DUR_WHOLE_Z
 F0K_WORD_DIFF_HIHI_N
 F0K_WORD_DIFF_HILO_N
 F0K_WORD_DIFF_LOLO_N
 F0K_WORD_DIFF_LOHI_N
 F0K_WORD_DIFF_MNMN_N
 F0K_WORD_DIFF_HIHI_NG
 F0K_WORD_DIFF_HILO_NG

F0K_WORD_DIFF_LOLO_NG
 F0K_WORD_DIFF_LOHI_NG
 F0K_WORD_DIFF_MNMN_NG
 F0K_WIN_DIFF_HIHI_N
 F0K_WIN_DIFF_HILO_N
 F0K_WIN_DIFF_LOLO_N
 F0K_WIN_DIFF_LOHI_N
 F0K_WIN_DIFF_MNMN_N
 F0K_WIN_DIFF_HIHI_NG
 F0K_WIN_DIFF_HILO_NG
 F0K_WIN_DIFF_LOLO_NG
 F0K_WIN_DIFF_LOHI_NG
 F0K_WIN_DIFF_MNMN_NG
 F0K_DIFF_LAST_KBASELN
 F0K_DIFF_MEAN_KBASELN
 F0K_DIFF_WINMIN_KBASELN
 F0K_LR_LAST_KBASELN
 F0K_LR_MEAN_KBASELN
 F0K_LR_WINMIN_KBASELN
 F0K_ZRANGE_MEAN_KBASELN
 F0K_ZRANGE_MEAN_KTOPLN
 F0K_ZRANGE_MEANNEXT_KBASELN
 F0K_ZRANGE_MEANNEXT_KTOPLN
 F0K_DIFF_MEANNEXT_KTOPLN
 F0K_DIFF_MAXNEXT_KTOPLN
 F0K_DIFF_WINMAXNEXT_KTOPLN
 F0K_LR_MEANNEXT_KTOPLN
 F0K_LR_MAXNEXT_KTOPLN
 F0K_LR_WINMAXNEXT_KTOPLN
 F0K_MAXK_MODE_N
 F0K_MAXK_NEXT_MODE_N
 F0K_MAXK_MODE_Z
 F0K_MAXK_NEXT_MODE_Z
 F0K_WORD_DIFF_BEGBEG
 F0K_WORD_DIFF_ENDBEG
 F0K_INWORD_DIFF
 SLOPE_DIFF_N
 LAST_SLOPE
 LAST_SLOPE_N
 FIRST_SLOPE_NEXT
 ENERGY_WORD_DIFF_HIHI_N
 ENERGY_WORD_DIFF_HILO_N
 ENERGY_WORD_DIFF_LOLO_N
 ENERGY_WORD_DIFF_LOHI_N
 ENERGY_WORD_DIFF_MNMN_N
 ENERGY_WORD_DIFF_HIHI_NG
 ENERGY_WORD_DIFF_HILO_NG
 ENERGY_WORD_DIFF_LOLO_NG
 ENERGY_WORD_DIFF_LOHI_NG
 ENERGY_WORD_DIFF_MNMN_NG
 ENERGY_WIN_DIFF_HIHI_N
 ENERGY_WIN_DIFF_HILO_N
 ENERGY_WIN_DIFF_LOLO_N
 ENERGY_WIN_DIFF_LOHI_N
 ENERGY_WIN_DIFF_MNMN_N
 ENERGY_WIN_DIFF_HIHI_NG
 ENERGY_WIN_DIFF_HILO_NG
 ENERGY_WIN_DIFF_LOLO_NG

ENERGY_WIN_DIFF_LOHI_NG
ENERGY_WIN_DIFF_MNMIN_NG
ENERGY_DIFF_LAST_KBASELN
ENERGY_DIFF_MEAN_KBASELN
ENERGY_DIFF_WINMIN_KBASELN
ENERGY_LR_LAST_KBASELN
ENERGY_LR_MEAN_KBASELN
ENERGY_LR_WINMIN_KBASELN
ENERGY_ZRANGE_MEAN_KBASELN
ENERGY_ZRANGE_MEAN_KTOPLN
ENERGY_ZRANGE_MEANNEXT_KBASELN
ENERGY_ZRANGE_MEANNEXT_KTOPLN
ENERGY_DIFF_MEANNEXT_KTOPLN
ENERGY_DIFF_MAXNEXT_KTOPLN
ENERGY_DIFF_WINMAXNEXT_KTOPLN
ENERGY_LR_MEANNEXT_KTOPLN
ENERGY_LR_MAXNEXT_KTOPLN
ENERGY_LR_WINMAXNEXT_KTOPLN
ENERGY_MAXK_MODE_N
ENERGY_MAXK_NEXT_MODE_N
ENERGY_MAXK_MODE_Z
ENERGY_MAXK_NEXT_MODE_Z
ENERGY_WORD_DIFF_BEGBEG
ENERGY_WORD_DIFF_ENDBEG
ENERGY_INWORD_DIFF
ENERGY_SLOPE_DIFF_N
ENERGY_LAST_SLOPE
ENERGY_LAST_SLOPE_N
ENERGY_FIRST_SLOPE_NEXT
AVG_PHONE_DUR_Z
MAX_PHONE_DUR_Z
AVG_PHONE_DUR_N
MAX_PHONE_DUR_N
AVG_PHONE_DUR_ZSP
MAX_PHONE_DUR_ZSP
AVG_PHONE_DUR_NSP
MAX_PHONE_DUR_NSP
AVG_VOWEL_DUR_Z
MAX_VOWEL_DUR_Z
AVG_VOWEL_DUR_N
MAX_VOWEL_DUR_N
AVG_VOWEL_DUR_ZSP
MAX_VOWEL_DUR_ZSP
AVG_VOWEL_DUR_NSP
MAX_VOWEL_DUR_NSP

S1ok:

../demo/data/demo_H.wav, H, male, 16, 51, ida, 51, 97, 51, 51, 0, S1:19_o:13_k:4, SUSP, ?, ?, ?, ?, ?, ?, ?, ?,
139.02508062164566, 139.02508062164566, 234.83447797558082, 233.0840175423176, 186.9536691565916,
185.8123559345546, 211.67439947397284, 233.0840175423176, Uf, UUf, U, -603.3102437040556, 5, 3, 0, 5,
4, 19, 139.02508062164566, 139.02508062164566, 234.83447797558082, 233.0840175423176,
186.9536691565916, 185.8123559345546, 211.67439947397284, 233.0840175423176, Uf, UUf, U,
-603.3102437040556, 5, 3, 0, 5, 4, 4, 80.67289597938922, 82.32080734983464, 236.2832594821913,
234.63801635039877, 110.95535441478432, 119.45182143158313, 233.8610169463582, 82.32080734983464,
fUfU, fUfU, -305.06865000265634, U, -205.19367383521234, U, 16, 17, 5, 18, 20, 0, 100.19550636412536,
101.49730061769934, 236.2832594821913, 234.63801635039877, 129.96878863224688,
138.11027394597048, 233.8610169463582, 101.49730061769934, f, f, -305.06865000265634, 18, ?, ?, ?, ?,
ff, -298.2415937013992, 41.4744506482582, 41.820994579845305, 62.582385553585134,

61.83891425207503, 50.2595575404426, 50.02837611086514, 42.796935462096926, 45.6739374744422, firff,
frf, -30.498152570363366, 136.83339157860632, 10.533070401230678, -44.58313111365682,
-33.6806918456533, -143.87432693793534, 42.42955097334462, 42.42955097334462, 62.582385553585134,
61.83891425207503, 52.91587141200776, 52.79343318665166, 49.680539218769944, 45.6739374744422, ff,
f, -44.58313111365682, -433.6806918456533, -143.87432693793534, 38.26992889395384,
38.33808017371905, 61.708175598025164, 61.70817559802514, 51.767148308090526, 51.67018371474762,
48.91832397553978, 38.33808017371905, rffrf, rfrf, 304.161234477898, -16.54053909241583,
-192.66495040034152, -39.08628577150598, 320.6246164221373, -115.3129708436157, 46.45482700980172,
48.445112345576675, 60.06328896928366, 58.65148347883252, 53.12676840515796, 53.090663089980715,
48.91832397553978, 48.445112345576675, rff, rf, 304.161234477898, -16.54053909241583,
-192.66495040034152, -39.08628577150598, fr, -448.03556141583334, 35, -3, -11.666666666666666, 0, ?, ?,
?, ?, ?, ?, ?, ?, ?, ?, ?, -0.006644991741483921, 1.0407750799446247, 0.5240304540466085,
-0.5233896176395001, 0.4418242032900518, 0.009731690969036843, 0.01204274008271854,
0.010901194354192336, 0.008809212348661836, 0.010643619481815704, -0.006644991741483921,
0.8313667756503376, 0.3146221497523214, -0.5233896176395001, 0.2966848728085111,
0.009731690969036843, 0.011496888956828863, 0.010407085109044966, 0.008809212348661836,
0.010330147208469885, 130.4520560909072, 83.1803944831442, 36.39311917023525, 0.8202495795321478,
0.5935879260841637, 0.30350495363413166, 0.8104726179526904, 0.18952738204730957,
0.16388520439742194, 0.8361147956025781, -85.81210147123768, 29.374093447577962,
29.374093447577962, -0.5413834577658334, 0.13374739071368635, 0.13374739071368635,
0.5325675070803668, 0.5392124988218508, 0.9377332452068012, 0.9528747156230797,
-0.09967775172383447, -0.003328015349992718, -0.09634973637384171, -1.9496991412989717,
-603.3102437040556, -2.5883810055510197, -305.06865000265634, 0.0021164189785320923,
0.4780891841048592, 0.0869548134672369, -0.3890179516590903, -0.03229053116707268,
0.030995965310442204, 0.03504188969087243, 0.03171882548470798, 0.0280565809402701,
0.03072647537445595, 0.052919978216041254, 0.24410139388462632, -0.13258637647301919,
-0.32376779214160406, -0.005614264927477919, 0.03138271754122069, 0.03292888384770804,
0.029921535307199813, 0.028516578189784966, 0.030936271291648276, 13.395110769247552,
17.749549405670493, 10.150724268149972, 0.3471163439370504, 0.43817887134717287,
0.27343358107628796, 0.5498821121281353, 0.45011788787186463, 0.6007454108123551,
0.3992545891876449, -12.8874696956416, -2.8494778123641495, -5.906169931556775,
-0.22267777804569977, -0.045142248104544085, -0.09594580734205337, 0.36243927898215234,
0.3603228600036202, 0.5824399634117822, 0.5783896720166231, -0.13368555104433266,
-0.06862421046021257, -0.0650613405841201, -2.4125327140221824, -143.87432693793534,
-3.1500311751848242, 304.161234477898, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

ida:

../demo/data/demo_H.wav, H, male, 51, 97, a, 125, 142, 97, 125, 28, i:9_d:11_a:25, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?,
80.67289597938922 82.32080734983464 236.2832594821913, 234.63801635039877, 110.95535441478432,
119.45182143158313, 233.8610169463582, 82.32080734983464 fUfU, fUfU, -305.06865000265634, U,
-205.19367383521234, U, 16, 17, 5, 18, 20, 0, 80.67289597938922, 82.32080734983464, 99.95838476292244,
111.04792168676437, 86.56867462526922, 96.68436451829949, 111.04792168676437, 82.32080734983464,
fU, fU, -205.19367383521234, U, ?, ?, ?, 14, 15, 0, 103.96671706036422, 109.15546284590134,
120.06342587530992, 120.06342587530992, 109.31326064755098, 114.60944436060562,
120.06342587530992, 109.15546284590134, f, f, -72.71975352939056, 16, 16, 0, ?, ?, ?, 103.45658413403169,
106.24667270472571, 120.06342587530992, 120.06342587530992, 108.61487112802209,
113.15504929001781, 120.06342587530992, 106.24667270472571, f, f, -72.71975352939056, 19, 20, 0, ?, ?, ?,
ff, -132.47392030582176, 38.26992889395384, 38.33808017371905, 61.708175598025164,
61.70817559802514, 51.767148308090526, 51.67018371474762, 48.91832397553978,
38.33808017371905, rffrf, rfrf, 304.161234477898, -16.54053909241583, -192.66495040034152,
-39.08628577150598, 320.6246164221373, -115.3129708436157, 38.26992889395384, 38.33808017371905,
61.378019967350816, 60.47817057569326, 49.49098476195482, 49.40812537470615, 60.47817057569326,
38.33808017371905, f, f, -115.3129708436157, 53.1109816337451, 52.277401148366955,
60.18830180780424, 59.39943676272692, 58.694816045571045, 57.8594886876932, 52.277401148366955,
55.608300373609524, r, r, 14.617160408337508, 50.63601847935072, 49.92159578993343,
60.18830180780424, 59.39943676272692, 57.639001397129555, 56.852434570493024, 52.277401148366955,
49.92159578993343, rf, rf, 14.617160408337508, -177.70951823987795, fr, -129.9301312519532, 46,
35.14285714285714, 1.3089430894308944, 0.028975508796136598, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

0.6700238185103393,	0.7652708398993883,	-0.28214923178672036,	-0.3773962531757692,
0.041382909732373636,	0.01110704244344693,	0.011332476889074499,	0.009157730071454751,
0.008975557381223577,	0.009828592768087237,	-0.07805831705728554,	0.04419834345892385,
-0.25513959265955977,	-0.3773962531757692,	-0.15730729705108615,	0.009584705856263039,
0.009835853204843401,	0.009210743256735054,	0.008975557381223577,	0.009419423556533524,
-20.311154101575767,	16.819859980172723,	-20.311154101575767,	-0.22052550041247687,
0.15176372279411185,	-0.22052550041247687,	0.16388520439742194,	0.8361147956025781,
0.11670324467944404,	0.883296755320556,	-90.65447854221519,	-85.20049702751089,
-85.20049702751089,	-0.5827663674982072,	-0.5362764277966529,	-0.5362764277966529,
0.5392124988218508,	-0.13081131968848858,	0.9528747156230797,	-0.16348892836059797,
0.6667068421188481,	-0.3773962531757692,	1.0441030952946173,	-0.8660237006490935,
-205.19367383521234,	-2.4926100756423706,	-72.71975352939056,	0.03813168355917774,
0.16585225049888871,	-0.31012051462743845,	-0.43784108156714935,	-0.11313656461598778,
0.03126929629077778,	0.0322786903183036,	0.02855180397515074,	0.02765895423052331,
0.03011634363288526,	0.017997738781345436,	0.19182879249999277,	-0.2640100278485021,
-0.43784108156714935,	-0.14034415180091303,	0.031116576570247228,	0.03249979900974809,
0.02888847528848471,	0.02765895423052331,	0.029903977120976578,	6.059253468524403,
19.391357009552976,	6.059253468524403,	0.17203216732907411,	0.4704694025142455,
0.17203216732907411,	0.6007454108123551,	0.3992545891876449,	0.7924904525226082,
0.20750954747739184,	-6.698164722696092,	-5.158216647662371,	-5.158216647662371,
-0.10954121342971196,	-0.0832739316637219,	-0.0832739316637219,	0.3603228600036202,
0.32219117644444234,	0.5783896720166231,	0.506864803520499,	-0.0664121275592495,
-0.31012051462743845,	0.24370838706818887,	-0.6996335094294047,	-115.3129708436157,
-3.0077920000455145,	14.617160408337508,	0.551262238721947,	1.2982269672237463,
1.2576068376068374,	1.6666666666666667,	0.551262238721947,	1.2982269672237463,
1.2576068376068374,	1.6666666666666667, ?, ? , ? , ? , ? , ? , ? .		

a:

../demo/data/demo_H.wav, H, male, 125, 142, b, 142, 164, 142, 142, 0, a:17, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?,
103.96671706036422, 109.15546284590134, 120.06342587530992, 120.06342587530992,
109.31326064755098, 114.60944436060562, 120.06342587530992, 109.15546284590134, f, f,
-72.71975352939056, 16, 16, 0, ?, ?, ?, 103.96671706036422, 109.15546284590134, 120.06342587530992,
120.06342587530992, 109.31326064755098, 114.60944436060562, 120.06342587530992,
109.15546284590134, Uf, Uf, U, -72.71975352939056, 16, 16, 0, ?, ?, ?, 103.45658413403169,
104.7922776341379, 121.76366067910342, 121.76366067910342, 111.36403692763508,
111.95724900388727, 108.42826531060743, 106.73396092711226, fUf, fUf, -72.71975352939056, U,
-115.6130750153166, 13, 14, 1, 5, 6, 0, 103.45658413403169, 104.7922776341379, 121.76366067910342,
121.76366067910342, 111.88864504429267, 112.47338485963157, 108.42826531060743,
109.0462224274186, fUf, fUf, -72.71975352939056, U, -115.6130750153166, 12, 12, 0, 5, 6, 0, ff, 0,
53.1109816337451, 52.277401148366955, 60.18830180780424, 59.39943676272692, 58.694816045571045,
57.8594886876932, 52.277401148366955, 55.608300373609524, r, r, 14.617160408337508,
37.407402480764375, 35.416107651411544, 60.18830180780424, 59.39943676272692, 55.04910736590773,
54.21485120127927, 35.416107651411544, 55.608300373609524, rr, r, 506.90207435453294,
14.617160408337508, 46.13045940081604, 46.13045940081604, 61.98893067685692, 61.83821369703302,
55.92191235835698, 55.59386899915398, 53.712732179050825, 47.27686074975728, firff, firf,
-177.70951823987795, 385.83291180117175, 52.517204977804276, -77.62991447936763,
-351.5357269632099, 46.13045940081604, 46.13045940081604, 61.98893067685692, 61.83821369703302,
56.67219307920548, 56.27199016114622, 53.712732179050825, 54.7762895916391, firf, firf,
-177.70951823987795, 385.83291180117175, 52.517204977804276, -77.62991447936763, rf,
192.32667864821545, 17, 15, 1.1333333333333333, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, -0.014061806261395266,
0.13604006963537318, 0.04079304824632444, -0.10930882765044422, 0.023413119303965838,
0.009715021646681492, 0.01002848880425557, 0.009828994303962418, 0.009521762878927824,
0.009791904705122709, -0.014061806261395266, 0.13604006963537318, 0.04079304824632444,
-0.10930882765044422, 0.018813597778859995, 0.009715021646681492, 0.01002848880425557,
0.009828994303962418, 0.009521762878927824, 0.009782368228527122, 6.5235013944909355,
11.977482909195217, 6.5235013944909355, 0.061623731374243436, 0.11038081306173814,
0.061623731374243436, 0.11670324467944404, 0.883296755320556, 0.09086143751517198,
0.909138562484828, -93.30667389893354, -83.50026222371739, -83.50026222371739, -0.606179486802173,

-0.5222146215352578, -0.5222146215352578, -0.13081131968848858, -0.11674951342709325,
-0.16348892836059797, -0.14692259970674623, 0.10193134699615772, 0.006684325607108722,
0.09524702138904899, 0, -72.71975352939056, -0.6662035195806153, -72.71975352939056,
-0.040236773764902256, 0.25281128786478624, 0.1250907209250753, -0.16795734070461324,
0.03994453707169263, 0.030677834360301994, 0.03302421102105738, 0.03199150365159186,
0.029718501051690153, 0.03128803988977304, -0.040236773764902256, 0.25281128786478624,
-0.2643067212981792, -0.5573547829278678, -0.03724202257093449, 0.030677834360301994,
0.03302421102105738, 0.028842961346539277, 0.026793663294045703, 0.030693781526587364,
23.329473668414877, 25.580661982498555, 3.1372809462168973, 0.5439209820909867,
0.5836059671302334, 0.09275523973325796, 0.7924904525226082, 0.20750954747739184,
0.7223014177961814, 0.27769858220381854, -8.96378441123531, -2.719439713356273,
-2.719439713356273, -0.1494857505014046, -0.04303715789881963, -0.04303715789881963,
0.32219117644444234, 0.3624279502093447, 0.506864803520499, 0.5824182601742882,
-0.027085894442846082, 0.03468240569162828, -0.061768300134474256, 1.0356195891053528,
14.617160408337508, 0.26285932693735936, -177.70951823987795, 0.25964539344474924,
0.25964539344474924, 1.1333333333333333, 1.1333333333333333, 0.25964539344474924,
0.25964539344474924, 1.1333333333333333, 1.1333333333333333, ?, ?, ?, ?, ?, ?, ?.

b:

../demo/data/demo_H.wav, H, male, 142, 164, d, 164, 187, 164, 164, 0, b:23, SUSP, ?, ?, ?, ?, ?, ?, ?, ?,
103.45658413403169, 104.7922776341379, 121.76366067910342, 121.76366067910342,
111.36403692763508, 111.95724900388727, 108.42826531060743, 106.73396092711226, fUf, fUf,
-72.71975352939056, U, -115.6130750153166, 13, 14, 1, 5, 6, 0, 103.45658413403169, 104.7922776341379,
121.76366067910342, 121.76366067910342, 112.20306296982986, 112.70833980892972,
106.24667270472571, 106.73396092711226, Uf, Uf, U, -115.6130750153166, 13, 14, 1, 3, 3, 0,
137.40824334782533, 140.45337965159138, 156.95922713599148, 142.1922792106645,
145.79546547246838, 141.32282943112793, 140.45337965159138, 142.1922792106645, Ur, Ur, U,
10.868122244207084, 16, 17, 1, 16, 17, 5, 137.40824334782533, 140.45337965159138, 156.95922713599148,
141.97491676578036, 145.95792770598186, 141.21414820868588, 140.45337965159138,
141.97491676578036, Ur, Ur, U, 10.868122244207084, 14, 15, 0, ?, ?, ?, fr, -126.48119725952368,
46.13045940081604, 46.13045940081604, 61.98893067685692, 61.83821369703302, 55.92191235835698,
55.59386899915398, 53.712732179050825, 47.27686074975728, firff, firf, -177.70951823987795,
385.83291180117175, 52.517204977804276, -77.62991447936763, -351.5357269632099, 46.13045940081604,
46.13045940081604, 61.98893067685692, 61.83821369703302, 56.51557639226279, 56.223319834930955,
48.02602759537473, 47.27686074975728, rrf, rf, 385.83291180117175, 52.517204977804276,
-77.62991447936763, -351.5357269632099, 46.17024996173063, 46.83300796170157, 63.35117582894084,
63.35117582894084, 56.147710883425205, 55.932187743078245, 47.165897552743345,
48.855184634273286, firff, firf, -10.402799720055652, 392.9566129532366, 123.23613289799019,
-48.76572863364754, -177.73413378303292, 46.17024996173063, 46.83300796170157, 63.35117582894084,
63.35117582894084, 57.08793423318845, 56.75026758757076, 47.165897552743345, 54.54267691533034,
firff, firf, -10.402799720055652, 392.9566129532366, 123.23613289799019, -48.76572863364754,
-177.73413378303292, ff, -341.13292724315426, 22, -1, -22, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?,
-0.15509826223989004, -0.14279365791226636, -0.292895533809035, -0.3052001381366586,
-0.2329297506345617, 0.009438701046372993, 0.00946218788219145, 0.009166421969923212,
0.00914366927778364, 0.009285152703115463, -0.1535684411320785, -0.14279365791226636,
-0.292895533809035, -0.30367031702884706, -0.22547410143061364, 0.009441614795916849,
0.00946218788219145, 0.009166421969923212, 0.009146491950316272, 0.009299756300073004,
4.101999475701859, 9.325287552476865, 2.160316182727499, 0.039189992547017105,
0.08696769375777225, 0.02083068312791907, 0.09086143751517198, 0.909138562484828,
0.3769865394030801, 0.62301346059692, -63.94109347169288, -63.0716436921563, -63.289006137040445,
-0.37324973616761137, -0.36711635929536757, -0.3686461804031792, -0.11674951342709325,
0.038348748812796825, -0.14692259970674623, 0.05212473645113002, -0.2587868111698194,
-0.2745362243899368, 0.01574941322011743, -0.8268473844538816, -115.6130750153166,
-1.0831892118598296, 10.868122244207084, -0.024171949451014228, 0.2779332653349916,
-0.015114796294696899, -0.31722001108070275, -0.0060670984268617, 0.03079955813258639,
0.03321850877503219, 0.030858327220804158, 0.028611243495249575, 0.030933352483498648,
-0.024171949451014228, 0.2779332653349916, -0.015114796294696899, -0.31722001108070275,
-0.009328757033511773, 0.03079955813258639, 0.03321850877503219, 0.030858327220804158,

0.028611243495249575, 0.030908500637452672, 14.998034044562637, 23.315042293959337,
13.85163269562139, 0.38160947857326877, 0.5436614300585407, 0.3570619610314372,
0.7223014177961814, 0.27769858220381854, 0.7327825529072608, 0.2672174470927392,
-8.625465667311047, -1.2064775814484534, -1.2064775814484534, -0.14341865207454288,
-0.01886520844780536, -0.01886520844780536, 0.3624279502093447, 0.3865998996603589,
0.5824182601742882, 0.6292899183583306, 0.12997894973783225, 0.0023498519117425595,
0.12762909782608975, -1.836895143330858, -351.5357269632099, -7.4356825175837145,
-10.402799720055652, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

d:

../demo/data/demo_H.wav, H, male, 164, 187, haitiyi, 187, 238, 187, 187, 0, d:23, 0, ?, ?, ?, ?, ?, ?, ?, ?,
137.40824334782533, 140.45337965159138, 156.95922713599148, 142.1922792106645,
145.79546547246838, 141.32282943112793, 140.45337965159138, 142.1922792106645, Ur, Ur, U,
10.868122244207084, 16, 17, 1, 16, 17, 5, 137.40824334782533, 140.45337965159138, 156.95922713599148,
142.1922792106645, 145.79546547246838, 141.32282943112793, 140.45337965159138, 142.1922792106645,
Ur, Ur, U, 10.868122244207084, 16, 17, 1, 16, 17, 3, 102.60401297201705, 114.24074853016559,
133.50271367105245, 122.98310375874283, 122.31099675327592, 117.982640610027,
114.24074853016559, 122.98310375874283, UrUr, UrUr, U, 30.64417548926255, U, 47.00849771463626, 15,
16, 0, 15, 17, 6, 102.60401297201705, 114.24074853016559, 130.050626269715, 117.9180495888771,
118.66298349385823, 116.07939905952132, 114.24074853016559, 117.9180495888771, Ur, Ur, U,
30.64417548926255, 13, 14, 0, ?, ?, ?, rr, -19.776053245055465, 46.17024996173063, 46.83300796170157,
63.35117582894084, 63.35117582894084, 56.147710883425205, 55.932187743078245, 47.165897552743345,
48.855184634273286, firff, frf, -10.402799720055652, 392.9566129532366, 123.23613289799019,
-48.76572863364754, -177.73413378303292, 46.523950178790024, 46.83300796170157, 63.35117582894084,
63.35117582894084, 57.60927035084922, 57.33385933055437, 46.83300796170157, 48.855184634273286,
rfff, rf, 392.9566129532366, 123.23613289799019, -48.76572863364754, -177.73413378303292,
42.81166150341454, 42.81166150341447, 59.29415513835917, 59.255216994428636, 53.479218299939454,
53.1898980687364, 46.95935387392093, 52.87557270267665, rfffrff, rfffrf, 276.08733949325415,
50.48200964258106, -94.01172946283309, -267.8811699921035, 142.7814041513301, -47.27252866326244,
79.1920102089292, -9.812220519148559, 43.16769235321622, 43.16769235321623, 59.29415513835917,
59.255216994428636, 54.06253539726378, 53.77911098181633, 46.95935387392093, 56.246841651617984,
rff, rf, 276.08733949325415, 50.48200964258106, -94.01172946283309, fr, -453.82147327628707, 23, 13,
1.7692307692307692, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 0.14513324248994894, 0.21887216978178387,
0.2065675654541601, 0.1328286381623251, 0.1803769307291842, 0.010037422859739624,
0.01019362815211654, 0.010168325751257035, 0.01001250818824057, 0.010111952898590116,
0.18719033242867064, 0.21887216978178387, 0.2065675654541601, 0.17488572810104702,
0.19677241212772625, 0.010125923004574496, 0.01019362815211654, 0.010168325751257035,
0.010100788659951475, 0.010146824746657911, 39.560317759254104, 38.690867979717524,
37.82141820018097, 0.3260308212645778, 0.319897444392334, 0.313726216936954, 0.3769865394030801,
0.62301346059692, 0.14972239049399097, 0.8502776095060091, -87.26565884181811, -82.28081914407798,
-87.34587331394371, -0.5536266668967954, -0.5122496017853164, -0.5543066917240383,
0.038348748812796825, -0.10678449367715216, 0.05212473645113002, -0.1350408909057629,
0.2065675654541601, 0.21887216978178387, -0.012304604327623649, -0.12928228270122233,
10.868122244207084, 0.07643257639963316, 30.64417548926255, 0.0668396407526865,
0.39188293692164256, 0.08977772213563678, -0.23526557403331927, 0.0502715314128721,
0.03148735342692737, 0.03421167318231807, 0.03172040094056965, 0.029194464413853056,
0.03137197446607108, 0.0668396407526865, 0.3836011141630248, 0.08149589937701883,
-0.23526557403331927, 0.06400624241608445, 0.03148735342692737, 0.03413642004454953,
0.03165062768834275, 0.029194464413853056, 0.03147767001103188, 16.57635792907864,
23.6533610378836, 14.554181256506922, 0.4144490113005481, 0.5497285284854024, 0.3721767573261341,
0.7327825529072608, 0.2672174470927392, 0.647826253244097, 0.3521737467559029,
-11.367755341652895, -5.302436415960656, -5.302436415960656, -0.19369018348741515,
-0.08570484920049187, -0.08570484920049187, 0.3865998996603589, 0.3197602589076723,
0.6292899183583306, 0.5023968663920694, 0.0043887056280221305, 0.039578090267043925,
-0.03518938463902185, -2.443688057137543, -177.73413378303292, -3.6379789599310492,
276.08733949325415, 1.091089451179962, 1.091089451179962, 1.7692307692307692, 1.7692307692307692,
1.091089451179962, 1.091089451179962, 1.7692307692307692, 1.7692307692307692, ?, ?, ?, ?, ?, ?.

haitiyi:

../demo/data/demo_H.wav, H, male, 187, 238, iS1gal, 238, 279, 238, 238, 0, h:6_a:8_i:12_t:7_i:8_y:3_i:7, SUSP, ?, ?, ?, ?, ?, ?, ?, ?, ?, 102.60401297201705, 114.24074853016559, 133.50271367105245, 122.98310375874283, 122.31099675327592, 117.9982640610027, 114.24074853016559, 122.98310375874283, UrUr, UrUr, U, 30.64417548926255, U, 47.00849771463626, 15, 16, 0, 15, 17, 6, 115.26155865691203, 115.9318291015474, 133.50271367105245, 122.98310375874283, 124.77731906082002, 119.45746643014509, 115.9318291015474, 122.98310375874283, Ur, Ur, U, 47.00849771463626, 15, 16, 0, 15, 16, 3, 105.13639370098308, 105.13639370098309, 141.24486016359108, 137.08401570013385, 130.57667537268827, 125.39299447233178, 113.98669910450877, 121.00879543049238, Uf, Uf, U, -133.96016891367884, 13, 14, 0, ?, 3, 6, 105.13639370098308, 105.13639370098309, 113.98669910450877, 113.98669910450877, 109.78966640378928, 109.56154640274593, 113.98669910450877, 105.13639370098309, U, UU, U, ?, 3, 11, ?, 3, 6, rf, 180.9686666283151, 42.81166150341454, 42.81166150341447, 59.29415513835917, 59.255216994428636, 53.479218299939454, 53.1898980687364, 46.95935387392093, 52.87557270267665, rrrfrfrf, rfrfrf, 276.08733949325415, 50.48200964258106, -94.01172946283309, -267.8811699921035, 142.7814041513301, -47.27252866326244, 79.1920102089292, -9.812220519148559, 45.86794551903271, 45.86794551903271, 58.56988626152594, 57.89422271136563, 56.29951440493003, 55.59963970746619, 45.86794551903271, 52.87557270267665, rf, rf, 79.1920102089292, -9.812220519148559, 44.93154639605038, 44.93154639605039, 63.669415043819626, 63.59762778432595, 55.50952759973103, 55.506644815639845, 48.90355954936352, 60.54380041217281, rfrfrfrf, rfrfrf, 167.11649863566396, -82.31865560749611, 96.10840417132866, -268.9177572754566, -14.39235722668729, 339.83125888393045, 106.33211381678089, -31.8107017932619, 44.93154639605038, 44.93154639605039, 58.68162780865706, 58.38632352197177, 52.74049645974597, 52.465243349942284, 48.90355954936352, 52.649411366762045, rfr, rfr, 167.11649863566396, -82.31865560749611, 96.10840417132866, fr, -176.9287191548125, 51, 33.42857142857143, 1.5256410256410255, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, -0.10854701302932311, 0.15678848329007905, 0.08304955599824414, -0.18228594032115808, -0.060782848176371015, 0.009528613196298465, 0.010071715005678659, 0.009917377888047745, 0.00938259847140728, 0.009620973129749329, 0.07596521105585496, 0.15678848329007905, 0.09774384336420354, 0.01692057112997943, 0.08647391959808029, 0.009899836857495034, 0.010071715005678659, 0.0099481333389472, 0.009778364165055963, 0.009922956322350608, 20.351142307332424, 15.366302609592296, 13.299867650136989, 0.1808975787746289, 0.13952051366314988, 0.12185293884875335, 0.14972239049399097, 0.8502776095060091, 0.22177334135523905, 0.7782266586447609, -79.87092843048903, -68.17990720268696, -91.27722379831204, -0.4928438187204245, -0.4037025887559933, -0.5882148128411716, -0.10678449367715216, 0.0017625193521711281, -0.1350408909057629, 0.002352098003124769, 0.0022628376402015, 0.07596521105585496, -0.07373892729183497, 1.1830491164840853, 47.00849771463626, 0.38223541509290004, -133.96016891367884, -0.07072234398396691, 0.2767136862924887, -0.04832960987646725, -0.39576564015292287, -0.042634247602703484, 0.030452440310536087, 0.033232962074092404, 0.030586578349270862, 0.02802747312767618, 0.03065121419460159, -0.00846407700014684, 0.2534774587551312, 0.020626378012683508, -0.24131515774259465, 0.058025802225664794, 0.030915587207845117, 0.03304378120155673, 0.031147992801528895, 0.029141897591297223, 0.031433993286398246, 20.596745997482003, 20.91107136354175, 13.589118813838063, 0.49352997324563447, 0.49945699707253016, 0.35135502307964817, 0.647826253244097, 0.3521737467559029, 0.7195992073251887, 0.2804007926748113, -9.051008594749447, -0.9600256260633415, -6.1713298884175245, -0.15105593588471167, -0.014982505216525001, -0.10047699973770251, 0.3197602589076723, 0.3904826028916393, 0.5023968663920694, 0.6369250148144682, -0.04056776990015851, 0.07809128231197181, -0.1186590522121302, -0.952706346929594, -9.812220519148559, -0.18557190055081613, 167.11649863566396, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

iS1gal:

../demo/data/demo_H.wav, H, male, 238, 279, mi, 279, 289, 279, 279, 0, i:8_S1:12_g:8_a:10_l:4, SUSP, ?, ?, ?, ?, ?, ?, ?, ?, ?, 105.13639370098308, 105.13639370098309, 141.24486016359108, 137.08401570013385, 130.57667537268827, 125.39299447233178, 113.98669910450877, 121.00879543049238, Uf, Uf, U, -133.96016891367884, 13, 14, 0, ?, 3, 6, 128.62636831239635, 121.00879543049238, 141.24486016359108, 137.08401570013385, 135.37367744243417, 129.04640556531314, 137.08401570013385, 121.00879543049238, Uf, Uf, U, -133.96016891367884, 13, 14, 0, ?, ?, ?, 107.53403058198809,

106.27317684998772, 134.0852790798437, 119.6691937413556, 119.86207746499721, 112.97118529567166, 119.6691937413556, 106.27317684998772, f, f, -133.96016891367884, ?, ?, ?, ?, ?, ?, 94.90565500106408, 94.21676164775661, 134.0852790798437, 119.6691937413556, 111.25164810808022, 106.9429776945561, 119.6691937413556, 94.21676164775661, f, f, -133.96016891367884, ?, ?, ?, ?, ?, ?, ff, 0, 44.93154639605038, 44.93154639605039, 63.669415043819626, 63.59762778432595, 55.50952759973103, 55.506644815639845, 48.90355954936352, 60.54380041217281, rfrf, rfrf, 167.11649863566396, -82.31865560749611, 96.10840417132866, -268.9177572754566, -14.39235722668729, 339.83125888393045, 106.33211381678089, -31.8107017932619, 48.4725705367772, 49.320399857903176, 63.669415043819626, 63.59762778432595, 58.2785587397161, 58.548046281337385, 49.78095528915718, 60.54380041217281, frf, frf, -14.39235722668729, 339.83125888393045, 106.33211381678089, -31.8107017932619, 57.09361452846416, 57.15065888755821, 57.75083821022652, 60.204486259711345, 57.45161164321204, 58.67757257363477, 60.204486259711345, 57.15065888755821, f, f, -31.8107017932619, 50.475045013294476, 52.200130614975734, 57.75083821022652, 60.204486259711345, 56.39664100494748, 57.00091934278746, 60.204486259711345, 52.200130614975734, f, f, -31.8107017932619, ff, 0, 41, 19.142857142857142, 2.1417910447761193, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 0.13586277393783877, 0.2545810720755778, -0.10275442424382446, -0.12947272238156332, 0.10431997237484497, 0.010020222303463515, 0.010275168819784943, 0.009721096075017468, 0.009479897159202139, 0.00958577477084273, 0.13586277393783877, 0.37499588858467603, 0.2502651298970451, 0.011132015250207858, 0.18787629906991288, 0.010020222303463515, 0.01054736238818157, 0.010280000253610578, 0.009766222495233571, 0.010135348565751481, 18.37683397908198, 22.761033020921374, 18.37683397908198, 0.16471383311632104, 0.20030336183952083, 0.16471383311632104, 0.22177334135523905, 0.7782266586447609, 0.10074077994851739, 0.8992592200514826, -92.29273760714915, -85.59472916146521, -85.59472916146521, -0.5971637910952694, -0.5395653626938322, -0.5395653626938322, 0.0017625193521711281, -0.13410025458566782, 0.002352098003124769, -0.16733015022149908, -0.048649450147339374, 0.011132015250207858, -0.05978146539754724, 0, -133.96016891367884, -1.1070283646499544, -133.96016891367884, 0.05482929853719376, 0.1068852513963737, -0.24055077888008186, -0.292606731739262, -0.05555484524131759, 0.03139458450528049, 0.031798538995045894, 0.029138031950025473, 0.0287678753578948, 0.030557401204858264, 0.05482929853719376, 0.1974911735738034, -0.056747211405822756, -0.1994090864424325, 0.026780324581442302, 0.03139458450528049, 0.0325270027200132, 0.030535561190730916, 0.029472474438252693, 0.03118526498098727, 28.26497370697816, 23.2278181104452, 17.04157315270853, 0.6289555815306921, 0.5420912446752336, 0.423926290363794, 0.7195992073251887, 0.2804007926748113, 0.8178347406968098, 0.1821652593031902, -5.880080836754523, -4.3531671506779475, -4.3531671506779475, -0.09550109064339406, -0.06981180375371876, -0.06981180375371876, 0.3904826028916393, 0.3356533043544456, 0.6369250148144682, 0.5318052855379222, -0.20789668587256402, 0.005620204724465512, -0.21351689059702944, 0, -31.8107017932619, -0.5254163362177394, -31.8107017932619, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

mi:

../demo/data/demo_H.wav, H, male, 279, 289, ediyor, 289, 345, 289, 289, 0, m:7_i:3, SUSP, ?, ?, ?, ?, ?, ?, ?, ?, ?, 107.53403058198809, 106.27317684998772, 134.0852790798437, 119.6691937413556, 119.86207746499721, 112.97118529567166, 119.6691937413556, 106.27317684998772, f, f, -133.96016891367884, ?, ?, ?, ?, ?, ?, 107.53403058198809, 106.27317684998772, 141.24486016359108, 131.7256089435867, 127.64606713188216, 118.99939289678721, 131.7256089435867, 106.27317684998772, f, f, -133.96016891367884, ?, ?, ?, ?, ?, ?, 80.97272635803911, 80.97272635803904, 297.20360068448406, 297.20360068448406, 121.4656435876833, 126.86463992871265, 104.93357516085092, 282.8988166990916, fUfUfU, fUfUfU, -133.96016891367884, U, -232.21777616429745, U, -357.61959963481155, U, 4, 6, 5, 12, 13, 0, 88.36990607989475, 90.19795658034624, 118.12757054432663, 118.12757054432663, 98.35002557956196, 102.24736856190509, 104.93357516085092, 108.83885949775473, fUf, fUf, -133.96016891367884, U, -232.21777616429745, 4, 4, 0, 12, 13, 0, ff, 0, 57.09361452846416, 57.15065888755821, 57.75083821022652, 60.204486259711345, 57.45161164321204, 58.67757257363477, 60.204486259711345, 57.15065888755821, f, f, -31.8107017932619, 57.09361452846416, 57.15065888755821, 63.669415043819626, 63.25831363186449, 59.379380823056614, 60.20448625971135, 63.25831363186449, 57.15065888755821, f, f, -31.8107017932619, 39.23704879152999, 39.237048791529986, 57.28791887195099, 56.811344735096746, 47.53616768440364, 47.42664678645171, 56.811344735096746, 42.73950833246955, ffrfrfr, frfrfr, -31.8107017932619, -120.71855650088668, 245.23566967035882, -61.245597611576585, 192.58927682514636, -182.44890257300247,

111.45513709639782, 45.64535041185501, 45.76180760159511, 57.28791887195099, 56.811344735096746, 52.2349979090338, 52.43688856286459, 56.811344735096746, 52.956062656523144, ffr, fr, -31.8107017932619, -120.71855650088668, 245.23566967035882, ff, 0, 10, 6.142857142857143, 1.627906976744186, 0, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 0.9096862109155243, 0.390618830869778, 0.2719005327320392, -1.0284045090532632, -0.11598790191621311, 0.008187015590309946, 0.010609715947802217, 0.010346468679801742, 0.007983880134296835, 0.009510205059602104, 0.10895589212562319, 0.37871426643249234, 0.16400614658316298, -0.1057522277237061, 0.15172333213498496, 0.009966032433570172, 0.010563192954073177, 0.010098506838909654, 0.009527616046092972, 0.010063025697446994, 3.641215398577316, 10.339223844261255, 3.641215398577316, 0.0348635197283742, 0.09598338946467586, 0.0348635197283742, 0.10074077994851739, 0.8992592200514826, 0.2361123974891082, 0.7638876025108918, -78.39928297410816, 91.93967778166325, -87.13635235849418, -0.48117588917905635, 0.370120848221692, -0.552531433107865, -0.13410025458566782, 0.7755859563298565, -0.16733015022149908, 1.5624858294446355, 0.13140368475805547, 0.01268538662031643, 0.11871829813773888, 0, -133.96016891367884, -1.2605266247265132, -133.96016891367884, 0.05801083508949457, 0.4281254494421527, 0.37606949658297273, 0.00595488223014354, 0.2128733460009334, 0.0314249357669935, 0.03459442565239852, 0.03415495346264417, 0.031025727366392315, 0.03268892646746825, 0.10749052275485586, 0.32377671171555633, 0.22224107119101502, 0.005954882230314354, 0.1381365482681694, 0.031804387220481796, 0.033603508427109546, 0.03278080108213365, 0.031025727366392315, 0.03206084313766426, 24.871832182363562, 26.398745868440123, 24.871832182363562, 0.5712794239470464, 0.5976460899165512, 0.5712794239470464, 0.8178347406968098, 0.1821652593031902, 0.4692803805913838, 0.5307196194086162, -17.13100662393758, -7.746308675292546, -0.3083744366443274, -0.12782263884321327, -0.12782263884321327, 0.3356533043544456, 0.277642469264951, 0.5318052855379222, 0.42668555626137633, 0.05801083508949457, 0.005954882230314354, 0.05205595285918019, 0, -31.8107017932619, -0.5566112869468096, -31.8107017932619, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

ediyor:

../demo/data/demo_H.wav, H, male, 289, 345, ?, ?, ?, 345, 3200, 2855, e:12_d:5_i:3_y:2_o:8_r:23, SUSP, ?, ?, ?, ?, ?, ?, ?, ?, 80.97272635803911, 80.97272635803904, 297.20360068448406, 297.20360068448406, 121.4656435876833, 126.86463992871265, 104.93357516085092, 282.8988166990916, fUfUfU, fUfUfU, -133.96016891367884, U, -232.21777616429745, U, -357.61959963481155, U, 4, 6, 5, 12, 13, 0, 282.8988166990916, 282.8988166990916, 297.20360068448406, 297.20360068448406, 290.58143920425647, 290.05120869178785, 297.20360068448406, 282.8988166990916, UfU, UfU, U, -357.61959963481155, U, 4, 6, 5, 4, 6, 9, ?, 39.23704879152999, 39.237048791529986, 57.28791887195099, 56.811344735096746, 47.53616768440364, 47.42664678645171, 56.811344735096746, 42.73950833246955, ffrfrfr, frfrfr, -31.8107017932619, -120.71855650088668, 245.23566967035882, -61.245597611576585, 192.58927682514636, -182.44890257300247, 111.45513709639782, 39.23704879152999, 39.237048791529986, 47.45419126940294, 47.454191269402905, 42.30696791849623, 42.44079283467879, 42.5034806641474, 42.73950833246955, frfr, frfr, -61.245597611576585, 192.58927682514636, -182.44890257300247, 111.45513709639782, ?, 56, 16.142857142857142, 3.4690265486725664, 2.9544670576060708, ?, 180.26685524768118, 24.232678477302244, 180.26685524768118, 1.013939896198607, 0.21197129138088897, 1.013939896198607, 0.2361123974891082, 0.7638876025108918, ?, ?, ?, ?, ?, ?, ?, ?, 0.7755859563298565, ?, 1.5624858294446355, ?, ?, ?, -0.9917617630905493, ?, -357.61959963481155, -1.2641254700446398, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, 10.460681627274901, 15.147820081257066, 6.95822208633534, 0.28071225082060264, 0.3847727439156179, 0.1952099273640739, 0.4692803805913838, 0.5307196194086162, ?, ?, ?, ?, ?, ?, ?, 0.277642469264951, ?, 0.42668555626137633, ?, ?, ?, 0.2846122908961293, ?, 111.45513709639782, 2.607777708376782, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?.

3.3.5. Cümle Bölütlemesi için Üretken (Generative) ve Ayırtedici (Discriminative) Yöntemler

Yukarıda da belirtildiği üzere, projenin amacı sadece belirtilen özelliklerin çıkarılması ile sınırlı olmayıp, çıkarılan özelliklerin özellikle cümle bölütlemesi için kullanılmalarıdır. Bu amaçla cümle

bölütleme işlevi için standart sınıflandırıcılar dışında aşağıda detayları ile verilmekte olan üretken (generative) ve ayırdedici (discriminative) yöntemler de kullanılmıştır.

Literatürde cümle bölütleme için özellikle sözcüksel bilginin yada sözcüksel özelliklerin üstünlüklerinin kullanıldığı istatistiksel modelleri baz alan çözümler geliştirilmiştir. Bununla birlikte, Türkçe geniş bir kelime üretme yeteneğine sahip üretken bir biçimbilgisel yapıya haiz olduğundan bu çözümler yetersiz kalabilmektedir. Bu çalışma, bir yönü ile de kelimelerden ve bu kelimelerin biçimbilgisel analizlerinden yeni bir biçimbilgisel özellik seti oluşturmaktadır. Biçimbilgisel bilgiyi işlemek için daha önce İngilizce ve benzeri dillere uygulanan HELM (Hidden Event Language Model) modelleme metodu ve fHELM (factored Hidden Event Language Model) modelleme metodu da kullanılmıştır. En nihayetinde Türkçe cümle bölütleme işlevi için, sözcüksel, bürünsel ve biçimbilgisel bilgilerin kullanıldığı Boosting, CRF gibi ayırdedici sınıflandırma teknikleri fHELM ile birleştirilmiştir.

Daha önceki çalışmalar ile İngilizce (ve benzeri diller) için konuşma tanıma, öğelere ayırma, kelime anlam açıklama gibi uygulamalar için istatistiksel dil modelleme teknikleri kullanılarak pek çok yararlı sonuçlara ulaşılmıştır. Ancak, özellikle Türkçe, Çekçe, Macarca gibi eklemeli veya çekimli bir biçimbilgisel yapı ile, göreceli olarak serbest öge sırasına sahip dillerde istatistiksel yaklaşımlar üzerinde kapsamlı bir biçimde çalışılmamıştır. Bu gibi dillerde daha zengin bir biçimbilgisel yapı mevcut olduğundan verilen bir veri büyüklüğü için kelime hazne genişliği diğer dillere göre daha büyüktür. (Waibel ve diğerleri, 2000; Hakkani-Tür ve diğerleri, 2000) Bu durum veri seyrekliğine neden olmaktadır. Ayrıca kelime odaklı istatistiksel modeller biçimbilgisel olarak çözümlenmiş bilgiyi de görememektedir. Genellikle bu iki problemin birleşmesinin sonucu olarak dil işleme performansı düşmektedir.

Benzer bir biçimde, makine öğrenme alanında, ayırdedici sınıflandırma tekniklerinde kaydedilen tüm gelişmelere rağmen, ayırdedici dizi sınıflandırması hala çözülmesi gereken bir sorundur.

Araştırmacılar özellikle maximum entropy Markov modelleri (McCallum ve diğerleri, 2000) veya CRF (Lafferty ve diğerleri, 2001; Altun, 2005; Kuo ve diğerleri, 2006) gibi çeşitli yöntemler önermişlerdir. Bununla birlikte bu teknikler sürekli değerli özellikleri etkili bir biçimde işlemeye çok yeterli değildirler. Diğer taraftan üretken dizi modellemesi için saklı Markov modelleri (HMMs) bu alanda hakimiyetini sürdürmekle birlikte sadece bir durum derecesi kullanılmaktadır. Örneğin; otomatik konuşma tanıma genel olarak kelime dizileri dil modeli için modellenmektedirler (Manning ve diğerleri, 1999). Grafiksel modellerdeki gelişmeler ile FLM (factored language models) modelleri herbir örnek için bir özellik yığını işleyebilmektedirler. Bu projede Türkçe için biçimbilgisel bilgiyi yada özellikleri kullanan istatistiksel sınıflandırma modelleri ile cümle bölütlemesinin gerçekleştirilmesi üzerinde de durulmuştur. Bu proje ile aşağıda belirtilen ve 3 bölümden oluşan bir katkı sunulmuştur.

- 1) HELM modelleri fHELM modellerine genişletilmiş ve bu iki model sınıflandırma modelleri ile birleştirilmiştir.
- 2) Kelimelerden ve bu kelimelerin biçimbilgisel analizlerinden yeni bir biçimbilgisel özellik seti oluşturulmuştur.
- 3) Daha önce başka diller ile yaptığımız çalışmalardan yararlanarak Türkçe için bürünsel bilginin çıkarılması gerçekleştirilmiştir.

(Tur, 2000) Türkçe cümle bölütlemeye biçimbilgisel özelliklerin kullanıldığı ilk çalışmadır. Gerçekleştirdiğimiz proje ile bu çalışma (Tur, 2000) bürünsel özelliklerin kullanılması ve daha gelişmiş sınıflandırma modelleri ile genişletilmiş olmaktadır.

Yöntemde üretken modeller (örneğin; hidden event modeling) baz alınmış ve biri kelimeleri diğeri ise kelimelerin biçimbilgisel analizini kullanan iki ayrı dil modeli olarak eğitilmişlerdir. Daha sonra herbir modelden elde edilen ağırlıklandırılmış olasılık kombinasyonları hesaplanmıştır.

(Guz ve diğeri, 2010) Sözlüksel, bürünsel, biçimbilgisel ve pseudo-biçimbilgisel özellikleri kullanan ayırdedici ve üretken modelleri birleştirerek cümle bölütleme alanında kullanmışlardır.

Literatürde cümle veya diyalog edim etiketleme (dialog act segmentation) bir sınır sınıflandırma problemi olarak ele alınmıştır. Amaç verilen bir özellik dizisi $X = X_1 \dots X_k$ için en olası sınır etiket dizisinin $Y = Y_1 \dots Y_n$ bulunmasıdır.

$$\arg \max_Y P(Y|X) \quad (7)$$

Bunu gerçekleştirmek için üretken, ayırdedici veya hibrid modeller kullanılmıştır. Aşağıda bu yaklaşımlar özetlenmektedir.

3.3.5.1. fHELM (Factored Hidden Event Language Models)

Burada HELM modelleri ile FLM modellerin kullanılması önerilmiştir. Aşağıda öncelikle HELM daha sonra da FLM anlatılmakta ve bu modellerin nasıl birleştirildikleri açıklanmaktadır.

1) HELM (Hidden Event Language Models)

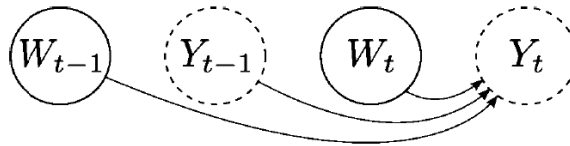
Cümle bölütleme için en yaygın kullanılan üretken model HELM modelidir (Stolcke ve diğeri, 1996). HELM modeli esasen önceleri deletion (DEL) ve repetition (REP) gibi konuşma akışındaki bozuklukları tasarlamada kullanılmaktaydı. Bu bozuklukları gidermek için kullanılan yaklaşım extra meta token lar olarak düşünölmeleri idi. Hesaplamaları kolaylaştırmak için iki kelime arasına hayali bir no disfluency (NODF) token ı koyulmaktadır. Bu durumlar sınırdan önce gelen kelimenin bir akış bozukluğunun parçası olmadığı durumlardır. Aşağıdaki örnek, akış bozukluğuna sahip bir dizinin kavramsal olarak gösterimidir.

...she NODF got REP got NODF real NODF lucky...

Cümle bölütleme için cümle sınırları basit bir biçimde saklı olaylar (hidden events) olarak ele alınmış ve kelime dizisi farazi cümle sınırı token ları (cümle sınırı için S, diğeri için N) ile arttırılmıştır. Bu durumda örnek aşağıdaki biçimde olmaktadır.

... she N got N real N lucky S however N there N were N...

Cümle bölütleme için kavramsal model Şekil 8 de gösterilmektedir.



Şekil 8. Cümle Bölütleme için HELM in kavramsal gösterimi

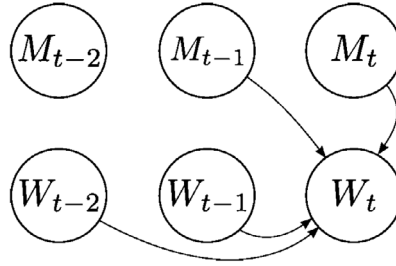
Bu modelde 1 durum sınır token larından S veya N in her biri için ayrılmış, kalan diğeri durumlar ise kelime üretimi için kullanılmıştır. En olası sınır token dizisi basit bir biçimde sadece sözcüksel özelliklerin kullanılması ile Viterbi çözümlemesi yardımı ile elde edilmektedir. Örneğin X ve Y nin sırası ile tüm kelimeleri ($X_t = (W_t)$) ve sınır token larını gösterdiği durumda $P(X, Y)$ nin modellenmesi için dil modeli

$$\arg \max_Y P(Y|X) = \arg \max_Y P(X, Y) \quad (8)$$

2) FLM (Factored Language Models)

FLM bir dizi özellik setini modellemede kullanılan ve bilinen dil modelleme yöntemlerini genişleten modellerdir. Bir başka ifade ile amaç, özellik setlerinin bir alt kümesini (factors) kullanarak olasılıksal dil modelleri inşa etmektir. Daha formal olarak, FLM, bir kelime dizisi W_1, \dots, W_n yerine özellik seti dizisinin X_1, \dots, X_n olasılığını kestirmeyi amaçlar. Burada W_t kelimesi için M_t bir biçimbilgisel özellik olmak üzere, $X_t = (W_t, M_t)$ olduğu varsayılmıştır. Örnek bir FLM Şekil 9 da gösterilmektedir. Şimdiki (ilgilenilen kelime) sadece önceki 2 kelimeye değil aynı zamanda şimdiki ve önceki biçimbilgisel analizlere de dayanmaktadır.

$$P(W_t | W_{t-1}, W_{t-2}, M_t, M_{t-1}) \quad (9)$$



Şekil 9. Okların olasılıkların kestirimi için kullanılan faktörleri, W ve M nin sırası ile kelime ve biçimbilgisel faktörleri gösterdiği FLM grafiği için bir örnek.

Daha düşük dereceli n-gram yaklaşımlar yapıldığında dahi olası gizli n-gram dizileri olduğunda, FLM ler için önemli olabilecek bir durum da, böyle olasılıkların güvenilir bir biçimde kestirilebilmesi için geri çekmenin (back off) nasıl olacağıdır. Bu problemin üstesinden gelebilmek için yeni geliştirilmiş paralel geri çekme tekniği önerilmiştir (Bilmes ve diğerleri, 2003). Temel olarak, sistem, geri çekme yollarını gösteren bir geri çekme grafiği olarak verilmiştir. Grafikteki yollar manual olarak seçilebilmektedirler. Literatürde, optimal geri çekme grafiğinin seçilmesi için genetik algoritmalar öne sürülmüştür (Duh ve diğerleri, 2004). Önemli bir nokta, bir çok geri çekme grafiğinin önerilebileceği ve sistemin onları paralel işleyebilmesidir.

En uzak kelimenin örnek bir geri çekme düşümü, W_t kelimeleri ve M_t biçimbilgisel analizleri kullanan bir FLM ler için aşağıdaki biçimde tanımlanabilir.

$$P(Y_t | C_t) = \begin{cases} P_{ML}(Y_t | C_t), & \text{if } N(C_t, Y_t) > \tau \\ \alpha(C_t) \times P_{BO}(Y_t | \hat{C}_t), & \text{otherwise} \end{cases} \quad (10)$$

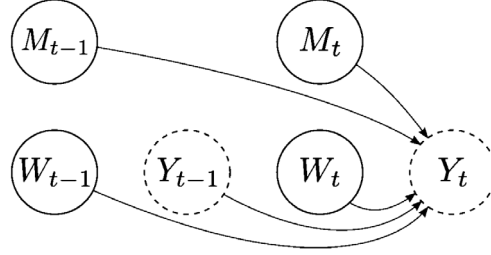
Burada $C_t = W_t, M_t, Y_{t-1}, W_{t-1}, M_{t-1}$ orjinal içerik, $\hat{C}_t = W_t, M_t, Y_{t-1}, M_{t-1}$ geri çekilen içerik, P_{ML} standart enyüksek olasılık kestirimi (maximum likelihood estimation) (smoothing yapılarak) $N(\cdot)$ meydana gelme sayısı olup, α ise sonucun hala bir olasılık dağılımı olduğunu sağlamak için kullanılmıştır.

3) HELM lerden fHELM lere

fHELM ler, HELM ve FLM lerin basit bir uzantısıdır. fHELM ler, FLM lerin çoklu özellik dizisi modelleme yeteneği ile HELM lerin sınıflandırma gücünü birleştirmektedir. Şekil 10, kullandığımız fHELM topolojisini göstermektedir. Sınır durumları potansiyel hibrid modelleri inşa etmek için hala mevcut olup, sınır kararı aşağıdaki ifade ile verilmektedir.

$$P(Y_t | W_t, M_t, Y_{t-1}, W_{t-1}, M_{t-1}) \quad (11)$$

Bu ifadede Y_t sınır kararı olan S veya N i göstermektedir (M_t biçimbilgisel analizi ile W_t kelimesinden sonra).



Şekil 10. Okların olasılıkların kestirimi için kullanılan faktörleri, W ve M ve Y nin sırası ile kelime, biçimbilgisel faktörler ve kelime sınırlarını gösterdiği HELM için oluşturulmuş bir FLM örneği.

Bir fHELM inşa etmek için bir sonraki adım, istenen n-gram için istatistiklerin güvenilir olmadığı durumda olası geri çekme yollarını belirten bir geri çekme grafiğinin yaratılmasıdır. Bu çalışmada sadece doğrusal geri çekme grafiği (düşme ve bir faktörü o zamanda unutma) ve en uzak özellikten başlayarak bütünüyle bağlantılı geri çekme grafiği denenmiştir. Deneylerde kullanılan geri çekme, trigram dil modelinde (M_{t-2}) en uzak biçimbilgisel analizi düşürmekte ve daha sonra en uzak kelimeyi (W_{t-2}) düşürmektedir. Bu işlem böyle devam etmektedir. Daha sonra, en olası durum dizisini (örneğin kelimeleri ve biçimbilgisel analiz gibi diğer özelliklerini veren) bulmak için standart Viterbi çözümü kullanılabılır. Bu, herbir örnek pozisyonu için çoklu özellikler kullanıldığı zaman üretken bir sınıflandırıcı inşa etmek için mükemmel bir yolaçar. Dahası, sıradan HELM lere benzer bir biçimde, performansı daha da arttırmak için diğer sınıflandırıcılardan (tercihen ayırdedici sınıflandırıcılar) elde edilen sonsal olasılıkları birleştirmek de mümkün olabilmektedir. Örneğin; sözcüksel ve biçimbilgisel bilgiyi işleyen bir fHELM, sadece bürünsel bilgiyi işleyen bir sınıflandırıcı ile birleştirilebilir.

Deneylerimizde Viterbi çözümü ve geleneksel ve Kneser-Ney yumuşatması (Kneser ve diğerleri, 1993) ile değiştirilmiş fHELM lerin inşa edilmesinde SRILM tolkit (Stolcke, 2002) kullanılmıştır.

3.3.5.2. Ayırdedici Sınıflandırma Modelleri

HELM lerin bir zayıflığı, kelime ve biçimbilgisel analizleri gibi sadece ayrıntı özellikleri içermesidir. Bu engelin ortadan kaldırılması için, literatürde çeşitli sınıflandırma metodları kullanılmıştır. Önde gelen çalışmalardan birinde, performansı arttırmak için, aynı zamanda bürünsel özelliklerin de eklenmesi ile bölütleme modellerinde karar ağaçları kullanılmıştır (Shriberg ve diğerleri, 2000). Ayırdedici sınıflandırma algoritmalarındaki gelişmeler ile birlikte araştırmacılar CRF (conditional random fields) leri (Liu, Stolcke ve diğerleri, 2005), boosting i (Cuendet ve diğerleri, 2006) ve boosting ile maximum entropy sınıflandırma algoritmalarını (Zimmermann ve diğerleri, 2006) kullanmayı denemişlerdir.

Sistemimiz, bağımsız veya birlikte kullanılabilen iki sınıflandırıcı tarafından sağlanan sınıra yönelik sonsal olasılıklara $P(Y_t | X_t)$ dayanmaktadır.

1) Boosting

İlk bileşen, ağırlıklandırılmış karar kütükleri (bir düzeyli karar ağaçları=sıg karar ağaçları) üzerinden sonsal olasılık kestirimlerini üreten bir Adaboost sınıflandırıcısıdır (Schapire ve diğerleri, 2000).

$$P(Y_t | X_t) = \left[1 + \exp \left(-2m \sum_{i=1}^m w_i s_i(X_t) \right) \right]^{-1} \quad (12)$$

Burada $s_i(\cdot)$ tek bir özellikten fazla karar kütüğü (ayrık bir özelliğin veya sürekli bir özelliğin eşik değerine ilişkin pozisyonun varlığında), w_i karar kütüğüne verilen ağırlık, ve m karar kütüklerinin sayısıdır. Adaboost en iyi karar kütüğü seçimi üzerinden iterasyonla eğitilmekte ve genel sınıflandırıcının hata yaptığı yerde örnekleri yeniden ağırlıklandırmaktadır. Deneylerin gerçekleştirilmelerinde icsiboost (<http://code.google.com/p/icsiboost>) kullanılmıştır. Deneylerimizin tümünde boosting 1000 iterasyon ile çalıştırılmıştır.

2) CRF (Conditional Random Fields)

Sistemimizin ikinci bileşeni (Lafferty ve diğerleri, 2001) tarafından önerilen CRF leri kullanmaktadır. Verilen bir gözlem dizisi ($X = X_1, \dots, X_n$) için sınır olay dizisinin ($Y = Y_1 \dots Y_n$) olasılığını kestirmek için zincir CRF ler kullanılmıştır.

$$Z(X) = \sum_Y \exp \left(\sum_{t=1}^n \sum_{i=1}^m \lambda_i s_i(Y_{t-1}, Y_t, X_t) \right) \quad (13)$$

olmak üzere

$$P(Y | X) = \frac{1}{Z(X)} \exp \left(\sum_{t=1}^n \sum_{i=1}^m \lambda_i s_i(Y_{t-1}, Y_t, X_t) \right) \quad (14)$$

Burada $s_i(\cdot)$ örneklere ve Y_t ye yakın özel sınır grubuna bağlı olan karar fonksiyonlarını, λ_i eğitme verisi üzerinde kestirilen s_i lerin ağırlıklarını ve $Z(X)$ ise bir normalizasyon faktörüdür.

Burada not edilmelidir ki, CRF ler sınır karar dizilerinin olasılıklarını vermektedir. Forward-Backward algoritması sınır düzeyinde sonsal olasılıkları kestirmek için kullanılabilir.

Deneylerimizde CRF++ toolkit (<http://crfpp.sourceforge.net>) kullanılmıştır. CRF toolkit'i şimdiki ve önceki sınıra ilişkin ikili karar fonksiyonlarını gözönüne almaktadır. X den çıkarılan özellikler sınır komşuluğundan oluşmaktadır ve CRF++ sürekli tipteki özellikleri işleyemediğinden ve nicelendirilmeye gereksinim duyduğundan (Bu durum sadece CRF++ toolkit için geçerli olup CRF lerin geneli için bir engel değildir) Adaboost ile kullanılan özelliklere uymaktadır. Farklı tipte nicelendirmeleri denedikten sonra, Adaboost tarafından öğrenilen karar kütüklerinden gelen eşikleri kullanmanın performansı biraz arttırdığı gözlenmiştir. Bu durum büyük olasılıkla karar kütük eşiklerinin negatif ve pozitif sınıfların bölünmesinde zaten optimize edilmiş olmalarından kaynaklanmaktadır. Bu nicelendirme (Friedman ve diğerleri, 2007) tarafından önerilen metoda benzerlikler göstermektedir.

3.3.5.3. Hibrid Üretken ve Ayırdedici Modelleme

Belirtmek durumunda olduğumuz önemli gözlemlerden biri de sıralı olmayan ayırdedici sınıflandırıcı algoritmalarının genellikle bölütleme görevi için kritik olan içeriği dikkate almamalarıdır. Her ne kadar ilave bir özellik olarak içerik ekleme veya basit bir biçimde zaten doğası gereği içeriği dikkate alan CRF leri kullanmak mümkün olsa da, bu yaklaşımlar durak süreleri ve pitch aralığı gibi gerçek değerli özellikler söz konusu olduğunda suboptimal dir. Önceki çalışmaların birçoğu bu sorunu basit bir biçimde özellik uzayını hem manual hem de otomatik olarak muhafaza etmek ile çözmüştür (Kuo ve diğerleri, 2006).

Bir alternatif de (Shriberg ve diğeri, 2000) tarafından önerilen bir hibrid sınıflandırma yaklaşımının kullanılmasıdır. Temel düşünce, boosting veya CRF gibi diğer sınıflandırıcılardan elde edilen P_c sonsal olasılıklarını kullanmaktır. Bu olasılıklar basit bir biçimde aşağıda verilen Bayes kuralı ile durum gözlem olasılıklarına dönüştürülmektedirler.

$$\arg \max_{Y_t} \frac{P_c(Y_t | X_t)}{P(Y_t)} = \arg \max_{Y_t} P_c(X_t | Y_t) \quad (15)$$

HMM'e Viterbi algoritmasının uygulanması, olasılığı en yüksek (en olası) bölütlemeyi verecektir. Durum geçiş olasılıkları ile gözlem olasılıklarının dinamik aralıklarını idare edebilmek için literatürde genellikle yapıldığı gibi bir ağırlıklandırma işlemi uygulanmıştır.

$$\arg \max_{Y_t} P_c(X_t | Y_t)^\alpha \times P(Y_t)^\beta \quad (16)$$

Bu ifadede $P(Y_t)$ fHELM tarafından kestirilmekte olup, α ve β held-out-set kullanılarak optimize edilmektedirler.

4. BULGULAR

4.1. Veri Setleri

Biçimbilgisel, bürünsel ve sözcüksel özelliklerin çıkarıldığı Türkçe konuşma verileri Boğaziçi Üniversitesi BUSİM Laboratuvarı (<http://www.busim.ee.boun.edu.tr>) tarafından Amerika'nın Sesi Türkçe Yayın Bölümü'nün (<http://voanews.com/turkish>) günlük sabah ve akşam saatlerinde yapmış olduğu yayınlardan kaydedilen herbiri 30 dakika olan (42 adet) toplam 21 saatlik haber yayını içerir. Türkçe konuşma verileri, audio (wav) formatında 32kHz örnekleme frekansında ve 16bit linear PCM de örneklenmişlerdir. 42 adet olan konuşma verileri toplamda 127915 adet kelime içerir. Belirtilen 42 dosyanın biçimbilgisel, bürünsel ve sözcüksel özelliklerinin çıkarılması, cümle bölütleme işleminde sınıflandırıcıların veri seti üzerinden eğitilmesi (training set), geliştirme seti (development set/held-out set) üzerinden bazı parametrelerin ayarlanması ve bir test seti (test set) üzerinde de test edilerek performans değerlendirilmelerinin yapılabilmesi için 3 bölüme ayrılması gerekmiştir.

Bu bağlamda, yukarıda sözü edilen 42 adet veri dosyası, aşağıda verilen bölümlere ayrılmıştır.

- 1) 22 adet veri dosyasında toplam 97330 kelimededen oluşan eğitim seti (training set)
- 2) 5 adet veri dosyasında toplam 14897 adet kelimededen oluşan geliştirme seti (development set)
- 3) 5 adet veri dosyasında toplam 15688 adet kelimededen oluşan test seti (test set)

Eğitim setinin kelime haznesi büyüklüğü 19328 olup geliştirme setinde yer alan kelime haznesindeki %33.5 ve test set te yer alan kelime haznesindeki %35.8 oranında kelime eğitim setinde bulunmamaktadır.

Böylece, toplamda eğitim ve geliştirme setlerinde 128,005 adet kelime bulunmaktadır. Bu kelimelerin %6.76 oranındaki kısmı biçimbilgisel analiz edici tarafından analiz edilmemiştir. Bunun nedeni bu orandaki verinin içerisinde özel isimlerin, yabancı kişi isimlerinin, şehir adlarının ve verideki yazım hatalarının yer almasıdır. Geriye kalan kelimeler ortalama kelime başına 1.95 adet ayrımsallaşmaya sahip olmuşlardır. Bu analiz edilme oranı, sadece en son çekimsel grup (en son IG) dikkate alındığında kelime başına 1.83 e düşmektedir. Aynı oran en son çekimsel grup un ögesi (son IG nin ögesi) dikkate alındığında 1.30 a düşmektedir. Tablo 2 de bu oranların yanısıra kelime-ifade belirsizliği olmama yüzdeleri de gösterilmektedir.

Tablo 2: Farklı Düzeylerdeki Biçimbilgisel Özelliklerin Kelime-İfade Belirsizlik İstatistikleri: Herbir kelime için biçimbilgisel analiz edici tarafından kelime başına ortalama ayırmsallaştırılma sayıları ve tek bir ayırmsallaştırılmaya sahip kelimelerin (kelime-ifade belirsizliği olmayan kelimeler) oranları

Biçimbilgisel Özellik	Ortalama ayırıştırma (parse)/ Kelime	Kelime-ifade belirsizliğinin olmama (unambiguity) yüzdesi
Tüm biçimbilgisel analiz	1.95	37.0
Son çekimsel grup (IG)	1.83	39.5
Son IG nin ögesi (POS)	1.30	62.9

4.2. Değerlendirme Ölçütleri (Evaluation Methods)

Performans değerlendirmesi için Otomatik konuşma tanıyıcısının çıktısı olan zorla hizalama (forced alignment) üzerinde F measure ve NIST hata ölçütleri kullanılmıştır (Arisoy ve diğerleri, 2009). NIST hata ölçütü, hatalı bir biçimde sınıflandırılmış kelime sınırlarının sayısının referans cümle sınırlarının sayısına oranı olarak tanımlanmaktadır (Liu ve diğerleri, 2006).

$$NIST = \frac{f_n + f_p}{t_p + f_n} \quad (17)$$

Burada f_n , f_p ve t_p sırası ile yanlış negatif, yanlış pozitif ve doğru pozitif olarak adlandırılır.

F-measure (kesinlik) precision ve yakalama oranı (recall) nin harmonik ortalamasıdır.

$$F - measure = \frac{2 \times precision \times recall}{precision + recall} \quad (18)$$

Burada $precision = \frac{t_p}{t_p + f_p}$ ve $recall = \frac{t_p}{t_p + f_n}$ olarak tanımlanmaktadır.

4.3. Sözcüksel ve Biçimbilgisel Özellikler ile Yapılan Deneyler

Tüm sınıflandırma metodları için sadece sözcüksel bilgiyi kullanan baseline ile tüm sonuçlar karşılaştırılmıştır. Tablo 3 ve 4, Boosting, CRF ve sözcüksel, biçimbilgisel ve pseudo-biçimbilgisel (kelimelerin son üç harfi) özellikleri kullanan HELM, fHELM sonuçlarını göstermektedir. Sadece sözcüksel özelliklerin kullanıldığı durumda HELM diğer metodlara göre üstünlük sağlamaktadır. Bunun bir nedeni, Türkçe nin çekimli doğasından dolayı çok sayıda sözcüksel özelliğin ortaya çıkmasıdır.

Özellik setlerine biçimbilgisel ve pseudo-biçimbilgisel özellikler eklendiğinde tüm sınıflandırıcılarda belirgin bir geliştirim gözlenmiştir.

İlginç olan bir başka gözlemimizde, hem biçimbilgisel hem de pseudo-biçimbilgisel özelliklerin kullanılması ile elde edilen Boosting performansının bu özelliklerden sadece birinin kullanılması ile elde edilen Boosting performansına göre düşmesidir. CRF kararlı bir biçimde Boosting den biraz daha iyi sonuçlar vermiştir. Sadece sözcüksel özellikler kullanıldığında HELM ile karşılaştırıldığında

fHELM in hata oranının göreceli olarak %26 oranında azaldığı belirlenmiştir. Bu durum üretken dizi sınıflandırması için fHELM in etkinliğini göstermiştir. Dahası, göreceli olarak NIST hata oranı azalmaları biçimbilgisel özellikler ile Boosting (%44) ve CRF (%50) de dahi daha fazladır. Bu sonuçlar Tablo 3 ve 4 de gösterilmektedir.

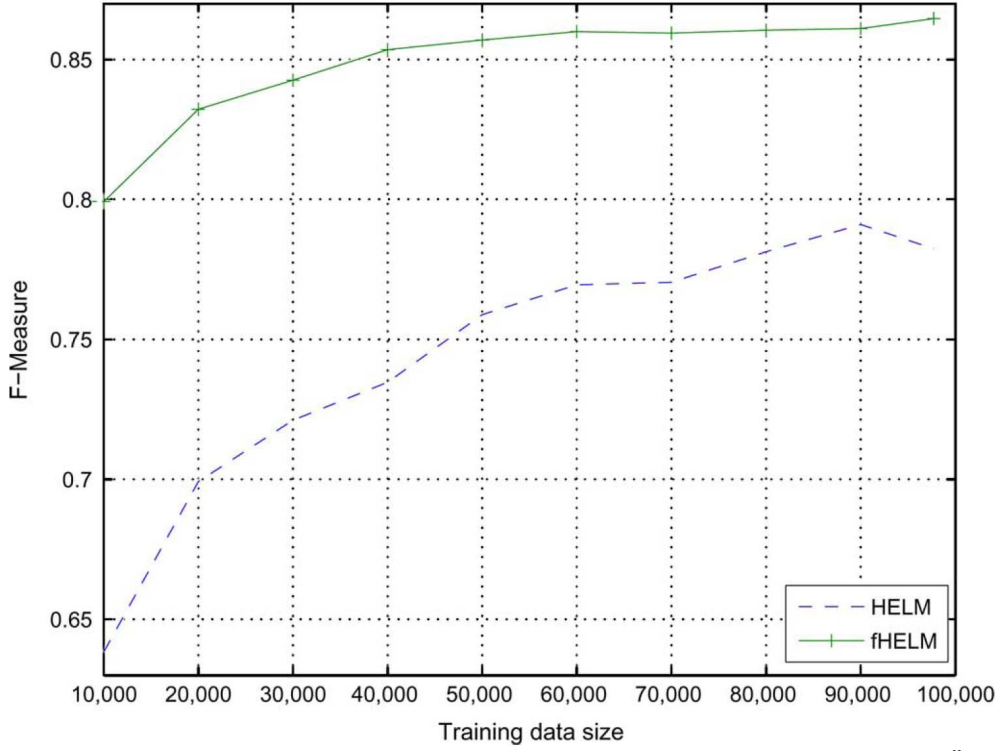
Tablo 3: Sözcüksel (L), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özelliklerin Boosting, CRF ve HELM/fHELM ile Kullanılması ile Elde Edilen F-Measure Sonuçları

Sınıflandırıcı	F			
	L	L+M	L+PM	L+M+PM
Boosting	0.749	0.884	0.853	0.869
CRF	0.756	0.887	0.864	0.891
HELM	0.782	-	-	-
fHELM	-	0.865	0.862	-

Tablo 4: Sözcüksel (L), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özelliklerin Boosting, CRF ve HELM/fHELM ile Kullanılması ile Elde Edilen NIST Hata Ölçütü Sonuçları

Sınıflandırıcı	NIST			
	L	L+M	L+PM	L+M+PM
Boosting	%44.0	%24.7	%30.0	%26.5
CRF	%43.3	%24.0	%26.0	%21.7
HELM	%36.7	-	-	-
fHELM	-	%25.9	%27.1	-

Çeşitli miktarlarda eğitim verisi olduğunda biçimbilgisel bilginin etkisini görebilmek amacı ile HELM ve fHELM için öğrenme eğrileri de çizilmiştir. Şekil 11 de gösterildiği üzere HELM ve fHELM arasındaki F-measure farkı, beklendiği üzere daha az eğitim verisi kullanıldığında artmaktadır. Örneğin; tüm özellik seti yerine sadece 10,000 adet örnek kullanıldığında F-measure ikiye katlanmaktadır.



Şekil 11. Çeşitli Eğitim Veri Seti Büyüklükleri ile HELM ve fHELM için F-Measure Öğrenme Eğrileri

Tablo 5 sözcüksel ve biçimbilgisel özelliklerin her ikisi de kullanıldığında ayırdedici ve üretken dizi sınıflandırma metodlarının kombinasyonları ile elde edilen sonuçları göstermektedir. Performans aşağı yukarı sadece ayırdedici sınıflandırıcıların kullanılması ile elde edilen performans ile aynıdır. Bunun olası bir nedeni HELM den gelen bilginin zaten katılmış olmasıdır.

Tablo 5: Sözcüksel (L) ve Biçimbilgisel (M) Özellikleri Kullanan Boosting ve CRF in fHELM ile Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları

Sınıflandırıcı	F	NIST
Boosting(L+M) + fHELM(L+M)	0.879	%23.8
CRF(L+M) + fHELM(L+M)	0.890	%21.5

4.4. Bürünsel Özellikler ile Yapılan Deneyler

Bürünsel özelliklerin cümle bölütlemesi için istatistiksel olarak bağlantısız bir bilgi sağlaması beklendiğinden, öncelikle sadece bürünsel bilgiyi kullanan Boosting ve CRF sınıflandırıcıların performanslarının etkinliği ölçülmüştür. Performansların sadece sözcüksel ve biçimbilgisel özelliklerin kullanılması ile elde edilen performanslara benzerlik gösterdiği tespit edilmiştir. Bu durum, esasen İngilizce için tasarlanmış bürünsel özelliklerin yararlarını ve kullanılabilirliğini göstermiştir.

Bürünsel bilginin sözcüksel bilgi ile birleştirilebilmesi için 2 farklı yaklaşım uygulanmıştır. İlk yaklaşımda, sadece bürünsel özellikleri kullanarak eğitilen sınıflandırıcı, fHELM ile birleştirilmektedir. Tablo 6 da elde edilen sonuçlar gösterilmektedir. Burada şu durum not edilmelidir ki, birleştirme işleminden önce Boosting ve fHELM karşılaştırılabilir performanslara sahiptirler (NIST hata oranları sırası ile %27.2 ve %25.9). Hibrid model, NIST hata oranını göreceli olarak %39 oranında düşürmektedir (%25.9 dan %15.8 e). Elde edilen bu sonuçlar göstermiştir ki, 2 farklı set

tarafından sağlanan bilgi tamamlayıcı niteliktedir. Bu biraz da özellikle haber verilerinde, bu verilerin doğası gereği, gazetecilerin yada konuşmacının cümle sınırlarını özellikle bürün ile vurgulamasından yada işaretlemesinden kaynaklanmaktadır. İkinci yaklaşım, bürünsel özelliklerin sözcüksel ve biçimbilgisel bilgi ile birlikte Boosting ve CRF sınıflandırıcılarda kullanılmasıdır. Tablo 7 ve 8 elde edilen sonuçları göstermektedir. Görüldüğü üzere her iki sınıflandırıcı için de performans belirgin bir biçimde geliştirilmiştir. Bu yaklaşım, ek olarak pseudo-biçimbilgisel özelliklerin de eklenmesi ile, önceki yaklaşıma oranla daha iyi bir performans sergilemiştir.

Tablo 6: Sadece Bürünsel (P) Özellikleri Kullanan Boosting ile Sözcüksel (L) ve Biçimbilgisel (M) Bilgiyi Kullanan fHELM in Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları

Sınıflandırıcı	F	NIST
CRF(P)	0.874	%24.7
Boosting(P)	0.862	%27.2
Boosting(P) + fHELM(L+M)	0.919	%15.8

Tablo 7: Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Elde Edilen F-Measure Sonuçları

Sınıflandırıcı	F			
	L+P	L+M+P	L+PM+P	L+M+PM+P
Boosting	0.894	0.922	0.918	0.927
CRF	0.895	0.921	0.916	0.923

Tablo 8: Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Elde Edilen NIST Hata Oranları

Sınıflandırıcı	NIST			
	L+P	L+M+P	L+PM+P	L+M+PM+P
Boosting	%20.4	%16.5	%15.8	%14.7
CRF	%20.2	%14.6	%16.9	%15.3

Deneylerimizin final setinde, tüm özellik setini kullanan Boosting ve CRF sınıflandırıcılar ile fHELM birleştirilmiştir. Tablo 9 elde edilen sonuçları göstermektedir. Sonuçlardan da görüleceği üzere Boosting i içeren hibrid model geliştirim yapmazken CRF i içeren bir miktar geliştirim gerçekleştirmiştir.

Tablo 9: fHELM in Sözcüksel (L), Bürünsel (P), Biçimbilgisel (M) ve/veya Pseudo-Biçimbilgisel (PM) Özellikleri Kullanan Boosting ve CRF ile Birleştirilmesi ile Elde Edilen F-Measure ve NIST Hata Oranları

Sınıflandırıcı	F	NIST
Boosting(L+P+M+PM) + fHELM(L+M)	0.925	%14.8
CRF(L+P+M+PM) + fHELM(L+M)	0.926	%14.9

5. TARTIŞMA VE SONUÇ

Bu projede Türkçe konuşma verileri için Bürünsel bilginin (özellikler) çıkarılması ve çıkarılan bürünsel bilginin özellikle Türkçe cümle bölütleme alanında kullanılması gerçekleştirilmiştir.

Ayrıca, bürünsel bilginin SRILM ve Praat ile çıkarılmasının yanısıra Türkçe için sözcüksel, biçimbilgisel ve pseudo-biçimbilgisel (kelimelerin son üç harfi) özellikler de çıkarılmıştır.

Çıkarılan bürünsel özellikler kelime bazında olup genellikle İngilizce için tasarlanmış özelliklerden oluşmaktadır. Bürünsel özellikler ayrıca hece düzeyinde de modellenerek geliştirilmişlerdir. Örneğin

cümlenin son kelimesindeki biçimbilgisel belirsizlik, anlambirim düzeyindeki bürünsel özelliklerin kullanılması ile ortadan kaldırılmıştır. Biçimbilgisel belirsizlik esasen fHELM için bir problem oluşturmaktadır. Gelecekteki çalışmalarımız biçimbilgisel açıklıştırmanın etkilerinin araştırılmasını da kapsayacaktır.

Ayrıca birçok dil işleme uygulamasının temeli olan cümle bölütleme işlevi için Boosting CRF gibi ayırdedici farklı sınıflandırıcı yapıları ile HELM ve fHELM gibi üretken dil modellerini birleştiren hibrid yaklaşımlar denenmiştir. Başka bir deyiş ile Türkçe cümle bölütleme için sözcüksel, bürünsel ve biçimbilgisel özellikleri kullanan, üretken, ayırdedici ve hibrid sınıflandırma yöntemleri ortaya konmuştur.

Bu projenin elde edilen en önemli sonuçlarından biri sözcüksel, biçimbilgisel ve bürünsel bilginin kullanıldığı ayırdedici sınıflandırma yaklaşımlarının Türkçe cümle bölütleme alanında en iyi sonuçları verdiğinin anlaşılmasıdır. CRF sınıflandırıcılar sadece bürünsel ve sözcüksel bilgi ile iyi sonuçlar verirken, Boosting biçimbilgisel bilgiden daha çok yararlanmaktadır. Bu durum büyük olasılıkla boosting in bilinmeyen özellik değerlerini işleme yeteneğinden kaynaklanmaktadır.

Deneylerimizde her ne kadar ayırdedici modeller tek başlarına en iyi sonuçları vermiş olsalar da üretken modeller cümle bölütleme için potansiyel kullanıma sahiptirler. Boosting ve CRF modelleri HELM lerde yer alması zor olan birçok bürünsel özelliğe erişebilmektedirler. Genellikle, sadece az miktarda bir konuşma verisinin mevcut olduğu bir durumda, gazete gibi yazılı text kaynaklarında belirgin bir çoklukta text verisi mevcuttur. Her iki veri kaynağından yararlanmak için, İngilizce dilinde cümle bölütlemesi için eğitilen modellerden kazandığımız deneyim göstermektedir ki, genel olarak boosting ve CRF modeller konuşma verisinden, HELM ler ise text verisinden eğitilmektedirler. Bu iki model bir HELM de test süresince birleştirilmektedirler. Bu işlem İngilizce için en iyi sonucu vermektedir. Bu çalışmada özellikle Türkçe cümle bölütleme alanında önemli sonuçlar elde edilmiştir.

Daha sonra yapılacak çalışmalardan bir diğeri de, Praat ile elde edilen bürünsel özelliklerin yine Praat ile dizayn edilecek yeni bürünsel özellikler ile birleştirilerek Türkçe cümle bölütleme performansının daha da geliştirilmesini sağlamaktır. Ayrıca bu çalışmada özellik çıkarımı için kullanılan ASR forced alignment verileri yerine manual olarak işaretlenmiş veriler üzerinde de aynı deneyler tekrar edilecektir. Farklı bir çalışma olarak, sözcüksel, bürünsel, biçimbilgisel ve pseudo-biçimbilgisel özellikleri işleyen ayırdedici sınıflandırıcılar ile co-training yaklaşımları birleştirilerek cümle bölütleme performansları da incelenecektir.

Yeni bürünsel özelliklerin tasarlanması ve özellikle cümle bölütleme alanında kullanılması konusunda, bu projede görev almış olan bursiyer yüksek lisans öğrencisinin tez çalışmalarına başlanmıştır. Tez, özelliklerin tümünü kullanmak yerine, varolan özellikler içerisinden cümle bölütleme için en uygun özelliklerin seçilmesi (feature selection) yöntemlerini de içerecektir.

Bu proje ile birlikte özellikle, SRI-International, Speech Technology and Research Laboratory ve International Computer Science Institute (ICSI), Speech Group ile daha önce başlatılmış araştırma işbirliğinin sürdürülmesi ve geliştirilmesi sağlanmıştır.

Ayrıca, Sözü edilen enstitüler ile yapılan ve süregelen araştırmalarımız çerçevesinde proje kapsamında bir yayın da çıkarılmıştır (Guz, U., Favre, B., Hakkani-Tür, D., Tur, G., “Generative and Discriminative Methods using Morphological Information for Sentence Segmentation of Turkish”, Special Issue on Processing Morphologically Rich Languages, IEEE Transactions on Audio, Speech and Language Processing, Volume 17, Issue 5, pp. 895-903, July 2009).

TEŞEKKÜR

Araştırmacılar, özellikle bu projenin gerçekleştirilmesinde birlikte çalışılan SRI-International, Speech Technology and Research Laboratory (STAR Lab.) de Senior Scientist olarak görev yapan Sayın Gökhan Tür ve International Computer Science Institute (ICSI), Speech Group da Senior Scientist olarak görev yapan Sayın Dilek Hakkani Tür'e, ICSI de postdoc olarak görev yapan Benoit Favre'e, Biçimbilgisel Analiz araçlarını paylaşan Sayın Prof. Dr. Kemal Oflazer'e, Türkçe veri setlerini ve forced alignment verilerini paylaşan Sayın Yrd. Doç. Dr. Murat Saraçlar ve Sayın Sıdıka Parlak'a, J. William Fulbright Post Doctoral Research Fellowship ve TÜBİTAK-BİDEB Yurtdışı Doktora Sonrası Araştırma Bursları çerçevesinde araştırma ortamını sağlayan SRI- International, Speech Technology and Research Laboratory, Menlo Park, CA, USA ve Projenin yürütüldüğü ve direktörlüğünü Prof. Nelson Morgan'ın yaptığı International Computer Science Institute, Speech Group, Berkeley, CA, USA ile Işık Üniversitesi Bilimsel Araştırma Fonu ve Işık Üniversitesi, Mühendislik Fakültesi, Elektronik Mühendisliği Bölümü'ne teşekkür ederler.

REFERANSLAR

Altun, Y., *Discriminative Methods for Label Sequence Learning*, (Ph.D. Dissertation), Department of Computer Science, Brown University, (2005).

Ang, J., Liu, Y., Shriberg, E., Automatic Dialog Act Segmentation and Classification in Multiparty Meetings, In Proc. of ICASSP, Philadelphia-USA, (2005) pp: 1061–1064.

Arisoy E., Can D., Parlak S., Sak H., Saraçlar M., Turkish Broadcast News Transcription and Retrieval, *IEEE Trans. Audio, Speech, Lang. Process., Special Iss. Process. Morphol. Rich Lang.*, 17, 5, pp. 874-883, (2009).

Beeferman, D., Berger, A., Lafferty, J., Statistical Models for Text Segmentation, *Machine Learning, Special Issue on Natural Language Learning*, 34, 177-210, (1999).

Bilmes, J.A., Kirchhoff, K., Factored Language Models and Generalized Parallel Backoff, In Proc. Human Lang. Technol. Conf. (HLT)-Conf. North Amer. Chap. Assoc. Comput. Lingist. (NAACL), Edmonton, AB-Canada, (2003).

Boersma, P., Weenink, D., Praat: Doing Phonetics by Computer, 2009.

Cuendet, S., Hakkani-Tür, D., Shriberg, E, Fung, J., Favre, B., Crossgenre Feature Comparisons for Spoken Sentence Segmentation, In Proc. IEEE Int. Conf. Semantic Comput. (ICSC), Irvine, CA-USA, (2007) pp: 265-274.

Cuendet, S., Hakkani-Tür, D., Tur, G., Model Adaptation for Sentence Segmentation from Speech, In Proc. IEEE/ACL Spoken Lang. Technol. (SLT) Workshop, Palm Beach-Aruba, (2006) pp:102-105.

Duh, K. Kirchhoff, K., Automatic Learning of Language Model Structure, In Proc. Int. Conf. Comput. Lingist. (COLING), Geneva-Switzerland, (2004).

Friedman, S., Fox, D., Pasula, H., Voronoi Random Fields: Extracting the Topological Structure of Indoor Environments via Place Labeling, In Proc. Int. Joint Conf. Artif. Intell. (IJCAI), Hyderabad-India, (2007).

Gotoh, Y., Renals, S., Sentence Boundary Detection in Broadcast Speech Transcripts, In Proc. of ISCA Workshop: Automatic Speech Recognition: Challenges for the New Millennium, ASR-2000, Paris-France, (2000) pp: 228–235.

- Guz, U., Cuendet, S., Hakkani, Tür, D., Tur, G., Co-training Using Prosodic and Lexical Information for Sentence Segmentation, In Proc. Interspeech, Antwerp, Belgium, (2007).
- Guz, U., Favre, B., Hakkani Tür, D., Tur, G., Generative and Discriminative Methods using Morphological Information for Sentence Segmentation of Turkish, *IEEE Trans. Audio, Speech, Lang. Process., Special Iss. Process. Morphol. Rich Lang.*, 17, 5, pp. 895-903, (2009).
- Hakkani-Tür, D., Tür, G., Stolcke, A., Shriberg, E., Combining Words and Prosody for Information Extraction from Speech, In Proceedings of the 6th European Conference on Speech Communication and Technology (EUROSPEECH'99), Budapest-Hungary, (1999).
- Hakkani-Tür, D., Oflazer, K., Tur, G., Statistical Morphological Disambiguation for Agglutinative Languages, In Proc. 18th Int. Conf. Comput. Lingist. (COLING), Saarbruecken-Germany, (2000).
- Hakkani-Tür D., Oflazer K., Tur G., Statistical Morphological Disambiguation for Agglutinative Language, *Computer and the Humanities*, 36, 4, 381-410, (2002).
- Hankamer, J., *Morphological parsing and the lexicon, Lexical Representation and Process*, ed: W. Marslen-Wilson, Cambridge, MA: MIT Press, (1989), pp: 392-408.
- Hearst M.A., Texttiling: Segmenting Text into Multiparagraph Subtopic Passages, *Computational Linguistics*, 23, 1, 33-64 (1997).
- Hirschberg, J., Grosz, B.J., Intonational Features of Local and Global Discourse, In Proceedings of the Workshop on Spoken Language Systems, DARPA, Arden House, New York-USA, (1992) pp: 441-446.
- Hirschberg, J., Nakatani, C., A Prosodic Analysis of Discourse Segments in Direction-Giving Monologues, In Proc. ACL, Santa Cruz, CA, (1996) pp: 286-293.
- Huang, Z., Chen, L., Harper, M., An Open Source Prosodic Feature Extraction Tool, In Proc. of Language Resource and Evaluation Conference, Genoa, Italy, (2006).
- Huang, J., Zweig, G., Maximum Entropy Model for Punctuation Annotation from Speech, In Proc. 7th Int. Conf. Spoken Lang. Process. (ICSLP), Denver, CO-USA, (2002).
- Lafferty, J.D., McCallum, A., Pereira, F.C.N., Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, In Proc. 18th Int. Conf. Mach. Learn. (ICML), San Francisco, CA-USA, (2001) pp: 282–289, Morgan Kaufmann.
- Levi, S.V., Limitations on Tonal Crowding in Turkish Intonation, *Phonologica: 9th International Phonology Conference*, Vienna (2002).
- Levi, S.V., Glides, Laterals, and Turkish Vowel Harmony, Proceedings from the 37th Meeting of the Chicago Linguistics Society, (2001) pp: 379-394.
- Levi S.V. Acoustic Correlates of Lexical Accent in Turkish, *Journal of the International Phonetic Association*, 35, 73-97, (2005).
- Liu Y., Shriberg E., Stolcke A., Hillard D., Ostendorf M., Harper M., Enriching Speech Recognition with Automatic Detection of Sentence Boundaries and Disfluencies, *IEEE Trans. Audio, Speech, Lang. Process.*, 14, 5, 1526–1540, (2006).

Liu, Y., Shriberg, E., Stolcke, A., Peskin, B., Ang, J., Hillard, D., Ostendorf, M., Tomalin, M., Woodland, P., Harper, M., Structural Metadata Research in the EARS Program, In Proc. of ICASSP, Philadelphia-USA, (2005) pp: 957–960.

Liu, Y., Stolcke, A., Shriberg, E., Harper, M., Using Conditional Random Fields for Sentence Boundary Detection in Speech, In Proc. Annu. Meeting Assoc. Comput. Lingist. (ACL), Ann Arbor, MI-USA, (2005) pp: 451-458.

Liu, D., Zong, C., Utterance Segmentation Using Combined Approach Based on Bi-directional N-gram and Maximum Entropy, in Proc. of ACL-2003 Workshop: The Second SIGHAN Workshop on Chinese Language Processing, Sapporo-Japan, (2003) pp: 16–23.

Kneser, R., Ney, H., Improved Clustering Techniques for Class-Based Statistical Language Modeling, in Proc. Eur. Conf. Speech Commun. Technol. (Eurospeech), Berlin-Germany, (1993).

Kolar, J., Svec, J., Psutka, J., Automatic Punctuation Annotation in Czech Broadcast News Speech, In Proc. of 9th Conference Speech and Computer, St. Petersburg-Russia, (2004) pp: 319–325.

Kozima, H., Text Segmentation Based on Similarity Between Words, In 31st Annual Meeting of the ACL, Columbus Ohio-USA, (1993) pp: 286-288.

Kuo H.-K. J., Gao Y., Maximum Entropy Direct Models for Speech Recognition, *IEEE Transaction Audio, Speech and Language Processing*, 14, 3, 873–881, (2006).

Manning, C.D., Schütze, H., *Foundations of Statistical Natural Language Processing*, Cambridge, MA: MIT Press, (1999).

McCallum, A., Freitag, D., Pereira, F., Maximum Entropy Markov Models for Information Extraction and Segmentation, In Proc. Int. Conf. Mach. Learn. (ICML), Palo Alto, CA-USA, (2000).

Oflazer K., Two-Level Description of Turkish Morphology, *Literary and Linguistic Computing*, 9, 2, 137-148, (1994).

Oflazer, K., Dependency Parsing with an Extended Finite State Approach, In Proc. 37th Annu. Meeting Assoc. Comput. Lingist., Maryland-USA, (1999) pp:254-260.

Oflazer, K., Gocmen, E., Bozsahin, C., An Outline of Turkish Morphology, Technical Report, Middle East Technical University, (1995).

Oskay, B., Salor, Ö., Özkan, Ö., Demirekler, M., Çiloğlu, T., Determination of Prosody from Turkish Text and its Application to Speech Synthesis, SIU 2001, Northern Cyprus, (2001) pp:238-243.

Passonneau R.J., Litman D.J., Discourse Segmentation by Human and Automated Means, *Computational Linguistics*, 23, 1, 103-139, (1997).

Roark, B., Liu, Y., Harper, M., Stewart, R., Lease, M., Snover, M., Shafran, I., Dorr, B., Hale, J., Krasnyanskaya, A., Yung, L., Reranking for Sentence Boundary Detection in Conversational Speech, In Proc. of ICASSP, Toulouse-France, (2006) pp: 545–548.

Sayli, O., Arslan, L.M., Türkçedeki Seslerin Süre Özellikleri, DilBilim Araştırmaları, Boğaziçi Üniversitesi Yayınevi, İstanbul, (2003).

Schapire R.E., Singer Y., Boostexter: A Boosting-Based System for Text Categorization, *Mach. Learn.*, 39, 2/3, 135–168, (2000).

Shriberg E., Stolcke A., Hakkani-Tür D., Tür G., Prosody-Based Automatic Segmentation of Speech into Sentences and Topics, *Speech Communication*, 32, 1-2, 127–154, (2000).

Stolcke, A., SRILM-An Extensible Language Modeling Toolkit, In Proc. Int. Conf. Spoken Lang. Process. (ICSLP), Denver, CO-USA, (2002) pp: 901-904.

Stolcke, A., Shriberg, E., Statistical Language Modeling for Speech Disfluencies, In Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Atlanta, GA-USA, (1996) pp: 405-409.

Stolcke, A., Shriberg, E., Hakkani-Tür, D., Tür, G., Rivlin, Z., Sönmez, K., Combining Words and Speech Prosody for Automatic Topic Segmentation. In Proceedings of DARPA Broadcast News Transcription and Understanding Workshop, Herndon, Virginia-USA, (1999) pp: 61-64.

Tur, G., A Statistical Information Extraction System for Turkish, (Ph.D. Dissertation), Department of Computer Science, Bilkent University, Ankara, Turkey, (2000).

Tur, G., Hakkani-Tür, D., Stolcke, A., Shriberg, E., Integrating Prosodic and Lexical Cues for Automatic Topic Segmentation. *Computational Linguistics*, 27, 31-57, (2001).

Waibel, A., Geutner, P., Mayfield-Tomokiyo, L., Schultz, T., Woszczyna, M., Multilinguality in Speech and Spoken Language Systems, *Proceedings of the IEEE Special Issue on Spoken Language Processing*, 88, 8, 1297–1313, (2000).

Zimmermann, M., Hakkani-Tür, D., Fung, J., Mirghafori, N., Gottlieb, L., Liu, Y., Shriberg, E., The ICSI+ Multi-Lingual Sentence Segmentation System, In Proc. Int. Conf. Spoken Lang. Process. (Interspeech), Pittsburgh, PA-USA, (2006).

Zimmermann, M., Liu, Y., Shriberg, E., Stolcke, A., A* Based Segmentation and Classification of Dialog Acts in Multiparty Meetings, In Proc. of ASRU, San Juan-Puerto Rico, (2005) pp: 215–219.

Zong, C., Ren, F., Chinese Utterance Segmentation in Spoken Language Translation, In the 4th International Conference on Computational Linguistics and Intelligent Text Processing, Mexico City-Mexico, (2003) pp: 516–525.

Generative and Discriminative Methods Using Morphological Information for Sentence Segmentation of Turkish

Umit Guz, *Member, IEEE*, Benoit Favre, *Member, IEEE*, Dilek Hakkani-Tür, *Senior Member, IEEE*, and Gokhan Tur, *Senior Member, IEEE*

Abstract—This paper presents novel methods for generative, discriminative, and hybrid sequence classification for segmentation of Turkish word sequences into sentences. In the literature, this task is generally solved using statistical models that take advantage of lexical information among others. However, Turkish has a productive morphology that generates a very large vocabulary, making the task much harder. In this paper, we introduce a new set of morphological features, extracted from words and their morphological analyses. We also extend the established method of hidden event language modeling (HELM) to factored hidden event language modeling (fHELM) to handle morphological information. In order to capture non-lexical information, we extract a set of prosodic features, which are mainly motivated from our previous work for other languages. We then employ discriminative classification techniques, boosting and conditional random fields (CRFs), combined with fHELM, for the task of Turkish sentence segmentation.

Index Terms—Prosodic and lexical information, sentence segmentation, Turkish morphology.

I. INTRODUCTION

MANY useful results have been obtained by applying statistical language modeling techniques to English (and similar languages)—in speech recognition, parsing, word sense disambiguation, part-of-speech (POS) tagging, etc. However, languages that display a substantially different behavior

than English, like Turkish, Czech, Hungarian (in that, they have agglutinative or inflective morphology and relatively free constituent order) have not been studied extensively using statistical approaches. In these languages, due to their richer morphology, the vocabulary size for a given corpus size is much larger than other languages [1], [2]. While this causes a data sparseness problem, the statistical models that look at only words are also blind to the information encoded in the morphology. Usually, the combined effect of these problems is a reduction in language processing performance.

Similarly, in spite of all the advances in discriminative classification techniques in the machine learning community, discriminative sequence classification is still a challenge. Researchers have proposed various techniques such as maximum entropy Markov models [3] or conditional random fields [4], [5]. However these techniques are typically not very successful in efficiently handling continuous valued features hence it is a common practice to discretize such features [6]. On the other hand, for generative sequence modeling, hidden Markov models (HMMs) still dominate the field; however usually only one level of states is employed. For example, for automatic speech recognition (ASR), typically word sequences are modeled for the language model [7]. With the advances in graphical models, factored language models (FLMs) handling bundles of features for each sample have been proposed [8]. FLMs have been successfully used for ASR of inflectional languages such as Arabic [9].

In this paper, we address the problem of exploiting morphological information in statistical classification models for sentence segmentation of Turkish speech. Our contributions are threefold: First, we extend the hidden event language models to factored hidden event language models and combine them with classification models. Second, we introduce a new set of morphological features, extracted from words and their morphological analyses. Third, we extract a set of prosodic features, which are mainly motivated from our previous work for other languages, for the task of Turkish sentence segmentation.

In the next section, we briefly summarize the related work on sentence segmentation of speech. Then, we present our approach, mainly the generative, discriminative, and hybrid modeling techniques, and we describe the feature sets for segmenting Turkish speech into sentences. Finally, we provide experimental results showing the effectiveness of the proposed techniques for this morphologically rich language.

Manuscript received August 11, 2008; revised November 17, 2008. Current version published June 17, 2009. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) CALO (NBCHD-030010) program, the DARPA GALE (HR0011-06-C-0023) program, the Scientific and Technological Research Council of Turkey (TUBITAK) fundings at SRI and ICSI, (TUBITAK CAREER Project 107E182, Extracting and Using Prosodic Information for Turkish Spoken Language), and the Isik University Research Fund (Project 05B304). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mari Ostendorf.

U. Guz is with the International Computer Science Institute (ICSI), Berkeley, CA 94704 USA, and also with Isik University, 34980 Istanbul, Turkey (e-mail: guz@icsi.berkeley.edu).

B. Favre and D. Hakkani-Tür are with the International Computer Science Institute (ICSI), Berkeley, CA 94704 USA (e-mail: favre@icsi.berkeley.edu; dilek@icsi.berkeley.edu).

G. Tur is with SRI International, Menlo Park, CA 94025 USA (e-mail: gokhan@speech.sri.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASL.2009.2016393

II. SENTENCE SEGMENTATION

Sentence segmentation for speech aims at finding sentential unit boundaries in a stream of words, output by a speech recognizer. It is a preliminary step for many speech processing applications, such as parsing, machine translation and information extraction, which generally assume the presence of punctuation. One typically leverages the word sequence generated by a speech recognizer and prosodic cues such as pitch, energy and pause duration in order to segment the audio in sentences.

Previous work on sentence segmentation has considered this task as a word boundary classification problem, by determining whether or not two consecutive words are separated by a sentence boundary. The features used are mainly limited to words neighboring the boundary [10]–[12], with the exception of [13], who included a reranking phase using sentence-level features. [14] showed that for segmentation of speech into sentences, prosodic and lexical cues provide complementary information. [15] evaluated different modeling approaches (HMM, maximum entropy, and conditional random fields) and various prosodic and textual features, in both conversational telephone speech and broadcast news speech.

There is also related work for sentence boundary detection in languages other than English, for example, in Czech [16] where an HMM approach was used, and in Chinese [17], [18] where a maximum entropy classifier was used with mostly textual features. [12] used lexical and prosodic features with several classifiers, including maximum entropy and boosting for English and Mandarin. [19] investigated the use of the same set of prosodic features and feature selection for English, Mandarin, and Arabic. [20] used syntactic dependency structure and support vector machines for sentence boundary detection in Japanese. [21] is the first work that used morphological features for sentence segmentation of Turkish; our work, in a way, extends that work to also include prosodic features and more sophisticated classification models. That study relied only on generative models, i.e., hidden event modeling, and trained two separate language models, one using words, the other using morphological analyses of the words, and then computed the weighted combination of the posteriors obtained from each model.

Sentence segmentation has also been studied according to various other aspects. [22] showed the benefits of speaker-adapted models and [23] focused on domain adaptation. Sentence segmentation can be optimized to improve downstream tasks, such as speech translation [24], [25] or information extraction [26]. For instance, [24] has shown about 10% relative BLEU score improvements for machine translation (MT), when using a sentence segmentation optimized for MT, in comparison to fixed length sentences. [26] has shown similar results for information extraction, and that a 4% relative gain on entity and relation extraction can be obtained by optimizing punctuation for these tasks.

III. APPROACH

In the literature, typically sentence or dialog act segmentation is treated as a boundary classification problem where the goal is

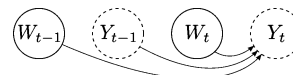


Fig. 1. Conceptual hidden event language model for sentence segmentation.

finding the most likely boundary tag sequence, $Y = Y_1 \dots Y_n$ given the features, $\mathcal{X} = \mathcal{X}_1 \dots \mathcal{X}_k$

$$\arg \max_Y P(Y|\mathcal{X}). \quad (1)$$

To this end, generative, discriminative, or hybrid models have been used. Below we summarize these approaches and explain how we extend them to handle the speech input of morphological languages.

A. Factored Hidden Event Language Models

We propose using factored language models with hidden event language models. Below, first we describe the Hidden Event Language Model (HELM) and the FLM and then describe how we combine them.

1) *Hidden Event Language Models*: The most popular generative model for sentence segmentation is the hidden event language model, as introduced by [27]. The HELM was originally designed for speech disfluencies, such as deletion (DEL) and repetition (REP). The approach was to treat such events as extra meta-tokens. To ease the computation, an imaginary “no disfluency” (NODF) token is inserted between two words, in cases the word preceding the boundary is not part of a disfluency. The following example is a conceptual representation of a sequence with disfluencies:

... she NODF got REP got NODF real NODF lucky ...

For sentence segmentation, sentence boundaries are simply treated as hidden events, and the word sequence is augmented with fictitious sentence boundary tokens (S for sentence boundary, N for else). So an example would be as follows:

... she N got N real N lucky S however N there N were N ...

Note that this is different from using an HMM as is typically done in similar tagging tasks, such as POS tagging [28] or named entity extraction [29]. For sentence segmentation, the conceptual model is depicted in Fig. 1. In this model, one state is reserved for each of the boundary tokens, S and N , and the rest of the states are for generating words. It has been shown that the HELM outperforms the conventional HMM approach, since it allows an explicit point to emit the boundary token, hence can incorporate nonlexical information via combination with other models as presented in the next subsection [14].

The most probable boundary token sequence is obtained simply by Viterbi decoding using only lexical features, i.e., the language model, to model $P(\mathcal{X}, Y)$, where \mathcal{X} and Y represent all the words ($\mathcal{X}_t = (W_t)$) and boundary tokens, respectively,

$$\arg \max_Y P(Y|\mathcal{X}) = \arg \max_Y P(\mathcal{X}, Y) \quad (2)$$

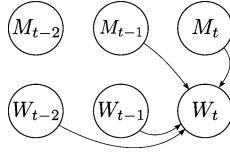


Fig. 2. Example factored language model seen as a directed graphical model over words W and morphological factors M . The arrows indicate the factors used for estimating the probabilities.

2) *Factored Language Models*: Factored language models aim to model a sequence of feature sets, extending the conventional language modeling. In other words, the goal is building probabilistic language models using the subsets of feature sets (or factors).

Factored language models have been successfully used for ASR [9] of inflectional languages, by defining factors or feature sets consisting of surface forms, stems, morphological analyses, etc., of the words.

More formally, the factored language model aims to estimate the probability of a feature set sequence, $\mathcal{X}_1, \dots, \mathcal{X}_n$ instead of a word sequence W_1, \dots, W_n . Here we consider $\mathcal{X}_t = (W_t, M_t)$ where M_t is a morphological feature for word W_t . An example factored language model can be seen in Fig. 2. The current word relies on not only the previous two words but also the current and previous morphological analyses. We provide example feature sets for Turkish in Section IV-C. More formally, it models

$$P(W_t|W_{t-1}, W_{t-2}, M_t, M_{t-1}). \quad (3)$$

Even with lower-order n -gram approximations, since it may be possible to have unseen n -gram sequences, one important issue with FLMs is how to back off to reliably estimate such probabilities. A new generalized parallel back-off technique was proposed to tackle this problem [8]. Basically, the system is given a back-off graph, which denotes the paths for back-off. Paths in this graph can be chosen manually. In the literature, with complex factors, methods based on genetic algorithms have been proposed to choose the optimal back-off graph [30]. The important point is that many back-off paths can be proposed and the system can process them in parallel.

An example back-off dropping the most distant word is defined as follows for factored language models using words W_t and morphological analyses M_t

$$P(Y_t|C_t) = \begin{cases} P_{ML}(Y_t|C_t), & \text{if } N(C_t, Y_t) > \tau \\ \alpha(C_t) \times P_{BO}(Y_t|\hat{C}_t), & \text{otherwise.} \end{cases} \quad (4)$$

where $C_t = W_t, M_t, Y_{t-1}, W_{t-1}, M_{t-1}$ is the original context, $\hat{C}_t = W_t, M_t, Y_{t-1}, M_{t-1}$ is the backed off context, P_{ML} is the standard maximum-likelihood estimate (with smoothing), $N(\cdot)$ is the number of occurrences, and α is used to ensure that the result is still a probability distribution.

3) *From HELM to fHELM*: The factored hidden event language models are straightforward extensions of hidden event language models and factored language models. They combine the strength of factored language models for multifeature sequence modeling with the classification power of hidden event

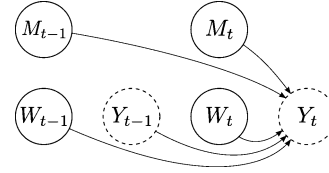


Fig. 3. Example factored language model created for a hidden event language model seen as a directed graphical model over word boundaries Y , words W , and morphological factors M . The arrows indicate the factors used for estimating the probabilities.

language models. Fig. 3 presents the factored hidden event language model topology employed in this paper. The boundary states still exist to potentially build hybrid models (as explained below) and the boundary decision is made according to the following formula:

$$P(Y_t|W_t, M_t, Y_{t-1}, W_{t-1}, M_{t-1}) \quad (5)$$

where Y_t indicates the boundary decision, S or N after the word W_t with a morphological analysis of M_t .

The next step for building an fHELM is creating a back-off graph indicating the possible back-off paths in case the statistics for the desired n -gram are not reliable. In this study, we tried only linear graph back-off (i.e., dropping and forgetting about one factor at a time) and fully connected graph back-off (i.e., backing off to all possible subsets) starting from the most distant feature. The back-off used in the experiments drop the most distant morphological analyses in a trigram language model M_{t-2} and then the most distant words W_{t-2} , and so on. Then standard Viterbi decoding may be employed to find the most probable state sequence, i.e., the boundary decisions given the words and their other features, such as morphological analysis. This results in an elegant method for building a generative classifier when multiple features are used for each sample position. Furthermore, similar to regular HELMs, it is possible to combine the posterior probabilities obtained from other classifiers (preferably discriminative) to improve the performance even more. For example fHELM may exploit the lexical and morphological information and then may be combined with a classifier that uses only prosodic features.

In our experiments, the SRILM [31] toolkit is used for Viterbi decoding and for building the conventional and factored hidden event language models with modified Kneser–Ney smoothing [32].

B. Discriminative Classification Models

One weakness of the hidden event language models is that one can incorporate only streams of discrete features such as words or morphological analyses. To overcome this obstacle, various classification methods have been used in the literature. In a pioneering study, decision trees were used to build segmentation models to improve the performance also by using additional prosodic features [14]. With the advances in discriminative classification algorithms, researchers tried using conditional random fields (CRFs) [33] and boosting [34], and hybrid approaches using boosting and maximum entropy classification algorithms [12].

Our system relies on boundary-wise posterior probabilities $P(Y_t|\mathcal{X}_t)$ provided by two classifiers that can be used independently or jointly.

1) *Boosting*: The first component is an Adaboost [35] classifier that generates posterior probability estimations out of weighted decision stumps (one-level decision trees)

$$P(Y_t|\mathcal{X}_t) = \left[1 + \exp \left(-2m \sum_{i=1}^m w_i s_i(\mathcal{X}_t) \right) \right]^{-1} \quad (6)$$

where $s_i(\cdot)$ is a decision stump (presence of a discrete feature or position relative to a threshold of a continuous feature) over a single feature, w_i is the weight given to that decision stump, and m is the number of decision stumps. Adaboost is trained by iterating over the selection of the best decision stump and reweighing of examples where the overall classifier makes mistakes. The implementation used in our experiments is *icsiboost*.¹ In all our experiments, we used boosting with 1000 iterations.

2) *Conditional Random Fields*: The second component of our system uses CRFs as proposed by [4]. We use chain CRFs to estimate the probability of a sequence of boundary events ($Y = Y_1 \dots Y_n$) given a sequence of observations ($\mathcal{X} = \mathcal{X}_1, \dots, \mathcal{X}_n$).

$$P(Y|\mathcal{X}) = \frac{1}{Z(\mathcal{X})} \exp \left(\sum_{t=1}^n \sum_{i=1}^m \lambda_i s_i(Y_{t-1}, Y_t, \mathcal{X}_t) \right) \quad (7)$$

where

$$Z(\mathcal{X}) = \sum_Y \exp \left(\sum_{t=1}^n \sum_{i=1}^m \lambda_i s_i(Y_{t-1}, Y_t, \mathcal{X}_t) \right). \quad (8)$$

Here, $s_i(\cdot)$ are decision functions that depend on the examples and a clique of boundaries close to Y_t , λ_i is the weight of s_i estimated on training data, and $Z(\mathcal{X})$ is a normalization factor. Note that CRFs give the probability of the sequence of boundary decisions. The forward-backward algorithm can be used to get boundary-level posterior probability estimates.

For our experiments, we use the CRF++ toolkit,² which allows binary decision functions dependent on the current boundary and the previous boundary. Features extracted from \mathcal{X} originate from a neighborhood of the boundary and match the features used with Adaboost, though CRF++ does not handle continuous features,³ and requires them to be quantized. After experimenting with different types of quantization, we observed that using thresholds from the decision stumps learned by Adaboost leads to slightly improved performance, probably due to the fact that the thresholds of decision stumps are already optimized on dividing positive and negative classes and the stumps embed the interaction between features (as in Adaboost training, classifiers are chosen in order to correct errors from previous iterations), thus this method is expected to be better than other quantization schemes. This is similar to the quantization method suggested by [36].

¹<http://code.google.com/p/icsiboost>.

²<http://crfpp.sourceforge.net/>.

³This is true for just the CRF++ toolkit, and not a drawback of CRFs in general.

C. Hybrid Generative and Discriminative Modeling

One important observation is that nonsequential discriminative classification algorithms typically ignore the context, which is critical for the segmentation task. While one may add context as an additional feature, or simply use CRFs, which inherently consider context, these approaches are suboptimal when dealing with real valued features, such as pause duration or pitch range. Most of the previous studies simply tackled this problem by binning the feature space either manually or automatically [6].

An alternative would be using a hybrid classification approach as suggested by [14]. The main idea would use the posterior probabilities P_c obtained from the other classifiers, such as boosting or CRF, by simply converting them to state observation likelihoods by dividing to their priors following the well-known Bayes rule:

$$\arg \max_{Y_t} \frac{P_c(Y_t|\mathcal{X}_t)}{P(Y_t)} = \arg \max_{Y_t} P_c(\mathcal{X}_t|Y_t). \quad (9)$$

Applying Viterbi algorithm to the HMM will then returns the most likely segmentation. In order to handle dynamic ranges of state transition probabilities and observation likelihoods, we apply a weighting scheme as is usually done in the literature

$$\arg \max_{Y_t} P_c(\mathcal{X}_t|Y_t)^\alpha \times P(Y_t)^\beta \quad (10)$$

where $P(Y_t)$ is estimated by the fHELM, α and β are optimized using a held-out set.

IV. FEATURES

In the classification models, three types of features—lexical, prosodic, and morphological—are used.

A. Lexical Features

The lexical features used in this work consist of six word n -gram features for each word boundary that were also used in our previous work for English [37]: three unigrams, two bigrams, and a trigram. Naming the word preceding the word boundary of interest as the *current* word, and the preceding and following words as the *previous* and *next* word respectively, the six lexical features are as follows:

- unigrams: {previous}, {current}, {next};
- bigrams: {current, next}, {previous, current};
- trigram: {previous, current, next}.

B. Prosodic Features

The prosodic features are also transferred from the ICSI+ sentence segmentation system [12]. We use about 200 prosodic features, defined for and extracted from the regions around each inter-word boundary. The features include the pause duration at the boundary, normalized phone durations of the word preceding the boundary, and a variety of speaker-normalized pitch features and energy features preceding, following, and across the boundary. These features are an extension of similar features described in [14]. The extraction region around the boundary focuses on either the single words or brief time windows around the boundary. Measures include the maximum, the minimum, or the average value in this range. Pitch features are normalized

by speaker, using the method to estimate a speaker’s baseline pitch values described in [14].

C. Morphological Features

Turkish is also a free-constituent-order language, in which constituents at certain phrase levels can change order rather freely according to the discourse context or text flow. However, the typical order of the constituents, especially for the news genre, is subject-object-verb (SOV).

Let us consider a simple complete sentence “*çocuk yemek yedi*” in Turkish, which means “*the child ate the meal*” in English. The correct morphological analyses are as follows:

çocuk: Noun+A3sg+Pnon+Nom (the child);
yemek: Noun+A3sg+Pnon+Nom (the meal);
yedi: Verb+Pos (+dH) +Past+A3sg (ate).

Turkish has an agglutinative morphology with productive inflectional and derivational suffixations [38]. The number of word forms one can derive from a Turkish root form may be in the millions [39]. For example, [40] shows that one can obtain thousands of new word forms from any noun, a verb, and an adjective root form by suffixing only three morphemes. Morphological information in Turkish can be represented in general form as given as follows:

$$\text{root} + IG_1 + \wedge DB + IG_2 + \wedge DB + \dots + \wedge DB + IG_n.$$

In this representation (adapted from [41]), the inflectional groups (IGs) denote the derivational boundaries and are marked with “ $\wedge DB$.” The root and derivational elements of a word are represented by different IGs. Each IG_i denotes the relevant sequence of inflectional features. Some of these inflectional features can be listed as follows:

+Adj: adjective, +Noun: noun, +Verb: verb, +A3sg: 3rd person singular agreement, +P1sg: 1st person singular possessive agreement, +Pnon: no possessive agreement, +Nom: nominative case, +Pos: positive polarity, +Past: past tense, +Fut: future tense, +FutPart: future participle.

As an example, let us consider the Turkish word “*yapabileceğim*,” which consists of the morphemes “(yap) + (abil) + (ecek) + (im)” which roughly corresponds to “(do) + (able to) + (will) + (I)” in English. It has three potential morphological analyses:

- (yap) yap+Verb+Pos (+yAbil) $\wedge DB$ +Verb+Able (+yAcAk) +Fut (+yHm) +A1sg (I’ll be able to do it);
- (yap) yap+Verb+Pos (+yAbil) $\wedge DB$ +Verb+Able (+yAcAk) $\wedge DB$ +Adj+Fut-Part (+Hm) +P1sg (The (thing that) I’ll be able to do);
- (yap) yap+Verb+Pos (+yAbil) $\wedge DB$ +Verb+Able (+yAcAk) $\wedge DB$ +Noun+FutPart+A3sg (+Hm) +P1sg+Nom (The one I’ll be able to do).

In this example, the root is a verb but the final IGs have three readings, that are verb, adjective, and noun, respectively.

Turkish presents an interesting problem for statistical models since the potential POS tag set size (that is, the number of possible morphological parses) is very large because of the productive derivational morphology. Following previous work [2], [42], our approach handles this by breaking up the morphosyntactic tags into inflectional groups, each of which contains the inflectional features for each (intermediate) derived form. To simplify our models further, we only extract morphological features from the final inflectional group of every word, which marks its final category in a sentence.

The morphological features used in this work are obtained using a morphological analyzer for Turkish [38], which outputs all possible morphological parses for all the words. We include the final inflectional group of every word as well as its POS tag, without resolving the ambiguity. For factored HELM, we arbitrarily chose one parse since fHELMs cannot handle multiple parses.⁴ With CRF and boosting we used all the possible parses as features. Boosting also exploited parse subsequences as additional features. For the POS tag, we mark the value of the feature as unknown when the word has multiple parses. We also include a single binary feature that checks if any of the possible morphological parses of a word is a Verb according to its final category. We hope, with this, to take advantage of the SOV nature of Turkish. To compare this approach, we also performed experiments with pseudo-morphological features, using the last three letters of each word. Like the “ed” suffix in English, in Turkish certain suffixes may indicate Verb categories.

The Verb information is linguistically the most important feature from the morphological analysis of Turkish. Even though the Turkish is a free word order language, the most frequent order is SOV. Especially in newswire and broadcast news, the order is almost always SOV. Therefore, if one of the morphological analyses is a verb form, this is a strong signal for a sentence end. However, due to morphological ambiguity, words that do not play the role of a verb can be assigned a verb analysis, as they may be verbs in other contexts. Prosodic features are expected to be useful in these cases.

V. EXPERIMENTS AND RESULTS

A. Data Sets

In our experiments, we use the Voice of America (VOA) Turkish Section⁵ part of the Turkish broadcast news (BN) speech corpus collected at the Boğaziçi University BUSIM Laboratory.⁶ The VOA part of the corpus contains approximately 21 hours of single-channel Turkish broadcast news speech data recorded at a 16 bit, 32-KHz sampling rate. For sentence segmentation experiments, 42 Turkish broadcast news programs (30 minutes each) are used. These 42 files are split into a training set (22 files, 97 330 words), a development set (five files, 14 897 words), and a test set (five files, 15 688 words). The development set is used to optimize the parameters, such as probability thresholds and combination weights α

⁴The performance is expected to be better when a good morphological disambiguator is used. We repeated experiments by randomly choosing a parse for each word, and the results did not change significantly.

⁵<http://www.voanews.com/turkish/>.

⁶<http://www.busim.ee.boun.edu.tr/>.

TABLE I

AMBIGUITY STATISTICS FOR DIFFERENT LEVELS OF MORPHOLOGICAL FEATURES: AVERAGE NUMBER OF PARSES PER WORD FOR EVERY WORD THAT WAS PARSED BY THE MORPHOLOGICAL ANALYZER AND PERCENTAGE OF WORDS THAT HAVE A SINGLE PARSE (I.E., UNAMBIGUOUS WORDS)

Morphological Feature	Avg. Parse/Word	% of Unamb
Full Morph. Analysis	1.95	37.0
Last IG	1.83	39.5
POS of Last IG	1.30	62.9

and β . The vocabulary size of the training set is 19 328 words, and 33.5% of the words in the development set vocabulary and 35.8% of the test set vocabulary are not observed in the training data (these correspond to 14.8% and 17.3% of the development and test set words, respectively).

There are in total 128 005 words in the training, test, and development sets. 6.76% of these are not parsed by the morphological analyzer, mainly because of foreign person and city names and typos in the data. The remaining words that are parsed have on average 1.95 parse. This drops down to on average 1.83 analyses per word if only the last inflectional group of each word is considered, and to 1.30 if only the POS tag category of the last IG is considered. Table I lists the average number of parses per word as well as the percentage of words that have a single parse in the overall data set with these different conditions.

B. Evaluation Methods

For performance evaluation, we report NIST error rate and F-measure on forced alignment output of an automatic speech recognizer [43]. The NIST error rate is the number of misclassified word boundaries divided by the number of reference sentence boundaries

$$\text{NIST} = \frac{f_n + f_p}{t_p + f_n} \quad (11)$$

where f_n , f_p , and t_p are false negative, false positive, and true positive, respectively. F-measure is the harmonic mean of precision and recall:

$$F - \text{measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

where precision = $t_p / (t_p + f_p)$ and recall = $t_p / (t_p + f_n)$. The NIST error rate is explained in detail with examples in [44].

C. Experiments With Lexical and Morphological Features

We compare our results with a baseline of using only lexical features for all classification methods. Tables II and III present results with boosting, CRF, and HELM/fHELM using lexical, morphological and/or pseudo-morphological features. In the case of only lexical features, HELM outperforms other methods probably because of the large number of lexical features they must tackle due to the agglutinative nature of Turkish.

When we add morphological and pseudo-morphological (last three letters of words) to the feature sets, we observe significant improvements in the performance with all classifiers. This is intuitive because of the morphological characteristics and SOV

TABLE II

F-MEASURE WITH BOOSTING, CRF, AND HELM/fHELM USING LEXICAL (L), MORPHOLOGICAL (M), AND/OR PSEUDO-MORPHOLOGICAL (PM) FEATURES

Classifier	F			
	L	L+M	L+PM	L+M+PM
Boosting	0.749	0.884	0.853	0.869
CRF	0.756	0.887	0.864	0.891
HELM	0.782	-	-	-
fHELM	-	0.865	0.862	-

TABLE III

NIST ERROR RATES WITH BOOSTING, CRF, AND HELM/fHELM USING LEXICAL (L), MORPHOLOGICAL (M), AND/OR PSEUDO-MORPHOLOGICAL (PM) FEATURES

Classifier	NIST			
	L	L+M	L+PM	L+M+PM
Boosting	44.0(%)	24.7(%)	30.0(%)	26.5(%)
CRF	43.3(%)	24.0(%)	26.0(%)	21.7(%)
HELM	36.7(%)	-	-	-
fHELM	-	25.9(%)	27.1(%)	-

sentence order of Turkish. One interesting observation is that with boosting the performance degrades when both morphological and pseudo-morphological features are employed instead of only one of them. CRF consistently performs a little better than boosting. The error rate of fHELM is reduced by 26% relative compared to HELM when only lexical features are used. This shows the effectiveness of factored hidden event language models for generative sequence classification. Furthermore, the relative NIST error rate reductions are even more with boosting (44%) and CRF (50%) with morphological features. These results are shown in Tables II and III.

In order to see the effect of morphological information when various amounts of training data is available, we also provide learning curves for HELM and fHELM. As Fig. 4 shows, the F-measure difference between the HELM and fHELM is larger when less training data is used, as expected. For example, the difference in F-Measure is doubled when only 10 000 examples are used instead of the whole set.

Table IV presents results with the combination of discriminative and generative sequence classification methods when both lexical and morphological features are used. The performance is more or less the same as using only the discriminative classifiers, suggesting that they probably already incorporate the information coming from hidden event language models.

D. Experiments With Prosodic Features

Since we expect the prosody to provide orthogonal information for sentence segmentation, we first check the effectiveness of using only prosodic features with boosting and CRF. The performance happens to be very similar to what we have got using the models trained with only lexical and morphological information. This shows the utility of the prosodic features that were originally designed for English.

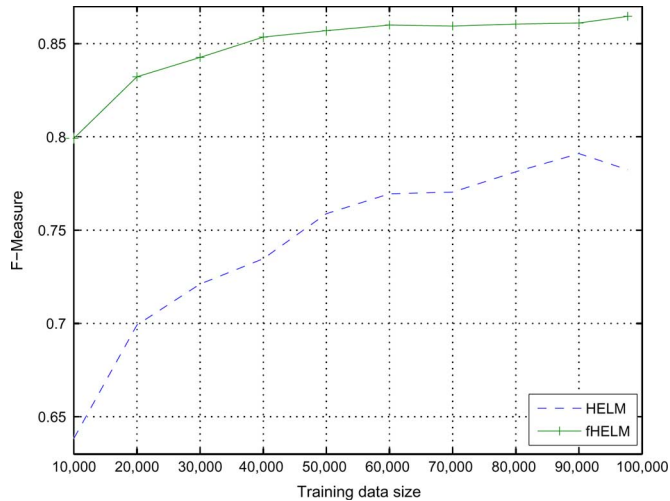


Fig. 4. F-measure learning curves for HELM and fHELM with various training data set sizes.

TABLE IV
F-MEASURE AND NIST ERROR RATES WHEN COMBINING BOOSTING AND CRF WITH fHELM WITH LEXICAL (L) AND MORPHOLOGICAL (M) FEATURES

Classifier	F	NIST
Boosting(L+M) + fHELM(L+M)	0.879	23.8%
CRF(L+M) + fHELM(L+M)	0.890	21.5%

TABLE V
F-MEASURE AND NIST ERROR RATES WHEN USING ONLY PROSODIC (P) INFORMATION WITH BOOSTING AND COMBINING WITH fHELM USING LEXICAL (L) AND MORPHOLOGICAL (M) INFORMATION

Classifier	F	NIST
CRF(P)	0.874	24.7%
Boosting(P)	0.862	27.2%
Boosting(P) + fHELM(L+M)	0.919	15.8%

In order to combine prosodic information with lexical and morphological information, we experiment with two approaches. In the first approach, we combine the classifier trained with only prosodic features with factored HELMs as presented in Section III-C. Table V presents these results. Note that, before combination, boosting and fHELMs have comparable performance (NIST error rates of 27.2% and 25.9%). The hybrid model reduces the NIST error rate by 39% relative (from 25.9% to 15.8%). This demonstrates that the information provided by two different sets are complementary. This is in part due to the nature of the data, i.e., broadcast news, in which the reporters and anchor people explicitly mark sentence boundaries with prosody.

As the second approach, we exploit the prosodic features along with lexical and morphological information with boosting and CRF. Table VI and VII present these results. As seen, for both classifiers, performance improves significantly. This approach happens to provide slightly better results than the previous one when also pseudo-morphological features are used.

TABLE VI
F-MEASURE WITH BOOSTING AND CRF USING LEXICAL (L), PROSODIC (P), MORPHOLOGICAL (M), AND/OR PSEUDO-MORPHOLOGICAL (PM) FEATURES

Classifier	F			
	L+P	L+M+P	L+PM+P	L+M+PM+P
Boosting	0.894	0.922	0.918	0.927
CRF	0.895	0.921	0.916	0.923

TABLE VII
NIST ERROR RATES WITH BOOSTING AND CRF USING LEXICAL (L), PROSODIC (P), MORPHOLOGICAL (M), AND/OR PSEUDO-MORPHOLOGICAL (PM) FEATURES

Classifier	NIST			
	L+P	L+M+P	L+PM+P	L+M+PM+P
Boosting	20.4(%)	16.5(%)	15.8(%)	14.7(%)
CRF	20.2(%)	14.6(%)	16.9(%)	15.3(%)

TABLE VIII
F-MEASURE AND NIST ERROR RATES WHEN COMBINING fHELM WITH BOOSTING AND CRF USING LEXICAL (L), MORPHOLOGICAL (M+PM), AND PROSODIC (P) INFORMATION

Classifier	F	NIST
Boosting(L+P+M+PM) + fHELM(L+M)	0.925	14.8%
CRF(L+P+M+PM) + fHELM(L+M)	0.926	14.9%

As the final set of experiments, we have tried combining fHELM with boosting and CRF using all the features. Table VIII presents these results. With this final combination, the performance of the hybrid model including boosting does not improve. The performance of the one with CRF improves, however only slightly.

VI. DISCUSSIONS

Discriminative classification approaches provide the best results for Turkish sentence segmentation using lexical, morphological, and prosodic features. While CRF results in better performance with prosodic and lexical features only, boosting benefits more from the morphological features. This is probably due to the ability of boosting to handle unknown feature values. For example, one of the morphological features is set to unknown in case the word is morphologically ambiguous. This requires further investigation, but a prior morphological disambiguation step may provide benefits.

Even though in our experiments, the discriminative models alone result in the best performance, the generative models have potential uses for sentence segmentation. The boosting and CRF models have access to several prosodic features, which are difficult to include in HELMs. However, usually, while there is only little speech data available, there is significantly more data from the written text sources, such as newspapers. In order to benefit from both data sources, our practice for training models for sentence segmentation in English has mainly been training boosting and CRF models from the speech data and the HELMs from the textual data (usually on the order of hundreds of millions), and

combine the two in HELMs during the test time, resulting in the best performance for English, as shown in [45], where for example boosting performance with all features is improved from an F-measure of 68.9% to 70.6% when combining boosting with a HELM trained from textual data of millions of words.

In this paper, we have used the same data for training all models, and investigating the use of more data for fHELMs is part of our future work, in addition to experimenting with real ASR output. In order to study this effect, we have done a simple experiment combining the Boosting model trained with all features, using 1% of the data, with fHELM. In this case, NIST error rate of the system decreases from 29.6% (Boosting only, with all features) to 28.5% (Boosting with all features combined with fHELM), encouraging the use of generative methods in combination with discriminative approaches. When 10% of the data is used, the NIST error rate decreases from 18.8% to 18.4%.

VII. CONCLUSION

We have presented generative, discriminative, and hybrid classification methods using lexical, morphological, and prosodic information for Turkish sentence segmentation. We have shown significant improvements over a lexical baseline.

The prosodic features are mainly transferred from English and model only word-level phenomena. They can also be improved by modeling at subword level. For example, the morphological ambiguity for the sentence final words may be resolved using morpheme-level prosodic features.

Morphological ambiguity is a problem for factored hidden event modeling. Our future work also includes checking the effect of morphological disambiguation for this task.

Note that fHELMs can be used for similar language processing tasks requiring sequence classification such as comma prediction, POS tagging, and named entity extraction and can easily be combined with state-of-the-art discriminative models.

ACKNOWLEDGMENT

The authors would like to thank K. Oflazer for providing us with his morphological analyzer, S. Parlak and M. Saraclar for providing us with the data sets and the forced alignments, and A. Stolcke, D. Vergyri, E. Shriberg, and J. Fung for many helpful discussions. The authors would also like to thank the anonymous reviewers for their useful comments on an earlier version of this paper.

REFERENCES

- [1] A. Waibel, P. Geutner, L. Mayfield-Tomokiyo, T. Schultz, and M. Woszczyna, "Multilinguality in speech and spoken language systems," *Proc. IEEE, Special Iss. Spoken Lang. Process.*, vol. 88, no. 8, pp. 1297–1313, Aug. 2000.
- [2] D. Hakkani-Tür, K. Oflazer, and G. Tur, "Statistical morphological disambiguation for agglutinative languages," in *Proc. 18th Int. Conf. Comput. Lingist. (COLING)*, Saarbruecken, Germany, Aug. 2000.
- [3] A. McCallum, D. Freitag, and F. Pereira, "Maximum entropy markov models for information extraction and segmentation," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Palo Alto, CA, Jun. 2000.
- [4] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. 18th Int. Conf. Mach. Learn. (ICML)*, San Francisco, CA, 2001, pp. 282–289, Morgan Kaufmann.
- [5] Y. Altun, "Discriminative methods for label sequence learning," Ph.D. dissertation, Dept. of Comput. Sci., Brown Univ., Providence, RI, 2005.
- [6] H.-K. J. Kuo and Y. Gao, "Maximum entropy direct models for speech recognition," *IEEE Trans. Speech Audio Process.*, vol. 14, no. 3, pp. 873–881, May 2006.
- [7] C. D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press, 1999.
- [8] J. A. Bilmes and K. Kirchhoff, "Factored language models and generalized parallel backoff," in *Proc. Human Lang. Technol. Conf. (HLT)-Conf. North Amer. Chap. Assoc. Comput. Lingist. (NAACL)*, Edmonton, AB, Canada, May 2003.
- [9] D. Vergyri, K. Kirchhoff, K. Duh, and A. Stolcke, "Morphology-based language modeling for Arabic speech recognition," in *Proc. Int. Conf. Spoken Lang. Process. (ICSLP)*, Jeju-Island, Korea, Oct. 2004.
- [10] A. Srivastava and F. Kubala, "Sentence boundary detection in Arabic speech," in *Proc. 8th Eur. Conf. Speech Commun. Technol.*, 2003, ISCA.
- [11] J. Huang and G. Zweig, "Maximum entropy model for punctuation annotation from speech," in *Proc. 7th Int. Conf. Spoken Lang. Process. (ICSLP)*, Denver, CO, 2002.
- [12] M. Zimmermann, D. Hakkani-Tür, J. Fung, N. Mirghafori, L. Gottlieb, Y. Liu, and E. Shriberg, "The ICSI+ multi-lingual sentence segmentation system," in *Proc. Int. Conf. Spoken Lang. Process. (Interspeech)*, Pittsburgh, PA, Sep. 2006.
- [13] B. Roark, Y. Liu, M. Harper, R. Stewart, M. Lease, M. Snover, I. Shafran, B. Dorr, J. Hale, A. Krasnyanskaya, and L. Yung, "Reranking for sentence boundary detection in conversational speech," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Toulouse, France, 2006.
- [14] E. Shriberg, A. Stolcke, D. Hakkani-Tür, and G. Tur, "Prosody-based automatic segmentation of speech into sentences and topics," *Speech Commun.*, vol. 32, no. 1–2, pp. 127–154, 2000.
- [15] Y. Liu, E. Shriberg, A. Stolcke, B. Peskin, J. Ang, D. Hillard, M. Ostendorf, M. Tomalin, P. Woodland, and M. Harper, "Structural metadata research in the EARS program," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2005.
- [16] J. Kolar, J. Svec, and J. Psutka, "Automatic punctuation annotation in Czech broadcast news speech," in *Proc. 9th Conf. Speech Comput.*, 2004.
- [17] C. Zong and F. Ren, "Chinese utterance segmentation in spoken language translation," in *Proc. 4th Int. Conf. Comput. Lingist. Intell. Text Process.*, 2003, pp. 516–525.
- [18] D. Liu and C. Zong, "Utterance segmentation using combined approach based on bi-directional n-gram and maximum entropy," in *Proc. ACL Workshop: 2nd SIGHAN Workshop Chinese Lang. Process.*, 2003, pp. 16–23.
- [19] J. Fung, D. Hakkani-Tür, M. M. Doss, E. Shriberg, S. Cuendet, and N. Mirghafori, "Prosodic features and feature selection for multi-lingual sentence segmentation," in *Proc. Int. Conf. Spoken Lang. Process. (Interspeech-Eurospeech)*, Antwerp, Belgium, 2007.
- [20] T. Kawahara, M. Saikou, and K. Takanashi, "Automatic detection of sentence and clause units using local syntactic dependency," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Toulouse, France, Apr. 2007.
- [21] G. Tur, "A statistical information extraction system for Turkish," Ph.D. dissertation, Dept. of Comput. Sci., Bilkent Univ., Ankara, Turkey, 2000.
- [22] J. Kolár, E. Shriberg, and Y. Liu, "On speaker-specific prosodic models for automatic dialog act segmentation of multi-party meetings," in *Proc. 9th Int. Conf. Spoken Lang. Process.*, 2006, ISCA.
- [23] S. Cuendet, D. Hakkani-Tur, and G. Tur, "Model adaptation for sentence unit segmentation from speech," in *Proc. SLT*, Aruba, 2006.
- [24] E. Matusov, A. Mauser, and H. Ney, "Automatic sentence segmentation and punctuation prediction for spoken language translation," in *Proc. Int. Workshop Spoken Lang. Translation (IWSLT)*, 2006, pp. 158–165.
- [25] S. Rao, I. Lane, and T. Schultz, "Optimizing sentence segmentation for spoken language translation," in *Proc. Int. Conf. Spoken Lang. Process. (Interspeech-Eurospeech)*, Antwerp, Belgium, 2007.
- [26] B. Favre, R. Grishman, D. Hillard, H. Ji, D. Hakkani-Tur, and M. Ostendorf, "Punctuating speech for information extraction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, NV, 2008.
- [27] A. Stolcke and E. Shriberg, "Statistical language modeling for speech disfluencies," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Atlanta, GA, May 1996.

- [28] K. W. Church, "A stochastic parts program and noun phrase parser for unrestricted text," in *Proc. Conf. Appl. Natural Lang. Process. (ANLP)*, Austin, TX, 1988, pp. 136–143.
- [29] D. M. Bikel, R. Schwartz, and R. M. Weischedel, "An algorithm that learns what's in a name," *Mach. Learn. J. Special Iss. Natural Lang. Learn.*, vol. 34, no. 1–3, pp. 211–231, 1999.
- [30] K. Duh and K. Kirchhoff, "Automatic learning of language model structure," in *Proc. Int. Conf. Comput. Lingvist. (COLING)*, Geneva, Switzerland, 2004.
- [31] A. Stolcke, "SRILM-An extensible language modeling toolkit," in *Proc. Int. Conf. Spoken Lang. Process. (ICSLP)*, Denver, CO, Sep. 2002.
- [32] R. Kneser and H. Ney, "Improved clustering techniques for class-based statistical language modeling," in *Proc. Eur. Conf. Speech Commun. Technol. (Eurospeech)*, Berlin, Germany, 1993.
- [33] Y. Liu, A. Stolcke, E. Shriberg, and M. Harper, "Using conditional random fields for sentence boundary detection in speech," in *Proc. Annu. Meeting Assoc. Comput. Lingvist. (ACL)*, Ann Arbor, MI, 2005.
- [34] S. Cuendet, D. Hakkani-Tür, and G. Tur, "Model adaptation for sentence segmentation from speech," in *Proc. IEEE/ACL Spoken Lang. Technol. (SLT) Workshop*, Aruba, 2006.
- [35] R. E. Schapire and Y. Singer, "Boostexter: A boosting-based system for text categorization," *Mach. Learn.*, vol. 39, no. 2/3, pp. 135–168, 2000.
- [36] S. Friedman, D. Fox, and H. Pasula, "Voronoi random fields: Extracting the topological structure of indoor environments via place labeling," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, Hyderabad, India, Jan. 2007.
- [37] S. Cuendet, D. Hakkani-Tür, E. Shriberg, J. Fung, and B. Favre, "Cross-genre feature comparisons for spoken sentence segmentation," in *Proc. IEEE Int. Conf. Semantic Comput. (ICSC)*, Irvine, CA, 2007.
- [38] K. Oflazer, "Two-level description of Turkish morphology," *Literary Linguist. Comput.*, vol. 9, no. 2, 1994.
- [39] J. Hankamer, "Morphological parsing and the lexicon," in *Lexical Representation and Process*, W. Marslen-Wilson, Ed. Cambridge, MA: MIT Press, 1989.
- [40] D. Hakkani-Tür, K. Oflazer, and G. Tur, "Statistical morphological disambiguation for agglutinative language," *Comput. Humanities*, vol. 36, no. 4, 2002.
- [41] K. Oflazer, "Dependency parsing with an extended finite state approach," in *Proc. 37th Annu. Meeting Assoc. Comput. Lingvist.*, 1999.
- [42] K. Oflazer, "Dependency parsing with an extended finite state approach," in *Proc. 37th Annu. Meeting Assoc. Comput. Lingvist. (ACL)*, 1999.
- [43] E. Arisoy, D. Can, S. Parlak, H. Sak, and M. Saraclar, "Turkish broadcast news transcription and retrieval," *IEEE Trans. Audio, Speech, Lang. Process., Special Iss. Process. Morphol. Rich Lang.*, vol. 17, no. 4, pp. XXX–XXX, May 2009.
- [44] Y. Liu, E. Shriberg, A. Stolcke, D. Hillard, M. Ostendorf, and M. Harper, "Enriching speech recognition with automatic detection of sentence boundaries and disfluencies," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 5, pp. 1526–1540, Sep. 2006.
- [45] M. M. Doss, D. Hakkani-Tür, O. Cetin, E. Shriberg, J. Fung, and N. Mirghafori, "Entropy based classifier combination for sentence segmentation," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Honolulu, HI, 2007, pp. 189–192.



Umit Guz (M'03) graduated from the Department of Computer Programming, Yildiz Technical University, Istanbul, Turkey, in 1990, the B.S. degree with high honors from the Department of Electronics Engineering, College of Engineering, Istanbul University, in 1994, and the M.S. and Ph.D. degrees in electronics engineering with high honors from the Institute of Science, Istanbul University, in 1997 and 2002, respectively.

He is an Assistant Professor in the Department of Electronics Engineering, Engineering Faculty, Isik University, Istanbul. From 1995 to 1998, he was a Research and Teaching Assistant in the Department of Electronics Engineering, Istanbul University. His research interest covers speech processing, speech modeling, speech coding, speech compression, automatic speech recognition, natural language processing, and biosignal processing.

Dr. Guz was awarded a Postdoctoral Research Fellowship by The Scientific and Technological Research Council of Turkey (TUBITAK) in 2006. He was accepted as an International Fellow by the SRI International Speech Technology and Research (STAR) Laboratory in 2006. He was awarded a J. William Fulbright Postdoctoral Research Fellowship for 2007. He was accepted as an Inter-

national Fellow by the International Computer Science Institute (ICSI) Speech Group at the University of California at Berkeley in 2007 and 2008. He received a TUBITAK Career Award for his project entitled Extracting and Using Prosodic Information for Turkish Spoken Language Processing in 2008–2010.



Benoit Favre (M'08) received the B.S., M.S., and Ph.D. degrees from the University of Avignon, Avignon, France, in 2001, 2003, and 2007. His Ph.D. thesis explored interactive speech summarization of broadcast news archives.

He was a Teaching Assistant at the University of Avignon from 2003 to 2007. He was also a Research Engineer at Thales Land, and Joint Systems, Paris, France, from 2004 to 2007. He currently holds a post-doctoral position at the International Computer Institute (ICSI), Berkeley, CA. His research interests include natural language processing, speech understanding, and text and speech summarization. He is also interested in machine learning on structured outputs and global inference.

Dr. Favre is a member of ISCA and was a member and Web Master of AFCEP.



Dilek Hakkani-Tür (S'00–M'01–SM'05) received the B.Sc. degree from Middle East Technical University, Ankara, Turkey, in 1994, and M.Sc. and Ph.D. degrees from Department of Computer Engineering, Bilkent University, Ankara, in 1996 and 2000, respectively. Her Ph.D. dissertation is on statistical language modeling for agglutinative languages.

She is a Senior Researcher at ICSI. Prior to joining ICSI, she was a Senior Technical Staff Member in the Voice Enabled Services Research Department at AT&T Labs-Research, Florham Park, NJ. She

worked on machine translation during her visit to the Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, in 1997, and her visit to the Computer Science Department, The Johns Hopkins University, in 1998. In 1998 and 1999, she visited SRI International, Speech Technology and Research Laboratory, and worked on using lexical and prosodic information for information extraction from speech. In 2000, she worked in the Natural Sciences and Engineering Faculty of Sabanci University, Turkey. Her research interests include natural language and speech processing, spoken dialog systems, and active and unsupervised learning for language processing. She has coauthored several papers in natural language and speech processing.

Dr. Hakkani-Tür is a member of ISCA, IEEE, and ACL. She was an associate editor of IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING from 2005 to 2008 and is a member of the IEEE Signal Processing Society (SPS), Speech and Language Technical Committee (SLTC) for 2009–2011.



Gokhan Tur (M'01–SM'05) received the B.S., M.S., and Ph.D. degrees from the Department of Computer Science, Bilkent University, Ankara, Turkey, in 1994, 1996, and 2000, respectively.

From 1997 to 1999, he visited the Center for Machine Translation, Carnegie Mellon University, Pittsburgh, PA, then the Department of Computer Science, The Johns Hopkins University, Baltimore, MD, and then the Speech Technology and Research Laboratory, SRI International, Menlo Park, CA. He worked at AT&T Labs—Research from 2001

to 2006. He is currently with the Speech Technology and Research Lab, SRI International. His research interests include spoken language understanding (SLU), speech and language processing, machine learning, and information retrieval and extraction. He coauthored more than 70 papers published in refereed journals and presented at international conferences.

Dr. Tur is the recipient of the Speech Communication Journal Best Paper awards by ISCA for 2004–2006 and by EURASIP for 2005–2006. Dr. Tur is the organizer of the HLT-NAACL 2007 Workshop on Spoken Dialog Technologies, and the HLT-NAACL 2004 and AAAI 2005 Workshops on SLU, and the editor of the *Speech Communication Journal* Special Issue on SLU in 2006. He is also the spoken language processing area chair for IEEE ICASSP 2007 and IEEE ICASSP 2008 conferences, spoken dialog area chair for HLT-NAACL 2007 conference, finance chair for IEEE/ACL SLT 2006 workshop, and SLU area chair for IEEE ASRU 2005 workshop. He is a senior member of ACL and ISCA, and a member of the IEEE Signal Processing Society (SPS), Speech and Language Technical Committee (SLTC) for 2006–2008.

TÜBİTAK
PROJE ÖZET BİLGİ FORMU

Proje No: 107E182
Proje Başlığı: Türkçe Dil İşleme için Bürünsel Bilginin Çıkarılması ve Kullanılması (Extracting and Using Prosodic Information for Turkish Spoken Language Processing)
Proje Yürütücüsü ve Araştırmacılar: Yrd. Doç. Dr. Ümit Güz (Proje yürütücüsü) Yrd. Doç. Dr. Hakan Gürkan Sinan Yiğit
Projenin Yürütüldüğü Kuruluş ve Adresi: Işık Üniversitesi, Mühendislik Fakültesi, Elektronik Mühendisliği Bölümü, Mesrutiyet Koyu, Üniversite Sokak, No:2, 34980, Sile, İstanbul.
Destekleyen Kuruluş(ların) Adı ve Adresi:
Projenin Başlangıç ve Bitiş Tarihleri: 01.02.2008 - 01.02.2010
Öz (en çok 70 kelime) <p>Bu projede genel olarak, konuşulan dili (Türkçe) anlamada, konuşulan dilin bürünsel/ezgisel (prosodic) ve sözcüksel (lexical) özelliklerinin ortaya çıkarılması ve bu özelliklerin konuşulan dilin bilgisayarla otomatik olarak işlenmesinde kullanılması amaçlanmaktadır. Bu daha özel olarak, otomatik konuşma tanıyıcısının (ASR) çıkışına ilişkin cümle bölütleme işlevini içermektedir. Otomatik konuşma tanıma sistemlerinden çıkan yazılı metnin özellikle noktalama (punctuation), büyük küçük harf farklılıkları ve vurgu, tonlama, perde, durak gibi konuşmaya ilişkin temel bazı parametrelerden yoksun olması veya bu özellikleri kaybetmiş olması, özellikle anlamda farklılıklara yol açmaktadır. Bu çıktının zenginleştirilmesi (enrichment) başka bir deyiş ile bu özelliklerin tekrar geriye kazandırılması, bu metinlerin hem insanlar tarafından okunmasını ve doğru algılanmasını hem de makineler tarafından işlenmesini kolaylaştıracaktır. Bu projedeki amaç, bu zenginleştirme ve geri kazandırma işleminin dilin bürünsel özelliklerinden yararlanarak yapılmasıdır.</p>
Anahtar Kelimeler: Bürünsel Bilgi, Dil İşleme, Cümle Bölütleme, Konu Bölütleme.
Fikri Ürün Bildirim Formu Sunuldu mu? Evet <input type="checkbox"/> Gerekli Değil <input checked="" type="checkbox"/> <small>Fikri Ürün Bildirim Formu'nun tesliminden sonra 3 ay içerisinde patent başvurusu yapılmalıdır.</small>
Projeden Yapılan Yayınlar: Umit Guz, Benoit Favre, Dilek Hakkani-Tür, Gokhan Tur, "Generative and Discriminative Methods using Morphological Information for Sentence Segmentation of Turkish", Special Issue on Processing Morphologically Rich Languages, IEEE Transactions on Audio, Speech and Language Processing, Volume 17, Issue 5, pp. 895-903, July 2009.