

ÖZET

Bu tez, sözdizim bilgisi kullanarak Türkçe'deki üçüncü kişi adlarını çözümlenmeye çalışan bir sistem sunar. Sistem, verilen bir metinde tümcelerin sadece sözdizim yapılarını kullanan Hobbs'un Naif Yaklaşımı (Hobbs, 1976, 1978) ile sözdizimsel kuram ve dil işleme modelleri içinde temel bir rol oynayan Chomsky'nin Bağlama Kuramına (Chomsky, 1981, 1995) dayanmaktadır. Algoritma, çeşitli metinlerden çıkartılmış, tümceler içeren bazı söylem parçacıklarına uygulanmıştır. Türkçe'deki üçüncü kişi adlarını çözümlenmek için oluşturulan bu model, derin söylem analizi yapmadan metinden çıkartılabilen basit sözdizimsel bilgi ve dünya bilgisini kullanmaktadır.

Algoritma Hobbs'un Naif Algoritmasında olduğu gibi belirli bir dilbilgisi kuralları tarafından üretilmiş, ayrıştırma ağaçları biçimindeki tümcelere uygulanmaktadır. Algoritma, adların olası aday öncüllerini kendi yerel alanları dışında, ilgili ayrıştırma ağaçları üzerinden giderek araştırmak için Chomsky'nin Bağlama Kuramı'ndan çıkartılmış sözdizimsel kısıtlamalarını kullanmaktadır. Tüm olası aday öncüller arasından en iyi öncülün belirlenmesinde önemli bir rol oynayan etkili bazı faktörler, bir kısıtlar ve tercihler listesi şeklinde tanımlanmıştır.

Bu çalışmada önerilen sistem, iki farklı dayanak algoritması ve Türkçe için kişi adlı çözümlenmesi yapan az bilgili bir sistem (Küçük ve Turhan-Yöndem, 2007) ile karşılaştırılmıştır. Asıl sistem, bu dayanak algoritmaları ve az-bilgili sisteme karşı, farklı kaynaklardan iki örnek metin üzerinde denenmiştir. Bu tezin temel amacı Bilgisayarlı Türkçe Adıl Çözümlenmesi yapmak ve elde edilen sonuçları Türkçe için bu alanda yapılmış bir başka önemli çalışma olan az-bilgili sistem ile kıyaslayarak gelecek için bir tartışma başlatmaktır.

Anahtar Kelimeler: Artgönderim, Artgönderim Çözümlenmesi, Adıl, Türkçe için Adıl Çözümlenmesi

ABSTRACT

This thesis presents a system intended to resolve third personal pronouns in Turkish using syntactic information. The system is based on Hobbs' Naïve Algorithm (Hobbs, 1978), which exploits only the surface syntax of sentences in a given text and Chomsky's Binding Theory (Chomsky, 1981, 1995), which has played a central role in both syntactic theory and models of language processing. The algorithm is applied to some discourse fragments which consist of sentences extracted from various texts. The model for third personal pronoun resolution in Turkish utilizes simple syntactic information, which can be extracted from the text without deep discourse analysis and world knowledge.

The algorithm is applied to the sentences which are presented as parse trees produced by a particular grammar like that used in Hobbs' Naïve Algorithm. The algorithm focuses on syntactic restrictions which are derived from Chomsky's Binding Theory to search for the possible candidate antecedents of the pronouns outside their local domain by traversing the relevant parse trees. Some effective factors, which come into play in order to determine the best antecedent among all possible candidates, are defined as a list of constraints and preferences.

The system proposed in this study is compared with two different baseline algorithms and a knowledge-poor pronoun resolution system for Turkish (Kucuk and Turhan-Yöndem, 2007). The original system is tested against these baseline algorithms and the knowledge-poor system on two sample Turkish texts from different sources. Thus, the main goal of the study is to provide a computational implemented system for resolving third personal pronouns in Turkish and to compare the results produced by our syntax-based system with the knowledge-poor system, which is an another important work about the same topic for Turkish, and to start a discussion about it.

Keywords: Anaphora, Anaphora Resolution, Pronoun, Pronoun Resolution for Turkish.

TEŐEKKÜR

Doktora tezi danıřmanlıđımı üstlenerek gerek konu seęimi, gerekse ęalıřmalarımın yürütülmesi sırasında yardımlarını esirgemeyen danıřman hocam Sayın Doę. Dr. Yılmaz KILIÇASLAN'a ve ęok deđerli katkılarından dolayı eř danıřman hocam Sayın Yrd. Doę. Dr. Rafet AKDENİZ'e teőekkür ederim.

Ayrıca, bana her zaman destek olan ve sabır gösteren eřime ęok teőekkür ederim. Onlardan ęaldıđım zaman dilimi için sevgili ikiz bebeklerime de ayrıca teőekkür ederim.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
ŞEKİL LİSTESİ.....	vii
TABLO LİSTESİ.....	ix
BÖLÜM 1	1
GİRİŞ	1
BÖLÜM 2	4
TEMEL KAVRAMLAR VE TERMİNOLOJİ.....	4
2.1 Metin, Söylem, Tutarlılık ve Bağdaşıklık.....	4
2.2 Artgönderim	6
2.3 Artgönderim Çeşitleri.....	8
2.4 Dilbilimsel Artgönderim Çalışmaları.....	14
2.4.1 Bağlama Kuramı	14
2.4.2 Diğer Dilbilimsel Artgönderim Çalışmaları	19
2.4.2.1 Konu Sürekliliği Modeli	19
2.4.2.2 Aşamalı Model.....	21
2.4.2.3 Ulaşılabilirlik Kuramı	23
2.4.2.4 Merkezleme Kuramı	24
2.4.3 Türkçe için Yapılan Dilbilimsel Artgönderim Çalışmaları.....	29
2.4.3.1 Enç (1986).....	29
2.4.3.2 Erguvanlı-Taylan (1986).....	31
2.4.3.3 Özsoy (1992).....	35
2.4.3.4 Turan (1996)	36
2.4.3.5 Oktar ve Yağcıoğlu (1997)	36
2.4.3.6 Çeltek ve Oktar (2004).....	37

2.4.3.7 Kılıçaslan (2004).....	37
BÖLÜM 3	38
BİLGİSAYARLI ARTGÖNDERİM ÇALIŞMALARI	38
3.1 Bilgisayarlı Artgönderim Çözümlemesi	38
3.1.1 Artgönderimin Tespit Edilmesi.....	39
3.1.2 Olası Aday Öncüllerin Belirlenmesi	39
3.1.3 Olası Aday Öncüller Arasından En İyi Öncülün Seçilmesi.....	40
3.1.3.1 Kısıtlamalar	41
3.1.3.2. Öncelikler/Tercihler	45
3.2 Bilgisayarlı Artgönderim Çalışmaları	46
3.2.1 Sözdizimsel Yaklaşımlar.....	47
3.2.1.1 Hobbs'un Naif Yaklaşımı	48
3.2.1.2 Lappin ve Leass'in Sözdizim Tabanlı Algoritması.....	53
3.2.1.3 Kennedy ve Borguraev'in Yöntemi	58
3.2.2 Sözdizimsel Olmayan Yaklaşımlar	58
3.2.2.1 Bilgi Tabanlı Yaklaşımlar	58
3.2.2.1.1 Baldwin'in CogNIAC Programı	58
3.2.2.1.2 Mitkov'un Az-bilgili Yaklaşımı.....	59
3.2.2.2 Öğrenme Tabanlı Yaklaşımlar	61
3.2.2.2.1 McCarthy ve Lefnert (1995)	61
3.2.2.2.2 Aone ve Bennet (1996)	61
3.2.2.2.3 Soon, Ng ve Lim (2001).....	62
3.2.2.3 Melez Yaklaşımlar	63
3.2.2.3.1 Mitkov, Evans ve Orasan (2002)	63
3.2.2.3.2 Preiss (2002)	63
3.2.3 Türkçe için Bilgisayarlı Artgönderim Çözümleme Çalışmaları	63
3.2.3.1 Tın ve Akman (1994).....	64
3.2.3.2 Turhan-Yöndem ve Şehitoğlu (1997)	65
3.2.3.3 Yüksel ve Bozşahin (2002)	65
3.2.3.4 Küçük ve Turhan-Yöndem (2007)	66
3.2.3.5 Yıldırım ve Kılıçaslan (2007)	67

BÖLÜM 4	69
SİSTEM TASARIMI VE UYGULAMASI.....	69
4.1 Sistem Tanıtımı	70
4.1.1 Bütünce	76
4.1.2 Algoritma	78
4.1.3 Kısıtlar.....	86
4.1.4 Tercihler	96
BÖLÜM 5	111
DEĞERLENDİRME.....	111
5.1 Performans Değerlendirme Ölçütleri	111
5.2 Performans Sonuçları	113
BÖLÜM 6	121
SONUÇ	121
KAYNAKÇA	123
EKLER	
Ek-1 Sistemden Elde Edilen Sonuçlar	
Ek-2 Program Kodu	

ŞEKİL LİSTESİ

- 2.1 Bağlama Kuramı İçinde İsim Öbeklerinin Sınıflandırılması
- 2.2 K-konumlandır Kavramı İçin Örnek Bir Gösterim
- 3.1 Tümce 3.10'un Sözdizimsel Yapısı
- 3.2 Tümce 3.11'in Sözdizimsel Yapısı
- 3.3 Tümce 3.24'ün Ağaç Yapısı ve Algoritmanın Çalışmasının Güvenliği
- 4.1 Otomatik Artgönderim Çözümleme
- 4.2 Adılın Tespit edilmesi
- 4.3 Program Tarafından Yapılan Adıl Tespiti
- 4.4 Olası Aday Öncülleri Tespiti
- 4.5 Program Tarafından Olası Aday Öncüllerin Gösterilmesi
- 4.6 En İyi Öncülün Seçimi
- 4.7 Program Tarafından En İyi Öncülün Gösterilmesi
- 4.8 Türkçe için Sözdizim Tabanlı Adıl Çözümleme Sistemi Akış Diyagramı
- 4.9 Seçilen Metine Uygulanan Ön İşlem
- 4.10 'S' Tümcesi için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.11 'NP' Ad Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.12 'VP' Eylem için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.13 'AdvP' Belirteç Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.14 'AdjP' Sıfat Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.15 'PP' İleç Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları
- 4.16 Tümce 4.2'nin Ağaç Yapısı ve 'onun' Adılı İçin Algoritmanın İşleyişi
- 4.17 Tümce 4.2'nin 'onun' Adılı İçin Program Çıktısı
- 4.18 Tümce 4.2'nin Ağaç Yapısı ve 'onu' Adılı İçin Algoritmanın İşleyişi
- 4.19 Tümce 4.2'nin 'onu' Adılı İçin Program Çıktısı
- 4.20 Söylem 4.3'ün Ağaç Yapısı ve 'onlar' Adılı İçin Algoritmanın İşleyişi
- 4.21 Söylem 4.3'ün 'onlar' Adılı İçin Program Çıktısı
- 4.22 Söylem 4.4'ün Ağaç Yapısı ve 'onlar' Adılı İçin Algoritmanın İşleyişi
- 4.23 Söylem 4.3'ün 'onlar' Adılı İçin Program Çıktısı
- 4.24 Filtre-1
- 4.25 Söylem 4.5'in Ağaç Yapısı ve 'o' Adılı İçin Algoritmanın İşleyişi

- 4.26 Söylem 4.5'in 'o' Adılı İçin Program Çıktısı
- 4.27 Söylem 4.5'in Ağaç Yapısı ve 'onları' Adılı İçin Algoritmanın İşleyişi
- 4.28 Söylem 4.5'in 'onları' Adılı İçin Program Çıktısı
- 4.29 Söylem 4.6'nın Ağaç Yapısı ve 'onun' Adılı İçin Algoritmanın İşleyişi
- 4.30 Filtre-2
- 4.31 Söylem 4.7'nin Ağaç Yapısı ve 'ona' Adılı İçin Algoritmanın İşleyişi
- 4.32 Söylem 4.7'nin 'ona' Adılı İçin Program Çıktısı
- 4.33 Söylem 4.8'in Ağaç Yapısı ve 'ona' Adılı İçin Algoritmanın İşleyişi
- 4.34 Söylem 4.8'in 'ona' Adılı İçin Program Çıktısı
- 4.35 Söylem 4.9'un Ağaç Yapısı ve 'ona' Adılı İçin Algoritmanın İşleyişi
- 4.36 Filtre-3
- 4.37 Söylem 4.9'un 'ona' Adılı İçin Program Çıktısı
- 4.38 Söylem 4.10'un Ağaç Yapısı ve 'onu' Adılı İçin Algoritmanın İşleyişi
- 4.39 Söylem 4.10'un 'onu' Adılı İçin Program Çıktısı
- 4.40 Söylem 4.11'in Ağaç Yapısı ve 'onlara' Adılı İçin Algoritmanın İşleyişi
- 4.41 Filtre-4
- 4.42 Söylem 4.12'nin Ağaç Yapısı ve 'onun' Adılı İçin Algoritmanın İşleyişi
- 4.43 Söylem 4.12'nin 'onun' Adılı İçin Program Çıktısı
- 4.44 Filtre-4 Çıkışı
- 4.45 Söylem 4.13'ün Ağaç Yapısı ve 'onların' Adılı İçin Algoritmanın İşleyişi
- 4.46 Söylem 4.13'ün 'onları' Adılı İçin Program Çıktısı
- 4.47 Söylem 4.14'ün Ağaç Yapısı ve 'o' Adılı İçin Algoritmanın İşleyişi
- 4.48 Söylem 4.14'ün 'o' Adılı İçin Program Çıktısı
- 4.49 Söylem 4.14'ün Ağaç Yapısı ve 'onlara' Adılı İçin Algoritmanın İşleyişi
- 4.50 Söylem 4.14'ün 'onlara' Adılı İçin Program Çıktısı
- 4.51 Filtre-5
- 4.52 Söylem 4.15'in Ağaç Yapısı ve 'o' Adılı İçin Algoritmanın İşleyişi
- 4.53 Söylem 4.15'in 'o' Adılı İçin Program Çıktısı
- 4.54 Söylem 4.16'nın Ağaç Yapısı ve 'onu' Adılı İçin Algoritmanın İşleyişi
- 4.55 Söylem 4.16'nın 'onu' Adılı İçin Program Çıktısı

TABLO LİSTESİ

- 2.1 Türkçe Adısal Artgönderimler
- 2.2 Merkezdeki Değişiklikler
- 2.3 Senaryolar
- 3.1 Belirgenlik Türleri ve Ağırlıkları
- 3.2 Söylem 3.28'in ilk tümcesindeki aday öncüllerin Belirgenlik Değerleri
- 3.3 Söylem Modeli
- 3.4 Söylem 3.28'in ikinci tümcesindeki 'o' adılıının Belirgenlik Değerleri
- 3.5 Söylem Modeli
- 3.6 Söylem 3.28'in ikinci tümcesindeki 'onu' adılıının Belirgenlik Değerleri
- 3.7 Söylem Modeli
- 3.8 Söylem 3.28'in ikinci tümcesindeki 'Jane' için Belirgenlik Değerleri
- 3.9 Söylem Model
- 3.10 Söylem Model
- 4.1 Bütünce İçeriği
- 4.2 Sistemin Bütünceleri
- 5.1 Karışıklık Matrisi
- 5.2 Sözdizim Tabanlı Sistemin Eğitim Bütüncesi İçin Başarı Oranları
- 5.3 Sistemde Kullanılan Bütünceler
- 5.4 Sözdizim Tabanlı Sistem Eğitim Bütüncesinin Yeniden Uyarlanmış Hali İçin Başarı Oranları
- 5.5 Sözdizim Tabanlı Adıl Çözümleme Sistemi ile Az-bilgili Adıl Çözümleme Sisteminin Kıyaslanması
- 5.6 Artgönderim Başarı Çalışmalarının Başarı Oranları

BÖLÜM 1

GİRİŞ

20. yüzyılda hızla gelişim gösteren dilbilim, yalnızca dilin çeşitli düzey ve düzlemlerinde yapılan çalışmalara tanıklık etmemiş, aynı zamanda kendisi dışındaki bilimsel alanların gelişimine de katkıda bulunan disiplinlerden birisi olmuştur. Dilbilimin açtığı çalışma alanlarından birisi de artgönderimsel ilişkilerin anlaşılması ve çözümlenmesidir. Artgönderimsel ilişkiler, son yıllarda, dilbilim çalışmalarının merkezinde yer almakla kalmamış, filozoflar, psikologlar ve yapay zekâ ile ilgilenen bilim insanlarının da dikkatini çekmeye başlamıştır. Bu ilginin pek çok sebebi vardır. Bunlardan ilki, artgönderimin dilin hem sözdizimsel hem anlambilimsel hem de edimbilimsel yönleriyle etkileşim halinde olmasıdır. Artgönderimsel ilişkilere olan ilginin bir diğer sebebi ise, artgönderim insan beyninin/zihninin doğasını anlamamız konusunda önemli bilgiler sunan ve Chomsky'nin (1981; 1982; 1986; 1995) dilbilimin en temel sorunu olarak gördüğü dil edinimi konusundaki sorularına yanıtlar bulunabilecek çok önemli bir inceleme alanı olarak görülmesidir. Chomsky'e göre, belirli yönleriyle artgönderim çalışmaları, insanların doğuştan, içsel ve bilinçsiz bir dil bilgisiyle donanmış olarak doğdukları konusundaki sava destek olabilecek kanıtlar sağlamaktadır (1981; 1982; 1986; 1995). Doğuştanlık varsayımına destek olabilecek kanıtlar sağladığı için artgönderim, 'beyne açılan bir pencere' olarak görülmüştür. Bu sebeplerden dolayı, artgönderim bu alanlarda üretilen varsayımların kanıtlanması için elverişli bir deneme alanı olarak görülmektedir. (Huang 2000).

Bu çalışmada artgönderimlerin, dilin sözdizimsel yönüyle olan ilişkileri üzerinde durulacak ve bazı söylem parçaları üzerinden Türkçe için algoritmik bir çözüm geliştirilecektir. Çalışmanın kuramsal çerçevesi Hobbs (1976; 1977; 1978) tarafından geliştirilen sözdizim tabanlı Naif Algoritma'ya dayanmaktadır. Sözdizimsel bir temele dayanan bu çalışmaya, Chomsky'nin Bağlama Kuramı (1981; 1982; 1986, 1995) ilkelerinin de uyarlanmasıyla Türkçe için oluşturulmuş kurallarla, Türkçe söylem yapısında 3. kişi adıları için kullanılan artgönderimsel biçimlerin çözümlenmesine çalışılmıştır. Sözdizimsel olarak denetlenen artgönderim kullanımları çalışma kapsamı içine alınmıştır.

Artgönderim terimi Yunancada “geriye taşımak” anlamına gelen αναφορα sözcüğünden gelmektedir. Çağdaş bilim yaklaşımları içinde bu terim, iki dilsel ifade arasındaki ilişkiyi açıklamakta kullanılır. Birbirleriyle ilişkili bu ifadelerden, yorumlanması bir diğerine bağlı olan ifadeye artgönderim, diğerine ise öncül denilmektedir (Huang 2000). Bir metindeki bütünlük içinde, daha önce ve daha sonra aynı sözcük, kavram ya da düşünce aynı biçimde ya da farklı biçimde yeniden kullanılabilir. Her metinde başka öğelere göre yorumlanabilecek artgönderimsel birimler bulunur. Bu tür yapılar bağlamsal bir durumu ilgilendirir ve sıralı tümcelerde dil ekonomisi adına yapılır (Günay,2003).

Ali, *Ahmet*₁'i ziyaret edecekti. Çünkü *ona*₁ yardım etmeliydi. (1.1)

Yukarıdaki tümcede geçen *ona* adılının gönderim amaçlı olarak kullanıldığı ve neye gönderimde bulunduğu bir insan tarafından çok çabuk ve kolay bir şekilde anlaşılabilir. Yani ikinci tümcede geçen *ona* adılının, ilk tümcedeki *Ahmet*'e gönderimde bulunduğunu rahat bir şekilde anlayabilmekteyiz. Ancak bu anlama yetisinin altında yatan birçok işlem vardır. Bu işlemlerin ne olduğu sorusu bilişsel bilimlerin konusu olup psiko-linguistik deneyler ile çözümlenmeye çalışılmaktadır. Esasen bu tip yapıların çözümlenmesi için insan beyninde ne tip bir özelleşmiş yapı bulunduğu tam olarak bilinmemektedir. Bu yapılarda, artgönderimsel biçimlerin çözümlenmesinin bilgisayarda modellenmesi için çeşitli sistemler kurulmuştur. Bu çalışma da bu amacın bir parçası olarak, yirmi yıldan fazla bir süredir bilimin değişik

dallarını meşgul eden ve zorlu bir konu olan, artgönderim çözümlemesini sözdizimsel tabanlı bir algoritma ve Bağlama Kuramının birlikte kullanılmasıyla bilgisayar ortamında gerçekleştirebilen bir model oluşturulmuştur.

Artgönderimsel biçimlerin çözümlenmesi, doğal dil içerisinde geçen önceki ifadelerin içerdiği bilgi ile sonradan gelen bilginin birleştirilerek nesne veya olay hakkında daha çok bilgi çıkarımı yapılmasını sağlayan işlemler bütünüdür. İnsanların çoğu zaman farkında bile olmadan yaptıkları ve düzgün bir iletişimin temelini oluşturan artgönderim çözümlemesi, günümüzde pratik amaçlı olarak doğal dil işleyen çeşitli hesaplamalı sistemlerde önemli bir yer tutmaktadır (Mitkov 2002). Artgönderim çözümlemesinin pratik amaçlarla kullanılabileceği sistemlere örnek olarak makine çevirisi, metin özetleme, bilgi çıkarımı ve soru cevaplama sistemleri verilebilir. Bu sistemlerde elde edilecek başarı, büyük oranda, işlenen söylem yapısı içinde geçen ifadeler arası gönderimlerin doğru çözümlenmesine bağlıdır.

Artgönderim çözümlemesine ait sürdürülmekte olan çalışmalar, sözdizimsel, anlambilimsel ve gerçek dünya bilgisine dayalı olan bilgi-yoğunluklu yaklaşımlar ve ön bilgiye dayanmayıp metin içinde saklı olan bilgiyi kullanan bilgi-yoksunu yaklaşımlar olarak iki ana kategori altında toplanırlar. Bu çalışmada, sözdizimsel tabanlı bilgi-yoğunluklu yaklaşımların kullanılmasıyla Türkçe için adıl çözümlemesi yapabilen bir sistem tasarlanmıştır. Bilgisayarlı bir alan içerisinde gerçekleştirilen bu sistem, deneylerle ortaya atılan diğer yaklaşımlarla kıyaslanacak ve detaylı bir şekilde tartışılacaktır.

BÖLÜM 2

TEMEL KAVRAMLAR VE TERMİNOLOJİ

Bu bölümde, artgönderime ilişkin tanımlamalar yapılmadan önce, bu tanımlamalarda sözü geçen *metin*, *söylem*, metinlerin biçimsel ve anlamsal bütünlüğünü oluşturmada rol oynayan *tutarlılık* ve *bağdaşıklık* ölçütleri gibi temel kavramlar açıklanacaktır. Bağdaşıklık ölçütünün dilsel öğelerinden biri sayılan artgönderim ve artgönderim çeşitleri detaylı bir şekilde incelenecektir. Daha sonra Türkçe ve diğer diller için yapılmış olan dilbilimsel artgönderim çalışmalarından bahsedilecektir.

2.1 Metin, Söylem, Tutarlılık ve Bağdaşıklık

Metin, en genel anlamda, bir tümceden daha uzun olabilen dilsel bütün olarak tanımlanabilir. Ancak terimin anlamı, daha birçok kavram gibi, farklı dilbilimciler tarafından farklı biçimlerde özelleştirilmiştir. Şöyle ki *metin*, kimi zaman, bir dilin türlü kullanım alanlarından derlenmiş örneklerinin dil bilgisi ve kuramsal dil bilimi araştırmalarında kullanılmak üzere bilgisayar tarafından okunabilecek biçimde bir araya getirilmiş olan *bütüncenin*, çoğu zaman da bir veya birçok tümceden oluşan, başı ve sonu olan bildiri anlamına gelen *söylemin* eşanlamlısı olarak karşımıza çıkmaktadır. Biz bu çalışmada ‘söylem’ ve ‘metin’ terimlerini eşanlamlı olarak kabul ettik.

Metin kavramını, daha teknik ifadelerle, kendisini oluşturan tümce dizilerinin birbirlerine *tutarlılık* ve *bağdaşıklık* ölçütleriyle bağlanarak bir anlam bütünü oluşturmasıyla meydana gelen, belli bir amaçla üretilmiş, başı ve sonu kesin çizgilerle sınırlandırılan yazılı ya da sözlü bir dilsel ürün olarak tanımlayabiliriz. Bir tümce için dilbilgisi kurallarına uygunluk ne ise, bir metin için de bağdaşıklık ve tutarlılık

ölçülerine uygunluk odur. Metin içinde her ne kadar birbirlerine bağlı olsalar, birbirlerini bütünleseler de, bağdaşıklık ve tutarlılık olguları düzey açısından birbirinden farklıdır. *Bağdaşıklık* metnin yüzeyinde, dilsel öğeler aracılığıyla görülebildiği halde, *tutarlılık* derin yapıda oluşan anlamlar arasındaki mantıksal bağlantıdır ve tutarlılığı gösteren belirli dilsel öğeler bulunmamaktadır.

Tutarlılık sözkonusu olduğunda, metinde sözü edilen izleğin dil dışı dünya gerçekleriyle kurduğu ilişkilerin incelenmesi gerekmektedir. Böylelikle, *bağdaşıklık*, kendisini yalnızca metnin yüzeyinde somut bir biçimde gösterirken, *tutarlılık* olgusu soyut anlam düzeyinde incelenir. Bu nedenle, *tutarlılık* ilk bakışta metnin yüzeyinde algılanamaz, belli bir yorum süreci gerektirir. Bu yorumu yapacak kişinin karmaşık bir belirtiler ağına dayanarak, sahip olduğu dilsel, söylemsel ve ansiklopedik bilgiler bütünlüğünü harekete geçirmesi gerekmektedir (Maingueneau, 1991). Bu olgu metinde yer alan eylem, durum ya da olayların, yani göndergelerin, alıcının dış dünya gerçekleriyle örtüşmesi gerekmektedir. Aşağıdaki örneği inceleyelim:

Artık kışlar eskisi gibi soğuk geçmiyor. Bilim adamları küresel ısınmanın (2.1)
birçok olumsuz etkisi olduğunu söylüyorlar. Okyanuslardaki adalar birkaç
yıla kadar sular altında kalacakmış.

Yukarıdaki söylemde gördüğümüz gibi, yazan öznenin ve okuyucuların dış dünya gerçekleriyle ilgili bilgileriyle bütünleşip birbirlerine eklenerek, kendi içinde bir anlam bütünlüğü oluşturmaktadır. Tutarlılık aynı zamanda, tümceler birbirlerine göre yorumlanabilmesiyle de oluşur. Metnin hem kendisini üreten özne; hem de alıcısı tarafından aynı ölçüde tutarlı olması için, dış dünya bilgilerinin her ikisi tarafından da paylaşılması gerekmektedir. Yukarıdaki örneğe bakıldığında bu tümce dizisini, küresel ısınma konusunda bilgisi olan alıcılar tarafından, gerekli mantıksal bağlantılar kurularak, şu biçimde yorumlanacaktır: kışlar küresel ısınmadan ötürü eskisi kadar soğuk geçmiyor ve yine küresel ısınma nedeniyle buzullar eriyor, böylece denizler yükseliyor ve bu nedenle de bazı adalar sular altında kalacak. Ancak, aynı tümce dizisi, küresel ısınma konusunda bilgisi olmayan alıcılara tutarsız gelebilir, çünkü onlar, bu

metnin evreniyle örtüşen bir dış dünya bilgisine sahip olmadıkları için metinde yer alan anlamsal bağlantıları kuramayabilirler (Onursal, 2003).

Yalnızca anlama ya da içerikle elde edilen bu metin bütünlüğü, biçimsel araçlarla da pekiştirilebilir. Bu araçlardan bir tanesi de bağdaşıklık ölçütüdür. Bağdaşıklık ölçütü, “önergelerin birbirlerine bağlanması ve metnin çizgisel biçimde düzenlenmesi sonucunda ortaya çıkar” (Maingueneau, 1996). Başka bir deyişle, bir metnin uyumlu olabilmesi için, kendisini oluşturan çeşitli bölümlerin bir dilsel bütünlük sağlayacak biçimde birbirlerine bağlanmaları gerekmektedir.

Metnin uyumluluğunu sağlayan dilsel öğeler, hem bölümlerin kendi içlerinde, hem de bölümler arasında sözdizimsel, anlamsal ve mantıksal bağlantılar kurulmasına yardımcı olurlar. Bağdaşıklık sağlayan dilsel öğeler metnin bir anlam bütünlüğü olarak algılanmasını sağlar. Bu öğeler arasında, eylemlerin zaman uyumlarını, gösterge dizilişlerini, *artgönderimleri*, *öngönderimleri* ve bağlantı öğelerini sayabiliriz. Bunlardan *artgönderimler* bu tezin kapsamını oluşturmaktadır. İleriki bölümlerde ayrıntılı olarak inceleyeceğimiz *artgönderimler*, bağdaşıklık için de örnek teşkil edecektir.

2.2 Artgönderim

Bir metnin bağdaşıklık özelliği taşıyabilmesi için içermesi gereken dilsel öğeler arasında bulunan *artgönderimler*, “daha önce kullanılmış bir biçime (öncül) daha sonra gelen bir öge (...) aracılığıyla gönderme” (Vardar, 1998) yapılmasını sağlayan dilsel öğelerdir. Bu öğeler, metnin konusunu oluşturan bilgilerin yeniden ele alınmasını sağlayarak, onun çizgiselliğini ve sürekliliğini oluşturur.

Bir başka tanıma göre ise *artgönderim*, bir söylem içindeki önce gelen bazı öğeleri işaret eden bağdaşıklık olarak tanımlanmıştır (Halliday ve Hasan, 1976). İşaret eden kelime ya da öbek, *artgönderim*; *artgönderimin* gönderme yaptığı öge, *öncül* olarak adlandırılır. Öncüller gerçek dünyada veya doğal dil cümlesini kuran kişinin zihninde

var olan olay veya nesnelere. Artgönderimin öncülünün tespit edilmesi işlemine ise *artgönderim çözümlemesi* denilir.

Ali Ayşe'ye kitap verdi. Ayşe de ona kalem verdi. (2.2)

Yukarıdaki ikinci tümcede geçen *ona* ifadesi artgönderimdir ve *Ali* öncülüne gönderimde bulunur. *Ona* artgönderiminin *Ali*'ye mi yoksa *Ayşe*'ye mi gönderimde bulunduğu çözümlemesi işlemi bir artgönderim çözümlemesi örneğidir. İnsan beyni sezgisel olarak artgönderimlerin neyi veya kimi işaret ettiğini rahatlıkla bilebilmesine karşın bu işlemin bir sistem tarafından yapılmasını sağlamak zor bir iştir. Bunun için birçok kuraldan oluşan bir sistemin oluşturulması gerekmektedir.

Artgönderim bir öncülü işaret ettiğinde, eğer her ikisi de gerçek dünyada aynı nesne veya kavrama karşılık geliyor ise, *eşgönderimli* diye adlandırılırlar (Mitkov, 2002). Aynı varlığa referansta bulunan isim öbeklerinin de *eşgönderim zinciri* oluşturduğu söylenir. Ancak her ne kadar artgönderim ile eşgönderim olguları birbirleriyle örtüşüyor gibi görünse de, eşgönderimin artgönderimden farklı olduğunu belirtmemiz gerekir. Artgönderim, eşgönderimden farklı olarak tek ve bütünleşik bir metin içinde gerçekleşir. Örnek olarak şu tümceleri inceleyelim.

Atatürk, Türkiye Cumhuriyeti'nin ilk cumhurbaşkanıdır. (2.3)

Ulu Önder, onbeş yıl süreyle cumhurbaşkanlığı görevini sürdürmüştür. (2.4)

2.3 ve 2.4 tümcelerinde *Atatürk* ve *Ulu Önder*, her ne kadar gerçek dünyada aynı varlığa gönderimde bulunsalar da, biri diğerine gönderimde bulunmamaktadır. Bu durumda artgönderim değil, eşgönderim yapılmıştır. Bir söylem içinde bir artgönderim ve öncülü arasında bağlantının kurulması işlemine nasıl artgönderim çözümlemesi deniliyorsa, bütün doğru öncüllerin veya eşgönderim zincirlerinin bulunmaya çalışılması işlemine de *eşgönderim çözümlemesi* denilmektedir (Güner, 2008). Bu tez kapsamında eşgönderimler daha sonra ele alınmayacak olup, bu ön bilgiyle aradaki fark konusunda bir fikir verilmesi amaçlanmıştır.

2.3 Artgönderim Çeşitleri

Artgönderimler, öncüllerin konumuna göre, gönderim yönüne göre ve gönderimde bulunulan öncüllerin dilbilgisel kategorilerine göre yani biçime göre farklı şekilde sınıflandırılırlar. Artgönderimler, öncülün konumuna bağlı olarak da iki gruba ayrılır: *tümce-içi artgönderim* ve *tümceler-arası artgönderim*. Eğer artgönderim ve öncülü aynı tümce içinde yer alıyorsa, *tümce-içi artgönderimden*; eğer öncül artgönderimin bulunduğu tümceden farklı bir tümcede yer alıyorsa *tümceler-arası artgönderimden* bahsedilir. Dönüşlü adılar, tümce-içi artgönderimlere örnektirler:

Ahmet_i kendi_i kamyonunda bir şoför gördü. (2.5)

Tümce 2.5’de de görüldüğü gibi, *kendi* dönüşlü adılı aynı tümcede yer alan *Ahmet’e* gönderimde bulunmaktadır. İyelik adıları da çoğunlukla tümce-içi artgönderim olarak kullanılırlar ve aynı tümcecikte artgönderim olarak yer alırlar. Buna karşılık, kişisel adılar ve isim öbekleri, genellikle öncülleri aynı kompleks tümcenin bir önceki tümcede yer alan tümceler-arası artgönderimler olarak rol alırlar.

Ahmet_j onun_k kamyonunda bir şoför gördü. (2.6)

Tümce 2.6’da yer alan *onun* kişi adılı bir önceki tümcelerde yer alan başka bir kişiye gönderimde bulunur.

Gönderimler, yön üzerinden yapılan sınıflamaya göre artgönderimler ve öngönderimler diye ikiye ayrılır. Artgönderimlerin yanı sıra, onlar kadar sık kullanılmayan öngönderimler, artgönderimlerin tersine, anlamlı bir birimin yerini tutan bir ögenin söylemde ondan daha önce anılmasıyla ortaya çıkar (Vardar, 1998).

Size şunu da söylemeliyim ki, ülkenize yaptığım ziyaretten çok memnun kaldım. (2.7)

Tümce 2.7’de de görüldüğü gibi, sözü edilen konu yada kişi, kendisinden önce yer alan bir adıyla belirtilmiştir.

Artgönderimler, söylem içinde değişik biçimlerde var olabilirler. Mitkov (2002), artgönderimleri biçime göre altı ana gruba ayırmıştır:

1. Adılsal Artgönderim,
2. Sözcüksel Ad Öbeği Artgönderimi,
3. Ad Artgönderimi,
4. Eylem Öbeği Artgönderimi,
5. Belirteç Artgönderimi,
6. Boş Artgönderim.

Artgönderim biçimlerine sırasıyla göz atalım.

1. Adılsal Artgönderimler: Biçime göre en yaygın olan artgönderim çeşiti adılsal artgönderimlerdir. Önce bir ad kullanılır, daha sonra ilgili adın yerine adıl kullanılır.

Babam koyun gütmeye gideceksin dedi. *O* şehre inecek bugün. (2.8)

Babam ve *o* kelimeleri arasında hem göndergenin yinelenmesiyle yapılan artgönderim hem de adıyla yapılan artgönderim vardır. *Babam* sözcüğünden sonra onu işaret etmek amacıyla *o* adılı kullanılmıştır.

Annesi “Ne oldu oğlum, ne bu hâlin?” dedi. *Çocuk babasının* karşısında (2.9)
durup “Dedem ölmüş.” diye bağırdı. Sanki yer sarsıldı, ev yıkıldı. Kimse bir şey söyleyemedi. Babası, başını öne eğdi, gizli gizli ağladı bir süre.
Onlara bir ölüm sessizliği eşlik etti.

Anne, baba, çocuk kelimeleri ile *onlara* adılı arasında adılsal artgönderim ilişkisi vardır.

Dede, bugün ceviz çırpmaya gelemeyeceğim. “*Ali Ağa* niye küser *bana*, (2.10)
ben ne yaptım *Ali*’ye?”

Dede sözcüğünden sonra onu işaretten *bana*, *ben* adıkları kullanıldığı için adısal artgönderim vardır. Ayrıca *Ali Ağa* ve *Ali* kelimeleri arasında gönderenin yinelenmesiyle yapılan artgönderim mevcuttur (Balci, 2006).

Türkçe’de adılar, 6 sınıfta incelenebilir. Bunlar şunlardır:

1. Kişi Adıları (ben, sen, o, biz, siz, onlar)
2. İyelik Adıları (benim, benimki, senin, seninki, onun, onunki, bizim, bizimki, sizin, sizinki, onların, onlarınki)
3. Dönüşlü Adılar (kendi, kendim, kendin, kendisi, kendimiz, kendiniz, kendileri)
4. İşaret Adıları (bu, şu, bunlar, şunlar)
5. Soru Adıları (ne, neyi, kim, kimin, hangi, hangisi, kaç, kaçını, nereye, ...)
6. Belgisiz Adılar (biri, birisi, birileri, birkaçı, kimisi, kimse, herkes, tümü, pek çoğu, hepsini, bazısı, hiçbirini, hiçbirisi, her biri, başkası, başkaları, bir kısmı, öbürleri, ...)

Türkçe adısal artgönderimler, Tablo 2.1’de de görülebileceği gibi bu 6 sınıfın, 5’inde oluşur. Kişi, iyelik ve dönüşlü adıların üçüncü kişi için olanları ile işaret ve soru adıları, adısal artgönderimlerdir.

Tablo 2.1 Türkçe Adısal Artgönderimler

Türkçe Adısal Artgönderimler				
Kişi	İyelik	Dönüşlü	İşaret	Soru
o	onun	kendi	bu	ne
onu	onunki	kendisi	şu	neyi
ona	onların	kendileri	bunlar	kim
onda	onlarınki		şunlar	kimi
ondan				kimin
onlar				hangi
onları				hangisi
onlara				...
onlarda				...
onlardan				...

- Kardeşler balık tutmaya çalışıyorlar. (2.11)
- Ağacın yaprakları arasında bir kuş *onları* gözlüyor. (2.12)
- O* da bir balık yemek istiyor. (2.13)
- Kardeşler *onu* umursamadan işlerine devam ediyor. (2.14)

Tümce 2.12, 2.13 ve 2.14'deki *onları*, *o* ve *onu* artgönderimleri, Tükçe'deki adısal artgönderimlere birer örnektir. İkinci tümcedeki *onları* artgönderimi, bir önceki tümcedeki yani tümce 2.11'deki *kardeşlere* gönderimde bulunur. Üçüncü tümcedeki *o* artgönderimi ise yine bir önceki tümcedeki yani tümce 2.12'deki *kuşa* gönderimde bulunur. Son olarak dördüncü tümcedeki *onu* artgönderimi ise bir önceki tümcedeki yani tümce 2.13'deki *o* diye geçen *kuşa* gönderimde bulunur.

2. Sözcüksel Ad Öbeği (AÖ) Artgönderimleri: Sözcüksel AÖ artgönderimleri, artgönderimin belirli bir AÖ veya bir özel ad olduğu durumdaki artgönderimlerdir. AÖ-artgönderimi ilişkisi içinde hem artgönderim, hem öncül, gönderimsel ifadelerdir ve AÖ'lerinin içinde yer alırlar. AÖ-artgönderimi, anlambilimsel olarak özdeş gönderimler olarak da tanımlanabilen öncül ile artgönderimin gönderimlerinin özdeş olduğu durumlardır (Huang 2000).

“(...) genç aktris *Juliette Drouet*'nin gözleri de, çalışmaları dikkatle izleyen *Victor Hugo*'nun üzerindedir (...) *Drouet* ile *Hugo*'nun tanışmasına ve bir yarım yüzyıl sürececek derin bir aşkın doğmasına vesile olur (...) Nasıl özetlemeli *Juliette*'nin o tarihe kadarki öyküsünü? (...)” (2.15)

İlk tümcede yer alan *Juliette Drouet*'dan, daha sonraki tümcelerde *Drouet* ve *Juliette* diye bahsedilmiştir ve bu sayede sözcüksel artgönderim ilişkisi yaratılmıştır.

Eşek sırtında, ağır aksak gelen biri var. Uzaktan seçemedi gözleri. (2.16)

Yaklaşınca tanıdı. “*Mahmut* geliyor, *Kösenin Mahmut* bu.” Altmış yaşlarında, ağarmış sakalı, yüzü kırış kırış, başında sekiz köşe kasketi...
”Bilirim *kardaşım*, ne yapacağın işte yaşlılık...”

Anlatım bir sözcük ya da sözcük grubu ile verilip, daha sonra bu sözcük farklı biçimlerde yinelendiği için; *ağır aksak gelen biri, Mahmut, Kösenin Mahmut* kelime grupları ve *kardaş* kelimesi arasında sözcüksel artgönderim vardır.

Şu toprak, ne kadar da *sıcak, samimi, sevecen bir ana* gibi. (2.17)

Toprak için önce *sıcak* ve *samimi* daha sonra da *sevecen bir ana* tabiri kullanılarak, sözcüksel bir artgönderim yapılmıştır.

Çocuk, yorgun bitkin bir halde eve vardı. Annesi “Ne oldu *oğlum*,
ne bu halin?” (2.18)

Çocuk ve *oğul* kelimeleri arasında sözcüksel artgönderim ifade edilmiştir (Balcı, 2006).

3. Ad (A) Artgönderimi : A-artgönderimlerin AÖ-artgönderimlerden farkı, olası gönderimsel ifadeler olmamalarıdır. AÖ-artgönderimi, anlambilimsel açıdan ‘gönderimin özdeşliği’ olarak tanımlanırken, A-artgönderimi, bir kavramın çağrıştırdığı kapsama giren niteliklerin veya taşıdığı özelliklerin bütünü (TDK Sözlük) olan işlemin özdeşliği olarak tanımlanabilir. Başka bir deyişle, öncül ile artgönderimin işlemlerinin özdeş olduğu durumlardır. *Bir-artgönderimi*, A-artgönderimine örnek olarak verilebilir.

Ali yeni *bir CD* aldı, ancak Ahmet ikinci el *bir CD* aldı. (2.19)

4. Eylem Öbeği (EÖ) Artgönderimleri: EÖ-artgönderimleri, öncüllerin bir eylem öbeği olması durumunda oluşurlar.

“Genel Kurmay Başkanlığı (...) gelen askeri heyetin ABD’nin olası Irak
operasyonu ile ilgisi olmadığını *açıkladı* (...) *Açıklamada* şu bilgilere
yer verildi (...)” (2.20)

5. Belirteç Artgönderimleri: Belirteç artgönderimleri, öncüllerin bir belirteç olması durumunda oluşurlar. Bunlar, “orada” gibi bir yerle ilgili olabilirler veya “o zaman” anlamında zamanla ilgili olabilirler.

Bahçe, dereye yakındı. *Bahçenin* giriş yerindeki çalıları bir kenara atıp (2.21)
içeriye girdi. Her gün ta köyden kalkıp da *buraya* gelmek yoruyor beni.

Tümce 2.21’de metnin çeşitli yerlerinde kullanılan *bahçe* kelimesiyle metnin son bölümlerinde kullanılmış olan *buraya* kelimesi arasında belirteçle yapılan artgönderim mevcuttur.

Bahçeye varmaya yakın yavaşladı. *Bahçeye* girdiğinde sağa sola bakındı, (2.22)
kimseyi göremedi. Çocuk, dedesinin yüzündeki tozu toprağı temizlemek
istedi ama daha fazla bakamadı, *orada* duramadı.

Tümce 2.22’de ise yine *orada* ve *bahçe* kelimeleri arasında aynı türde bir artgönderim yapılmıştır (Balcı, 2006).

6. Boş Artgönderim: Boş artgönderimler, tümceden çıkartılmış yani ihmal edilmiş olan ama anlam olarak anlaşılabilir artgönderimlerdir. Başka bir deyişle, söylemde yüzeyde görünmeyen fakat dilbilgisi kurallarınca açığa çıkarılan artgönderim tipidir. Aşağıdaki söylemi ele alalım.

*Alin*_i baleye gitmek için hazırlık yapıyordu. (2.23)

[Ø_i[Ø_i Çantasına gerekli eşyaları koyarak] evdençıktı.] (2.24)

Tümce 2.24’de özne pozisyonundaki ilk öge boş artgönderimdir ve *Alin* ile eşgönderime sahiptir. Benzer bir şekilde boş artgönderim pozisyonunda bulunan ikinci Ø gönderimi de boş artgönderime örnektir (Yıldırım, 2008).

2.4 Dilbilimsel Artgönderim Çalışmaları

Bu bölümde, temelde İngilizce için yapılmış dilbilimsel artgönderim çalışmalarından ve çeşitli kuramlardan bahsedilmiştir ve ardından da Türkçeye ilgili yapılmış dilbilimsel artgönderim çalışmalarına yer verilmiştir. İlk olarak birçok yöntem tarafından kullanılan ve bizim de bu çalışmada ilkelerine dayandığımız Bağlama Kuramı (Chomsky, 1981) detaylı bir şekilde anlatılmıştır. Daha sonra, diğer dilbilimsel artgönderim çalışmalarına kısaca değinilmiştir.

2.4.1 Bağlama Kuramı

Chomsky tarafından (1981) sunulan İlkeler ve Değiştirgenler Kuramı'nın bir parçası olan Bağlama Kuramı, artgönderimlerin ve adılların nasıl yorumlanacağı konusunda ipuçları verir. Bağlama Kuramının temel kavramı gönderimdir. Gönderim, bir ad ile bu adın sözlüksel içeriğinin ait olduğu kişi, nesne, varlık vb. arasındaki özdeşlik ilişkisidir. Dil kullanıcıları bu ilişkiden hareketle ad ile gönderimde bulunduğu öğeyi özdeş olarak yorumlar. Her ad ve her gönderim kaynağı belirli bir dizine aittir. Adlar ile gönderimde buldukları kaynaklar aynı dizine aittir. Başka bir deyişle, gönderimde bulunan AÖ ile, eğer yapıda varsa, gönderim kaynağı eşdizinlenir (Uzun, 2000).

Ali_i kendini_i vurdu. (2.25)

Ali_i onu_j vurdu. (2.26)

Tümce 2.25'de *Ali* ve *kendini* eşdizinlenir. Gönderim kaynağı tümce dışında olduğunda da eşdizinleme kurulabilir. Bu durumda tümce içindeki AÖ'leri farklı dizinler taşırlar. Tümce 2.26'da olduğu gibi, *Ali* ve *onu* farklı dizinlerdedir.

Tümcede bulunan her adın bir gönderimi bulunur. Ancak bu gönderimler, yukarıda da belirtildiği gibi, kaynaklarının tümce içinde ve dışında olması bakımından farklılaşırlar. Bu açıdan, Bağlama Kuramı AÖ'lerini aşağıdaki gibi adlandırmıştır (Cook ve Newson 1996; Uzun, 2000):

Adılsıl: Gönderimi tümce dışında olan adılar, adılsıl

Gönderge: Gönderimi tümce içinde olan işteş ve dönüşlü adılar, gönderge

Gönderimsel İfade: Ali gibi sıradan adlar ise gönderimsel ifade olarak tanımlanır.

Şu örnekleri inceleyelim:

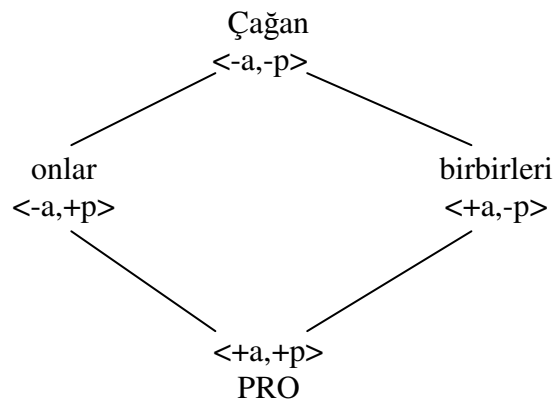
Çocuklar_i birbirlerine_i baktılar. (2.27)

Çocuklar_i onlara_j baktılar. (2.28)

Çocuklar Çağan'a baktılar. (2.29)

Tümce 2.27'deki *birbirleri* işteşlik adılının gönderimi tümce içinde olduğundan gönderge, tümce 2.28'deki *onlar* adılının gönderimi tümce dışında olduğundan adılsıl ve tümce 2.29'daki *Çağan* ise gönderimsel ifadedir.

Chomsky (1981), Bağlama Kuramı olarak adlandırdığı yaklaşımında, isim öbeklerini iki temel özelliğin varlığına göre 4 gruba ayırmış, Bağlama Kuramı prensipleri ile de gönderim bağlantılarının nasıl olması gerektiğini ortaya koymuştur. İsim öbeklerinin artgönderimsel ve adılsıl özellikler taşıyan öğeler olduğunu ve öğelerin bu özelliklerden bir kısmı veya hepsini taşıyabileceği gibi hiçbirini de taşıyamayabileceğini söylemiştir. Bu konuda çıkardığı model aşağıdaki gibidir:



Şekil 2.1 Bağlama Kuramı İçinde İsim Öbeklerinin Sınıflandırılması

Şekil 2.1’de özetlenen gönderimsel ifadeler, Chomsky’nin dilbilgisi yaklaşımı içerisinde olan bağlama kuramına uymak zorundadırlar. Bağlama kuramı temelde şu 3 ilkeden oluşur:

- Yalnız artgönderimsel özellik taşıyan (yani $\langle +a, -p \rangle$ olan) dönüşlü artgönderimler, kendilerine en yakın konumdaki öge ile bağlanabilirler:

Ali_i kendini_i yaraladı. (2.30)

Tümce 2.30’daki *kendini* dönüşlü adılı *Ali* ile artgönderimsel ilişki kurmak zorundadır.

- Yalnız adılsıl özellik taşıyan (yani $\langle -a, +p \rangle$ olan) adıllar, kendilerine en yakın olan hariç diğer özne görevindeki öğelerle bağlanabilirler:

Ahmet_i [Ali’nin_j onu_{i/k} yaraladığını] biliyor. (2.31)

Tümce 2.31’deki *onu* adılı *Ahmet* ile artgönderimsel ilişkiye girebilmektedir.

- Ahmet gibi doğrudan gerçek dünyada referansını bulan öğeler başka bir ifadeyle bağlanamaz. Bu tip öğeler, artgönderimsel veya adılsıl özelliklerden herhangi birini taşımadıkları için (yani $\langle -a, -p \rangle$ olduklarından), hiçbir gönderimsel ilişkiye giremezler.

Belirtilen bu üç ilkenin haricinde Şekil 2.1’de gösterilen bir diğer gönderimsel ifade de, hem artgönderimsel hem de adılsıl özellik taşıyan (yani $\langle +a, +p \rangle$ olan) ve PRO ile ifade edilen öğelerdir. Bunlar tümce içinde ses içeriği kazanamazlar. Örneğin,

[PRO Sevilmek] güzeldir. (2.32)

Tümce 2.32’de *sevilmek* eyleminden önce gizli ve kişi-dışı bir öge vardır.

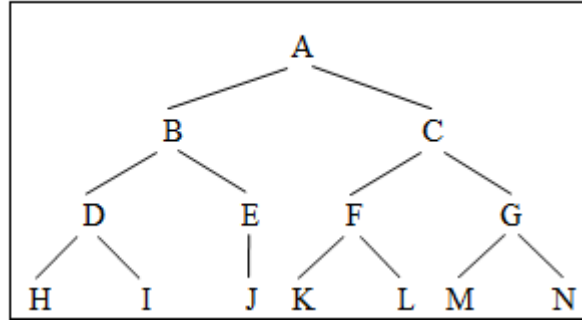
Diğer bir deyişle Bağlama Kuramına göre, adılsıllar, göndergeler ve gönderimsel ifadeler arasındaki en önemli fark bağlı oldukları tümce alanıdır. Göndergeler tümce içinde bağlı olurlar; adılsıllar başka tümceciklerdeki AÖ’lerine ya da tümce dışındaki

varlıklara bağılırlar. Bağlama Kuramı, bağlamanın yapılabileceği ve yapılamayacağı alanları belirlemeye çalışır ve bağlamayı şöyle tanımlar (Cook ve Newson, 1996; Uzun, 2000):

A, B'yi k-komutlandırıyorsa ve
A ile B eşdizinli ise
A, B'yi bağlar.

K-buyurmanın yapısal ilişkisi, bağlama için önemli bir rol oynar. Bu ilişki şöyle olmaktadır: Bir A düğümü, bir B düğümünü k-komutlandırıyorsa:

- (i) A, B'nin üstünde değildir.
- (ii) B, A'nın üstünde değildir.
- (iii) A'nın üstündeki ilk dallanan düğüm, B'nin de üstündedir.



Şekil 2.2 K-komutlandır Kavramı için Örnek Bir Gösterim

Şekil 2.2, k-komutlandır kavramının anlaşılması için bir örnek teşkil eder. Şöyle ki:

- B, C'yi ve C'nin altındaki her düğümü k-komutlandırır.
- C, B'yi ve B'nin altındaki her düğümü k-komutlandırır.
- D, E'yi ve J'yi k-komutlandırır, ancak C veya C'nin altındaki herhangi bir düğümü k-komutlandırmaz.
- H, I dışındaki başka bir düğümü k-komutlandırmaz.

Bağlama Kuramı çerçevesinde tümce kavramı yerine yerel alan kavramı kullanılmaktadır. Yerel alan bağlanmanın sözdizimsel boyutunu belirleyen koşul durumundadır. Bu kavramı kullanarak Bağlama Koşullarını şu şekilde özetleyebiliriz:

A İlkesi: Bir gönderge yerel alanı içinde bağlı olmalıdır.

B İlkesi: Bir adıslı yerel alanı içinde özgür olmalıdır.

C İlkesi: Bir gönderimsel ifade her yerde özgür olmalıdır.

Çağan_i kendine_i güvenir. (2.33)

Kerem_i [Çağan'ın_j kendine_{i/j} güvendiği]-ni bilir. (2.34)

Kerem_i onu_{i/j} seviyor. (2.35)

Çağan_i [Kerem'in_j onu_{i/*j/k} sevdiği]-ni biliyor. (2.36)

O_i [onun_j Çağan'ı_{i/*j/k} sevdiği]-ni biliyor. (2.37)

Tümce 2.33'de ve tümce 2.34'de görüldüğü gibi *kendi* göndergesi bulunduğu yerel alandaki *Çağan*'a gönderimde bulunurken, tümce 2.34'de yerel alanı dışındaki AÖ'lere bağlanamamaktadır; dolayısıyla bu tümceler Bağlamanın A İlkesine uymaktadır. Tümce 2.35'de ve tümce 2.36'da *onu* adıslı yerel alan içindeki AÖ'ne bağlanamaz ve bu nedenle de Bağlamanın B ilkesine uyar. Tümce 2.37'de *Çağan* ne ana tümcedeki ne de yerel alanındaki bir AÖ'ne bağlanabilir; dolayısıyla da Bağlamanın C İlkesine uyar.

Türkçe'nin basit tümcelerinde işteşlik ve dönüşlülük adılları, Bağlama Kuramının ilkelerine uymaktadır.

Ali_i kendini_i yaraladı. (2.38)

[Ali ve Ayşe]_i birbirlerini_i seviyor. (2.39)

Ali_i onu_{i/j} yaraladı. (2.40)

Ali_i Ayşe'yi_j yaraladı. (2.41)

Tümce 2.38 ve tümce 2.39'da görüldüğü gibi *kendi* ve *birbiri* adılları buldukları tümcelerdeki AÖ'lerine gönderimde bulunurken, 2.40'da *onu* adıslı ve 2.41'de *Ayşe*

gönderimsel ifadesi tümce dışına gönderimde bulunmuşlardır. Bu gönderimsel ifadelerin bileşik tümcelerdeki görünümüne bakalım:

Ahmet_i [Ali'nin_j kendini_{i/j} yaraladığı]-nı biliyor. (2.42)

[Ali ile Ayşe]_i [onların_j birbirlerini_{i/j} sevdiği]-ni biliyor. (2.43)

Ahmet_i [Ali'nin_j onu_{i*/j/k} yaraladığı]-nı biliyor. (2.44)

O [onun_j Ayşe'yi_{i*/j/k} sevdiği]-ni biliyor. (2.45)

2.42 ve 2.43 tümcelerinde dönüşlülük ve işteşlik adıllarının, buldukları tümce dışındaki AÖ'lerine bağlanamadığı görülmektedir; dolayısıyla bu tümceler Bağlamanın A ilkesine uymaktadırlar. Tümce 2.44'deki kişi adlı *onu*, yerel alanı içindeki AÖ'ne bağlanamaz ve bu nedenle de Bağlamanın B İlkesine uyar. Tümce 2.45'de *Ayşe* ne ana tümcedeki ne de yerel alanındaki bir AÖ'ne bağlanabilir; dolayısıyla da Bağlamanın C ilkesine uyar. Örneklerde görüldüğü gibi, hem basit hem de bileşik yapılı Türkçe tümcelerde Bağlama İlkelere uyulduğu gözlemlenmektedir. Ancak Türkçede bu ilkelerden sapmaların yaşandığı durumlar da söz konusudur. Bu sapmaları belirlemeye çalışanlardan Sezer (1991), Bağlamanın A ve B İlkelere yetersizliklerini belirlemiş ve bu ilkelerle ilgili AÖ türlerini ayrı ayrı incelemiştir (Çeltek, 2003).

2.4.2 Diğer Dilbilimsel Artgönderim Çalışmaları

Bunlar, Givón'un (1983) Konu Sürekliliği Modeli, Fox'un (1987) Aşamalı Modeli ve Ariel'in (1988) Ulaşılabilirlik Kuramı ve son olarak da artgönderimlerin yorumlanmasına ilişkin kural ve kısıtlamaların nasıl uygulandığını gösteren Merkezleme Kuramıdır (Grosz ve Sidner, 1986). Şimdi sırasıyla bu çalışmalara kısaca bir göz atalım.

2.4.2.1 Konu Sürekliliği Modeli

Givón (1983; 1985; 1990) tarafından tanımlanan Konu Sürekliliği Modeli, gönderimsel ifadelerin söylem içinde nasıl kullanıldığını açıklamayı amaçlar. Bu modelde temel kavram olan konu, en basit anlamda söylem içinde neden konuşulduğudur. Givón,

konuyu, söylem yapısı içinde ‘meseleleri, sıkıntıları, yaptıkları’ anlatılan kişiler olarak tanımlar (1983). Dolayısıyla, konular genellikle adsız sözcüklerdir.

Bu model, söylem içinde artgönderimsel kodlamanın konu sürekliliği ile belirlendiğini ileri sürmektedir. Konu sürekliliği, söylem içinde, çizgisel uzaklık (aynı göndergenin iki kullanımı arasına giren tümce sayısı), gönderimsel müdahaleler (araya giren başka göndergeler) ve konusal bilgi (söylem başkişisinin aynı kalması ya da değişmesi) olmak üzere üç temel etken tarafından ölçülür.

Çizgisel uzaklık ölçütü, söylem içinde konunun önceki ifadesi ile belirli bir dilbilgisel kodlama yöntemi ile belirtildiği yeni ifadesi arasındaki boşlukla ilgilidir. Bu boşluk, sola doğru olan, aradaki tümce sayısı olarak da açıklanabilir. Bu boşluk için en küçük değer bir tümce olarak belirlenmiştir. En yüksek değeri belirlemek, bir anlamda sınırı belirlemek, imkansızdır. Ancak kısa erimli bellek gibi psikolojik etkenlerin bu ölçüt ile karşılıklı etkileşim durumunda oldukları göz önüne alınarak gönderimsel boşluk için en büyük değer belirlenebilir. Givón, bu çalışması için bu değeri yirmi tümce olarak belirlemiştir. Dolayısıyla, eğer konu, bu iki erim arasında tekrar edilmezse otomatik olarak süresiz olarak tanımlanır. Başka bir deyişle, bu ölçüte göre çizgisel uzaklık ne kadar kısa olursa süreklilik o kadar yüksek olur (Givón, 1983).

Gönderimsel müdahaleler ölçütü, aynı söylem yapısında ‘hemen takip eden kesit’ içindeki diğer konuların/gönderimlerin, konunun ya da söylem başkişisinin belirlenebilmesi veya ulaşılabilmesi üzerindeki bozucu veya aksatıcı etkisi ile ilgilidir. Gönderimsel uzaklık için belirlenen en yüksek değer olan yirmi tümce, aradaki ilişkiyi en aza indirmek için yeniden belirlenmektedir. Gönderimsel müdahaleler ‘hemen takip eden kesit’ sola doğru bir ile beş tümce arasında olmalıdır. Bu belirleme, eğer konu kesit içindeki son beş tümce içinde sürekli veya baskın bir konumda ise ve tanımlanmasında bir belirsizlik yoksa, daha ileride devam eden kesit içinde araya giren gönderimlerin konunun belirlenmesine olumsuz bir etkisi olmayacağı varsayımı ile yapılmıştır. Bu ölçüte göre ‘hemen takip eden kesit’ içinde araya giren başka gönderimler bulunmuyorsa konu sürekliliği yüksek olur (Givón, 1983).

Yukarıda açıklanan iki ölçüt de, artgönderimsel ifadeden önce bulunan dilsel öğeler ile ilgiliydi. Üçüncü ölçüt olan konusal bilgi ise, konunun söylem içinde sonradan gelen tekrarlanma durumu ile ilgilidir. Bu ölçüt, konunun söylem içindeki öneminin bir yansıması ve konuşucunun konusal amacının ölçüsüdür. Givón, çalışmasında bu ölçütü sağa doğru olan, aradaki tümce sayısı olarak açıklamıştır. Bu ölçüt için belirlenen en küçük değer sıfırdır. Yani konunun söylemin sonraki bölümlerinde hemen yok olmasıdır. Bu ölçüt için en büyük değer belirlenmemiştir (Givón, 1983).

Givón'ın (1983)'deki çalışmasında 'Çok İmgelilik İlkesi'nden bahseder: *Bir konunun sürekliliği ne kadar az ve takip edilmesi ne kadar zor ise, dilde onu kodlayabilmek için o kadar fazla öge kullanılır.*

Psikolojik açıdan bakıldığında artgönderimsel kodlama ile konusallık arasındaki ilişki dil kullanıcısının bilişsel durumu ile ilgili bilgi de vermektedir. Givón'a göre göndergesel uzaklık ve araya giren gönderge sayısı arttıkça, hem o göndergeye gönderimde bulunmak için daha fazla öge kullanılmakta hem de dil kullanıcısı kastedilen göndergeyi belirleyebilmek için daha çok zaman ve zihinsel çaba harcamaktadır. (Givón, 1985).

2.4.2.2 Aşamalı Model

Aşamalı modelde, artgönderim seçimlerinin söylemin aşamalı yapısıyla belirlendiği varsayılmaktadır. Bu varsayıma göre, söylem içinde yeni yapısal birimlere ilişkin göndergeler tam AÖ ile kodlanırlar. Tam AÖ ile açıklanan bu göndergeye aynı söylem yapısı içinde sonradan gelen önermelerde yapılan gönderimler indirgenmiş artgönderimsel ifadeyle gerçekleştirilmektedir. Söylem içindeki yapısal birimler, dönüşler, paragraflar, bölümler, olaylar ve konular şeklinde belirlenebilir (Huang, 2000). Bu şekilde ifade edilen bu yaklaşımı, Hinds (1978; 1979), Tai (1978), Longacre (1979) ve Givón (1983) gibi birçok araştırmacı uygulamaya çalışmıştır. Hinds (1978) Japonca anlatı türünde artgönderimsel kodlamayı açıklamaya çalışmıştır. Bir artgönderimsel ifadenin bir diğerine tercih edilmesinin büyük ölçüde söylemin aşamalı

düzeni ile belirlendiğini öne sürmüştür. Tai (1978) de, Çince anlatı türünde kullanılan artgönderimsel örüntüler üzerinde yaptığı çalışmada benzer bulgular elde etmiştir. Aynı biçimde Longacre (1979) Nepal’de konuşulan Gurung dilinde ve Yeni Gine’de konuşulan Sanio-Hiowe dilinde adılların paragraf sınırları dışında kullanılmadığını belirlemiştir. Ancak bu konudaki en kapsamlı çalışma Fox’ın (1987) İngilizce yazılı ve sözlü metinlerde söylem artgönderimini incelediği çalışmadır (Huang, 2000).

Fox’un çözümlemesindeki temel sav artgönderimsel kodlama ile söylemin düzenlenmesinin çok yakından ilişkili olduğudur. Fox çalışmasında iki önemli tanım geliştirmiştir: “*kullanımı belirleyen bağlam*” ve “*bağlamı tamamlayan kullanım*”. İlkinde belli bir artgönderim seçimini söylemin aşamalı yapısının belirlediği varsayılmaktadır: İkincisinde ise belli bir artgönderimsel biçimin kullanılmasının belli bir söylem örüntüsü oluşturduğu varsayılmaktadır. Başka bir deyişle, Fox’a göre artgönderim dağılımı hem söylemin aşamalı yapısına göre belirlenir, hem de kendisi bu yapıyı belirler (Fox, 1987).

Fox (1987), açıklayıcı anlatım biçimli İngilizce yazılı metinlerdeki artgönderim çözümlemesini yaparken, araştırma yöntemi olarak Retorik Yapı Çözümlemesi (Mann, 1988) yöntemini benimsemiştir. Bu çözümleme yöntemine göre metinler aşamalı bir şekilde düzenlenmiş önermeler grubundan oluşmaktadır. Her önermenin içsel olarak bir Retorik yapısı (R-yapısı) vardır. Daha geniş bir bakış açısıyla bakıldığında Retorik Yapı Çözümlemesi (RYÇ) yöntemi, söylem çözümlemesi içinde yazılı ya da sözlü söylemin kurucu parçaları arasında retorik açıdan ya da anlambilimsel açıdan belli ilişkiler varsa, o söylem bağdaştırıcı görüşünü benimsemiş daha genel bir bağdaşıklık modeli olarak da görülebilir (Huang, 2000). Fox, aynı söylem alanı içinde sisteme dahil edilmiş öncüllerin adıyla ya da boş artgönderimlerle kodlandığını vurgulamıştır. Öte yandan başka bir söylem alanındaki öncüller ise Tam Ad Öbeği ile kodlanırlar. [Fox, 1987] söylemin aşamalı olarak sıralanmış önermeler dizisi olduğunu belirtir. Bu aşamalı yapının artgönderim seçimlerinde önemli bir rol oynadığını vurgular. Aşamalı modeldeki temel yaklaşım şu şekildedir: Söylem sıralı bir şekilde okunuyor olsa da anlamlandırmaları aşamalı olarak gerçekleşir. Kuşku yok ki bunun artgönderim davranışı üzerinde çok büyük bir etkisi vardır.

2.4.2.3 Ulaşılabilirlik Kuramı

Ariel (1988; 1990) tarafından geliştirilen bu kuram, yukarıda temel savları aktarılan Konu Sürekliliği Modeli ile birçok açıdan benzerlikler göstermektedir. Bu kurama göre, boş adılardan tam AÖ'lerine kadar tüm adsıl biçimler verilen bağlam içindeki göndergenin *ulaşılabilirliğini* yansıtan bir ölçek üzerinde sınırlanırlar. Ulaşılabilirlik, alımlayıcının gönderge ile zihinsel temas kurması için harcaması gereken çabanın miktarının belirlenmesi olarak da tanımlanabilir (Van Hoek, 1997).

Ariel'e (1988; 1990) göre artgönderimsel ifadelerin kullanımı ile temsil ettikleri zihinsel varlıklar arasında karşılıklı bir ilişki vardır. Olası tüm artgönderimsel ifadeler birer ulaşılabilirlik belirticisidir. Bu belirticilerden her biri aynı zamanda, konuşucunun zihninde tasarladığı göndergenin seslenen kişinin zihninde de aynı şekilde temsil edilebilirliğini belirleyen ulaşılabilirlik dereceleridir (Huang, 2000).

Ulaşılabilirliğin derecesini etkileyen temel etkenler nelerdir? Ariel, bir öncülün ulaşılabilirlik derecesini belirlemede payı olan dört temel etken belirlemiştir (1990).

- (i) Uzaklık: Artgönderim ile öncülü arasındaki uzaklık.
- (ii) Rekabet: Öncül ile artgönderim arasına giren olası öncüllerin sayısı.
- (iii) Belirginlik: Varlığın/göndergenin/öncülün belirginliği.
- (iv) Bütünlük: Artgönderimin öncülü ile aynı çerçeve / dünya / bakış açısı / bölüm / paragraf içinde olup olmaması durumu.

İlk iki etken, Givón'un bir önceki bölümde sözü edilen gönderimsel uzaklık ve gönderimsel müdahaleler ölçütleri ile benzerlikler göstermektedir. Üçüncü etken olan Belirginlik ise konu olan ve konu olmayan öncüller arasındaki farkla ilgilidir. Ariel'e (1990) göre, bir varlık ya konudur ya da değildir ve bir söylem birimi içinde yalnızca bir konu bulunabilir.

2.4.2.4 Merkezleme Kuramı

Adıl çözümleme kuramlarından biri olan, [Grosz ve Sidner, 1986] ve [Brennan vd., 1987] tarafından geliştirilen Merkezleme Kuramı, tümceler arası ve tümce içi göndermelerini incelemek amacıyla önerilmiştir. Grosz ve Sidner'ın (1986) çalışmasında yansıttığı, konuşucu ve dinleyicinin bir metin ilerledikçe dikkatlerini odakladıkları ve metin göndermelerini biriktirdikleri yapı olan dikkat odağı, dinamik bir yapıdır; bir gönderge sözcelerde kullanıldığı sürece dikkat odağında kalır. Eğer bir kaç sözcüde kullanılmazsa dikkat odağını terk ederek, yerini yeni göndermelere bırakır.

Merkezleme Kuramına göre, söylemin merkezindeki değişiklikler belli bir davranış ilkesi uyarınca oluşmaktadır. Söylemin, merkezindeki öğeyi mümkün oldukça koruma eğilimi içinde olduğu vurgulanmıştır. Bilgi tabanlı bu kuram, birçok uygulamada başarılı olmuş ve birçok çalışmaya ilham vermiştir. Söylem tutarlılığıyla ilgili olan, yerel bağdaşıklığı ele alan ve merkezin hareketine ilişkin bazı kurallar içeren Merkezleme Kuramı, yerel bağdaşıklık ve belirginlik modeli olarak sunulmuştur.

Bu kurama göre her söylem bir dizi bölüttten oluşur ve her bölüm bir tümceden (sözce) oluşur. Her sözcüde konuşucu ve dinleyicinin odaklandığı tek bir merkez ya da *geriye dönük merkez* (M_g) vardır. M_g kavramı Reinhart'ın (1981) konu kavramına benzer ve sözcenin en çok ilgili olduğu öge olarak tanımlanır. Ayrıca her sözce daha sonraki adılar için öncül olması olası bir dizi söylem varlığına yani ad öbeklerine sahiptir. Bunlar dikkat odağındaki belirgin olma derecelerine göre sıralanmışlardır ve ileriye dönük merkez (M_i) listesini oluşturur. Bu ileriye dönük merkez listesi söylem boyunca sürekli güncellenir ve sıralama değişir. Bu liste yeni bir listeye değişecektir. Dolayısıyla M_i listesi başka bir liste olacaktır. Eski M_i listesi ise bu aşamadan sonra geriye dönük M_g listesi olacaktır. M_i listesindeki en belirgin öge ya da ilk adılın olası öncülü olabilecek söylem varlığı o listenin odağı olacak ve M_o olarak ifade edilecektir. M_o , yani dikkat odağındaki en belirgin metin göndergesi genellikle tümcenin öznesidir. Bir sözcüde bir önceki sözcenin M_o 'su varsa o, bu sözcenin M_g 'sidir. Eğer bir sözcüde bir tek adıl varsa, o M_g 'dir. Bir sözcüde M_g adıl olduğu sürece bir başka adıl da bulunabilir. Yapılan araştırmalar vurgulanmayan ya da ses olarak ifade edilmeyen adılın

en belirgin ögeye gönderge yapmak için kullanıldığını göstermektedir. Bu, adıl düşmeli dillerde boş adıl, diğer dillerde ise vurgusuz adıldır (Turan, 1998).

Merkezleme Kuramı dikkat odağındaki değişimleri ve dönüşümleri de incelemektedir. Söylemdeki merkezde meydana gelen değişiklikler **devam**, **içerme**, **yumuşak-dönüş** ve **sert-dönüş** olmak üzere dört şekilde ifade edilmektedir. [Brennan vd., 1987] bu geçişleri 2.2 nolu tablodaki gibi açıklamıştır.

Tablo 2.2 Merkezdeki Değişiklikler

	$Mg(T_n) = Mg(T_{n-1})$	$Mg(T_n)! = Mg(T_{n-1})$
$Mg(T_n) = Mo(T_n)$	Devam	Yumuşak Dönüş
$Mg(T_n)! = Mo(T_n)$	İçerme	Sert Dönüş

Şimdi Merkezleme Kuramındaki iki bitişik sözcük arasındaki bu geçişlere bir göz atalım:

1. Devam geçişi: Metnin aynı metin varlığı üzerinden sürdürülmesi devam geçişi olarak adlandırılır.

- (a) George Orwell, yazarlığıyla olduğu kadar, kişiliğiyle de dikkati çeker. (2.46)
- (b) Çok genç yaşta, önce döneminin ve ülkesinin toplumsal düzenine karşı çıktı.

7. İçerme geçişi: Metin aynı metin varlığı üzerinden sürdürülürken, yeni bir metin varlığının metne katılması içerme geçişi olarak adlandırılır.

- (a) Sertab Erener Eurovision şarkı yarışmasında büyük başarı kazandı. (2.47)
- (b) Tüm dünya onu ayakta alkışladı.

8. Dönüş geçişi: Söylemde dikkat odağının değiştirilmesi dönüş geçişi olarak adlandırılır. İki tür dönüş geçişi bulunmaktadır: Yumuşak dönüş ve sert dönüş.

- (i) **Yumuşak dönüş;** eğer izleyen sözcüde bir önceki sözcenin en belirgin ögesi yerine, ileriye dönük merkez listesindeki daha az belirgin bir öge merkez oluşturursa bu tür geçiş yumuşak dönüş olarak adlandırılmaktadır.

(a) Selanik Zirvesi'nde AB liderlerine sunulan Avrupa Anayasası taslak (2.48) metni, AB ülkeleri dillerinin yanı sıra Türkçe olarak da basılıp dağıtıldı.

(b) Türkçe, Türkiye'den önce AB'ye girdi.

(ii) *Sert dönüş*; birbirini izleyen sözceler arasında merkez içerme ilişkisi yoksa bu durum sert dönüş olarak adlandırılır.

(a) Bronzlaşmak şu günlerde çok moda. (2.49)

(b) Güneş enerjisinin yararları dün yapılan toplantıda konuşuldu.

Bu merkezleme geçişleri söylem varlıklarının belirginlik derecelerine göre sıralanır. Merkezdeki her türlü değişiklik bu geçişlerle ifade edilmiştir. Merkezleme Kuramına göre söylem bu geçişlerden bazılarını daha çok tercih eder. Mesela devam geçişi ilk tercihtir. Devam geçişi iki önceki ve bir önceki tümcede merkezde yer alan ögenin yine merkezde olması durumudur. Yani merkezin değişmemesi söylem yapısının ilk tercihidir. Bu anlamda kurama göre söylemler merkezlerinin fazla değişmemesini tercih eder. Dolayısıyla bu geçişleri belli bir öncelikle sıralamak gerekiyor.

Devam > İçerme > Yumuşak-dönüş > Sert-dönüş

Söylemdeki adılların gönderimsel ilişkileri çözümlenirken bu sıralamadan yararlanılır. Tümcedeki adıllar aday öncüllerle birçok biçimde eşleşebilir. Bu eşleşmeler farklı senaryolar oluşturur. Bu farklı senaryolar ise farklı geçiş durumlarına denk gelir. Bu aşamada hangi geçişin seçileceği işte yukarıdaki sıralamaya göre seçilir. Yani *Devam* geçişi *İçerme* geçişine tercih edilir. Benzer bir şekilde *İçerme* *Yumuşak Dönüşe*, *Yumuşak Dönüş* *Sert-dönüşe* tercih edilir. Bu tercih yapıldıktan sonra hangi adılın hangi öncüle gönderimde bulunduğu ortaya çıkar. Bu da Merkezleme Algoritmasının çözümü olmuş olur. Merkezlemenin, söylem tutarlılığı için güçlü bir araç olduğu kanıtlanmıştır ve artgönderim çözümlenmesinde de başarılı bir şekilde kullanılmıştır. Basit bir örnekle, bu kurama açıklık getirilebilir. Aşağıdaki üç tümceli söylem ele alındığında yöntem şu şekilde uygulanabilir.

T_1 : Ayşe Ali'yi yolda gördü. (2.50)

T_2 : Ona elindeki kitapları gösterdi.

T_3 : Ve onları eve götürmesini rica etti.

Söylem 2.50 T_1 tümcesi için oluşturulacak ileriye dönük liste, $M_i(T_1)$, *Ayşe* (yalın hal), *Ali* (belirtme hali) ve *yol* (bulunma hali) söylem varlıklarını içerir. Bu söylem varlıklarından *yol* ögesinin gönderimsel ifade elemesinden geçemediği varsayalım. Merkezleme Kuramı yaklaşımında en önemli aşama gönderimsel ifadelerin ya da aday öncüllerin nasıl sıralanacağı aşamasıdır. Bu sıralama, Türkçe için Yıldırım (2008) tarafından basit olan ad öbeklerinin isim halleri üzerinden yapılmış istatistiksel bir sonuçla aşağıdaki şekilde bulunmuştur.

Yalın Hal > Belirtme Hali > İlgi Hali > Yönelme Hali > Ayrılma Hali > (2.51)
Bulunma Hali

Bu sıralama da yalın hale sahip söylem varlıkları daha belirgin ya da daha önemli oluyor. Sıralamada sağa doğru gidildikçe önem değerleri azalmaktadır. Bu sıralamaya göre, 2.50'nin T_1 tümcesi için oluşturulan M_i listesi sıralandığında

Ayşe > Ali (2.52)

gibi bir sıralam oluşur.

Söylem 2.50 T_2 tümcesi için oluşturulacak ileriye dönük liste, $M_i(T_2)$ ise, \emptyset (boş artgönderim, üçüncü tekil kişi), *ona* (yönelme hali) ve *kitap* (çoğul, belirtme hali) söylem varlıklarını içerir. Bu söylem varlıklarının elemekten başarılı bir şekilde geçtiği varsayalım. Bu liste 2.51'e göre sıralandığında ise;

\emptyset > Ona > Kitapları (2.53)

gibi bir sıralama oluşur. Bu listedeki adılların her biri için aday öncül belirlenmesi gerekiyor. Her iki adıl için {Ayşe, Ali} listesi olası öncül listesidir.

Öncül Listesi (\emptyset) : (Ali, Ayşe) (2.54)
 Öncül Listesi (Ona) : (Ali, Ayşe)

Bu aşamada 2.3 nolu tabloda gösterildiği gibi iki seçenekli senaryo oluşur.

Tablo 2.3 Senaryolar

	\emptyset	Ona
Senaryo 1	Ali	Ayşe
Senaryo 2	Ayşe	Ali

Bu aşamada hangi senaryonun Merkezleme Kuramı açısından en iyi geçişi verdiği bakılmalıdır. Birinci senaryo seçildiğinde 2.2 nolu tabloda belirtilen geçişlerden DEVAM geçişi seçilmiş olacaktır. Eğer ikinci senaryo seçilirse bu durumda geçişin ismi YUMUŞAK-GEÇİŞ olacaktır. DEVAM geçişi YUMUŞAK-GEÇİŞ'e tercih edildiğinden bu aşamada birinci senaryo tercih edilecektir. Sonuç olarak \emptyset boş artgönderimi Ayşe'ye, *ona* açık adlı *Ali*'ye gönderimde bulunacaktır.

Söylem 2.50 T₃ tümcesi için bu algoritma tekrar çalıştırıldığında bu tümcedeki onları artgönderimi Ali ve Ayşe öğeleriyle say ve kişi uyumu sebebiyle eşleşemez. Ardından bir tek kitaplar söylem varlığı kalır ve bu öge onlar adının öncülü olarak seçilir.

Merkezleme Kuramı, yalnızca yerel metin bölümlerindeki göndermeleri inceleyecek bir araç sağlar. Bu kuramda metnin yapısının çok katmanlı olduğu savı benimsenmiştir. Diğer yandan, Givón (1983) ve Ariel'in (1990) adlı çözümlemesi kuramlarında metin göndermelerinin dikkat odağını terk etmeleri adıl ve öncül arasında tümcelerin sayısı ile belirlenir. Bu öneri metnin çok katmanlı olan yapısını göz ardı ederek, düz bir metin yapısı görüşünü varsayar. Bu ise gerçeği yansıtmamaktadır (Grosz, 1995; Turan, 1998).

Fox (1987) adlı çözümleme kurallarında metin yapısının çok katmanlı olduğu görüşünü benimser. Ancak, hem Givón ve Ariel'in hem de Fox'un (1987) adlı

çözümleme kuralları Merkezleme Kuramı kuralları ile karşılaştırıldığında yeterli kesinlik ve matematiksel sağlamlık sağlamadığı için tercih edilir nitelikte değildir. Diğer yandan, Merkezleme Kuramı kesin kuralları ile çıkarım yükünü en aza indirerek, adıl kullanımındaki çekirdek kuralları açıklamayı amaçlamaktadır (Esmer, 2003).

2.4.3 Türkçe için Yapılan Dilbilimsel Artgönderim Çalışmaları

Türkçe’de artgönderim ile ilgili farklı görüşler içeren pek çok çalışma vardır. Bu çalışmalar arasında şunları sayabiliriz: Türkçe’deki artgönderimlerin açık ve gizli gösterimleri üzerine (Enç, 1986; Erguvanlı-Taylan, 1986), Merkezleme Kuramı perspektifinden Türkçe’deki söylem artgönderimleri (Turan, 1995), Türkçede Aşamalı Modeli kullanan (Oktar ve Yağcıoğlu, 1997) çalışma, retorik yapı çözümlemesi yapan (Çeltek ve Oktar, 2004) ve Türkçedeki artgönderimsel ilişkilerin bazı kısıtlarla tam olarak ifade edilemeyeceğine dair bir çalışma (Kılıçaslan, 2004).

2.4.3.1 Enç (1986)

Enç (1986) Türkçe’nin özne konumundaki adılın silinebildiği durumları belirlemek amacıyla gerçekleştirdiği çalışmasında, Türkçede iki tür söylem artgönderimi olduğunu ileri sürmektedir. Enç’e göre kastedilen gönderge söylem yapısından anlaşılabilirdiği zaman, konuşucu o gönderge yerine ya adıl kullanır ya da AÖ’ni tamamen siler. Basit tümcelerde eylem, öznesi ile sayı ve kişi bakımından uyum gösterdiği için özne konumundaki adıl silinebilir. Ancak Enç’e göre bu silinmeyi kısıtlayan süreçler vardır. Aşağıdaki iki örneği inceleyelim:

Ben çarşıya gidiyorum. (2.55)

Ø çarşıya gidiyorum. (2.56)

Doğruluk derecesi açısından 2.55 ve 2.56 tümceleri aynı değere sahiptir. Ancak, Enç’e göre, 2.55 ve 2.56 arasındaki seçim rastlantısal değildir. Bir bağlam içinde konuşucunun 2.55’deki tümceyi mi yoksa 2.56’daki tümceyi mi kullanacağını tümce kuruluşunu yöneten kurallar belirler. 2.55 nolu örnekteki ilk tümcede özne açık adıyla ifade

edilmiştir. İkinci örnekte ise boş artgönderim kullanılmıştır. Enç ilk tümcedeki açık adılın söylemin konusunu belirlediğini söyler. İkinci tümce “Neden ceketini giydin?” gibi bir soruya cevap olarak verilmiştir. Bu durumda ceketin neden giyildiği söylemin konusudur. Bu konuya yorum yapmak için ikinci tümce kullanılır. Bu tümcede boşartgönderim kullanılmasının sebebi, söylemin konusunun özne olarak belirlenmesinden kaçınmaktır.

Enç, önce Konu/Yorum ilişkisinden yola çıkarak şu saptamayı yapar: iletişimin sağlıklı olabilmesi için konuşucu ve dinleyicinin söylemin konusu üzerinde uzlaşmış olmaları gerekmektedir. Bu uzlaşma ya konunun açıkça belirtilmesi ile ya da konuya ilişkin imaların yapılması ile sağlanır. Söylemin başında konu üzerinde uzlaşıldıktan sonra, üretilen diğer tümceler konu üzerine yapılan yorumları içermelidir. Eğer konuşucu konuyu değiştirmek isterse bunu dinleyiciye açıkça belli etmelidir; aksi takdirde ürettiği tümceler eski konunun yorumları olarak algılanabilir. Bazı dillerde konu değişimi belli yapılarla sağlanır. Türkçede öznesi adıl olan tümceler bu işlevi görür. Aşağıdaki iki tümcelik söylem parçası, konu değişimini örnelemektedir:

Ali_i yarın Ankara’ya gidiyor. (2.57)

O_i bugünlerde çok dalgın. (2.58)

Enç, Türkçede bulunan altı açık adılın söylem yapısı içinde konu değişimini sağlamak için kullanılabildiğini belirtir ve açık adılların kullanımının söylem içinde bir AÖ’nin göndergesi ile adılın göndergesinin farklı/karşıt olduklarını göstermektedir.

Arabayı Ahmet yıkamadı. Ben yıkadım. (2.59)

Bu havada kimse top oynamaz. Ben oynarım. (2.60)

Sonuç olarak Enç, söylem içinde açık adıl kullanılmasının konu değişimi ve karşıtlık olmak üzere iki işlevi olduğunu belirlemiştir. Enç’e göre açık adılların bu iki işlevi birbirleriyle çok yakından ilişkilidir. Açık adıl kullanılarak konu değişimi sağlanmakta; bir anlamda söylem içinde eski ve yeni bilgi arasında bir karşıtlık yaratılmaktadır (Çeltek, 2003).

2.4.3.2 Erguvanlı-Taylan (1986)

Türkçede açık ve boş adılların tümce boyutunda nasıl bir dağılım gösterdiklerini saptamayı hedeflediği çalışmasında Taylan (1986), Türkçede artgönderimsel ilişkilerin;

- (i) boş adıllar,
- (ii) açık adıllar,
- (iii) boş ya da açık adılların seçmeli kullanımı

olmak üzere üç şekilde ifade edilebileceğini ileri sürmektedir.

Taylan şu iki sorunun yanıtını bulmaya çalışmıştır:

1. Bir dil bir ad öbeğinin bir diğeriyle artgönderimsel ilişkisini kurabilmek için ne zaman bir adıla ihtiyaç duyar, aynı ilişkiyi belirtmek için ne zaman boş gösterim kullanır?
2. Bu iki tür artgönderimsel ifade birbirlerinin yerine kullanılabilirler mi?

Taylan, Türkçede bulunan altı kişi adılının eylem üzerinde uyum ekleri ile belirtildiğini, bundan dolayı bu bağımsız adılların kullanımlarının bazı durumlar dışında seçmeli olabildiğini belirtir. Taylan'a göre, adılların seçmeli olmadığı durumlar, tümcenin öznesinin vurgulama ve karşıtlık yaratma işlevleri olduğu durumlardır. Bu durumlarda adıl kullanımı zorunlu olmaktadır.

(a) Ben işe geciktim. (2.61)

(b) Ø İşe geciktim.

(a) Ben işe geciktim ama sen henüz gecikmedin. (2.62)

(b) *Ø işe geciktim ama henüz gecikmedin.

(a) Bu raporu kim yazdı? (2.63)

(b) i. Ben yazdım.

ii. *Ø yazdım.

2.61-a ve 2.61-b'de adıl kullanımı seçmeli olabilirken, 2.62-b ve 2.63-b-ii'de adıl kullanılmaması dilbilgisi-dışı bir durum yaratmaktadır. Çünkü 2.62-b'de birleşik iki tümce karşıtlık ilişkisi içindedir; 2.63-b'de ise tümce yeni bilgi sağlıyor olduğu için öznenin eylem üzerinde uyum ekleriyle belirtilmesi yeterli olmamaktadır. Dolayısıyla iletişimsel değeri uyum eklerinden daha yüksek olan adılın kullanılması gerekmektedir.

Taylan'a göre, tümce içinde bir AÖ ile eşgönderimsel ilişkinin yalnızca boş artgönderim ile sağlandığı belli durumlar vardır. Bu durumlarda adıl kullanımı uzak bir gönderimi belirtir. Eğer yapıda öncül bir AÖ varsa, iyelik yapısındaki iye AÖ silinebilir.

- (a) Erol_i Ø_i karısı için her şeyi yapar. (2.64)
- (b) Ø_i karısı için Erol_i her şeyi yapar.
- (c) Erol_i onun^{*i/j} karısı için her şeyi yapar.
- (d) Onun^{*i/j} karısı için Erol_i her şeyi yapar.

2.64-a ve 2.64-b tümcelerinde karısı ifadesinden hemen önce boş bir artgönderim vardır. Bu artgönderimin boş kullanılması zorunludur ki *Erol* ile eşgönderime sahip olsun. Boş artgönderim öncülden önce olsa da fark etmez. Bu durumda yine iki öge aynı eşgönderime sahiptir. Öte yandan 2.64-c ve 2.64-d nolu tümcelerdeki *karısı* ögesi açık adıl ile kullanılmıştır. Bu durumda söz konusu açık adıl *onun*, *Erol* ile aynı eşgönderimli değildir. *Erol* ile eşgönderim ilişkisi ancak boş artgönderim ile sağlanabilir. 2.64-c nolu tümcede açık adıl öncülünden önce yer almasına rağmen 2.64-d nolu tümceyle aynı davranışa sahiptir. Taylan, özne konumundaki AÖ'leri eşgönderimsel ilişkinin, öznenin tümce içindeki konumuna bakmaksızın boş artgönderimlerle sağlandığını; özne olmayan AÖ'leri ile eşgönderimsel ilişkinin de, eğer AÖ gönderimsel ifadeden önce geliyorsa, ya boş artgönderimler ile ya da adıllarla sağlanabildiğini saptamıştır. Taylan, bazı durumlarda ise, bir AÖ ile eşgönderimsel ilişkinin yalnızca adıllar ile sağlanabildiğini ve bu durumlarda boş artgönderim kullanılmasının tümceyi dilbilgisi-dışı yapacağını belirlemiştir.

- (a) Ben Erol'u_i iyice tanıyınca kadar onu_i başka türlü bir insan zannetmiştim. (2.65)
- (b) *Ben Erol'u iyice tanıyınca kadar Ø başka türlü bir insan zannetmiştim.

- (a) Ahmet'in_i işe aldığı kızlar onu_i severler. (2.66)
- (b) * Ahmet'in işe aldığı kızlar \emptyset severler.
- (c) Onu_{*i/j} Ahmet'in_i işe aldığı kızlar severler.

2.65 ve 2.66'da eşgönderimsel ilişki adıyla sağlanmaktadır. 2.65-b ve 2.66-b tümcelerindeki durumlarda boş artgönderim kullanılması tümcelerin dilbilgisi-dışı olmasına neden olmaktadır. Taylan, bu tümcelerdeki eylemlerin belirtme durumunda nesne almak zorunda olduklarını ve bunların kullanılmamasının tümcenin dilbilgisi-dışı olmasına neden olduğunu ileri sürmektedir.

Bu örnekler ışığında Taylan şu sonuca ulaşır: Türkçede uyum dizgesi ile artgönderim seçimleri arasında bir etkileşim vardır. Tümcenin öznesi eylem üzerinde uyum biçimbirimleriyle belirtildiği için, bağımsız adların kullanımı seçimsel olabilir. Ancak, özne konumundaki adın vurgulama ya da karşıtlık yaratma işlevi varsa, adıl kullanımı zorunlu olur. Türkçede nesne uyumu eylem üzerinde belirtilmediğinden eylemin nesnesi açık adıl kullanımı ile belirtilmelidir; aksi takdirde tümce dilbilgisi-dışı olur. Ancak öncüller birden fazla olduklarında, tümce içindeki pozisyonları önemli olur. Örnek olarak aşağıdaki 2.67 nolu tümcede boş artgönderimin öncülü belirsizdir.

Ahmet_i Erol_j'a $\emptyset_{i/j}$ karısını sordu. (2.67)

Ancak *karısını* ögesi değişik pozisyonlarında yer aldığı bu belirsizlik ortadan kalkar. *Karısını* ifadesi 2.67 nolu tümcedeki gibi yerleştiğinde, \emptyset boş artgönderim *Ahmet* öncülüne gönderimde bulunur.

*Ahmet_i $\emptyset_{i/*j}$ karısını Erol_j'a sordu.* (2.68)

2.68 nolu tümcedeki gibi bir diziliş söz konusu olduğunda boş artgönderim Erol'a gönderimde bulunur.

*Erol_j'a $\emptyset_{*i/j}$ karısını Ahmet_i sordu.* (2.69)

Çalışmadaki bir başka bulgu ise şudur: Ana tümce öznesi ile aynı eşgönderime sahip yan tümce öznesi boş adıl olmak zorundadır.

Erol_i bana [\emptyset_i / onun_{*i/j} toplantıya gelmeyeceğini] söyledi. (2.70)

2.70 nolu tümcede görüldüğü gibi eğer yan tümcenin öznesi açık adıl olursa gönderimi *Erol*, yani ana tümcenin öznesi, olmaz. Eğer bu özne boş artgönderim olursa *Erol'a* gönderimde bulunabilir. Dolayısı ile bu tip örneklerde adılların açık ve boş kullanılması seçime bağlı değildir.

Çalışmada altı çizilen bir diğer önemli başlık ise söylemin gönderim çözümüne etkisidir.

(a) *Erol_i* yemeğe gelecekmiydi? (2.71)

(b) *Nazan* bana [onun_i / \emptyset_i yedide geleceğini] söyledi.

Daha önce belirtildiği gibi boş artgönderim, yantümcenin öznesi ile ana tümcenin öznesi arasında eşgönderimsel ilişki olduğu zaman kullanılabilirken, 2.71'de boş artgönderim kullanımı, aynı tümce içinde değil, iki ayrı tümce arasında eşgönderimsel ilişki kurmak için kullanılmaktadır. 2.71 nolu söylemde onun / \emptyset gönderimleri *Erol'a* gönderimde bulunur. 2.71-b tümcesinde, eğer söylem dikkate alınmasaydı, boş artgönderim *Nazan* ile eşgönderimli olarak yorumlanacaktı. Ancak burada boş artgönderim söylem içinde daha önce söylenmiş olan bir önceki tümcenin öznesi ile eşgönderimlidir. Taylan, Türkçede boş artgönderim kullanımının söylem içinde özgür bir kullanıma sahip olduğunu, bunun nedeninin de artgönderimin göndergesinin yakın söyleminden çıkarılabilmesi olduğunu belirlemiştir. Son olarak, söylemin de kendi içinde belli kuralları ve kısıtlamaları olduğunu ve bunları belirlemek için ayrı çalışmalar yapmak gerektiğini belirtmiştir (Çeltek, 2003).

2.4.3.3 Özsoy (1992)

Artgönderime sözdizimsel olarak yaklaşan bir diğer çalışma da Özsoy (1992) tarafından yapılmıştır. Özsoy, Yönetim ve Bağlama Kuramı çerçevesinde Türkçede bir adıl ile aynı tümce içinde bulunan öncül olabilecek bir ad arasındaki eşgönderge örüntüsünü incelemiş ve aşağıdaki iki örnekten yola çıkarak şu sorunun yanıtını bulmaya çalışmıştır: Türkçe tümcelerdeki eşgönderge örüntüsü rastlantısal mı yoksa Türkçe içindeki yapısal bir kısıtlamadan mı kaynaklanmaktadır?

*Hande_i [onun_i defteri]-ne resim çizdi. (2.72)

*Hande_i [onun_i işten çıkarılması]-nı kınadı. (2.73)

Özsoy'a göre 2.72 ve 2.73'deki adılar kendi yönetim alanlarında serbesttirler; ancak bir üst tümcedeki olası öncül tarafından bağlanamamakta; dolayısıyla da Bağlamanın, bir adılısıl yerel alanı içinde özgür olmalıdır diyen B İlkesine karşı gelmektedir.

Alev_i Ayşe'nin_j Mehmet'e_k onun_i/*_j/k bugün okula gelmeyeceğini (2.74)
söylediğini duydu.

Özsoy'a göre, 2.74'de de görüldüğü gibi Türkçede bir adıl, Bağlamanın B İlkesinin de öngördüğü şekilde kendi yönetim ulamında serbest durumdadır. Ancak bir üst tümcedeki olası öncül tarafından bağlanması gerektiği durumlarda belli kısıtlamalar gözlenmektedir. Özsoy, bu sapmanın Türkçenin yapısı içindeki eşgönderge sınırlamasından kaynaklandığını belirlemiştir. Sonuç olarak, Özsoy'a göre, Türkçede ana tümcenin belirleyicisi olan özne ile yantümcenin öznesi veya ad öbeğinin belirleyicisi olan adılar arasında eşgönderge ilişkisi yoktur ve dolayısıyla öncül ile adıl arasındaki eşgönderge ilişkisinde yapısal bir sınırlama vardır (Çeltek, 2003).

2.4.3.4 Turan (1996)

Turan (1996) çalışmasında, Merkezleme Kuramı (Grosz vd., 1986; 1995) perspektifinden Türkçe'deki söylem düzeyi artgönderimsel ilişkilerin çözümlenmesine yönelik en geniş analizi sunmuştur. Bu çalışmada Merkezleme Kuramı bakış açısıyla üç önemli konu üzerinde durulmuştur. İlk olarak, gönderimsel olmayan ifadelerin de gönderimsel ifadeler kadar adılların öncülleri olabileceği gösterilerek, ileriye dönük merkezler listesine katkı yapan isim öbekleri oldukları ortaya konmuştur. Üzerinde durulan ikinci konu, ileriye dönük merkezler kümesindeki elemanların sıralanması mevzudur. Turan, İngilizceden farklı olarak, Türkçedeki sıralamanın dilbilgisel roller üzerinden değil de tematik roller üzerinden yapılması gerektiğini söylemiştir. Bu çalışmada ayrıca, açık ve gizli adıllar ile tam isim öbeklerinin söylemdeki görevleri üzerine bir tartışmaya gidilmiştir.

2.4.3.5 Oktar ve Yağcıoğlu (1997)

Oktar ve Yağcıoğlu (1997), [Fox, 1987]'un çalışmasını örnek olarak Türkçede Aşamalı Modeli kullanarak bir çalışma yapmışlardır. Türkçe açıklayıcı anlatım biçimli yazılı metinler üzerine yaptıkları bu çalışmada artgönderim dağılımını ve artgönderimsel örüntülemenin dilsel düzenleme üzerindeki etkisini saptamaya çalışmışlardır. Bu çalışmada, artgönderimi özgün işlevi açısından, yani söylemde söz konusu olan kişiye ilişkin bilgilerin sürekliliğini izlemeyi sağlayan bir öge olarak ele almışlar ve çözümlmelerini üçüncü-tekil-kişi insan göndermeleri ile sınırlandırmışlardır (1997).

Oktar ve Yağcıoğlu, retorik yapı çözümlemesinden elde ettikleri bulgulara dayanarak Türkçe açıklayıcı anlatım biçimli yazılı metinlerde artgönderimin temel örüntüsünü şöyle açıklamaktadırlar (1997): “Eğer etken ya da denetleyici önermede söz konusu kişiye gönderimde bulunmak için adıl kullanılır; aksi takdirde tam AÖ kullanılır. Böylece, adıl kullanarak yazar okuyucuya, hedeflenen gönderenin etken ya da denetleyici önermede bulunduğu işaret etmektedir. Buna karşın, tam AÖ kullanarak yazar okuyucuya, hedeflenen gönderenin bu birimlerin dışında olduğunu belirtmektedir.”

2.4.3.6 Çeltek ve Oktar (2004)

Çeltek ve Oktar ise yine [Fox, 1987]'un çalışmasındaki retorik yapı çözümlemesi yöntemini temel alarak metinlerin sadece bir tümceler dizisi olmadığı, aşamalı biçimde düzenlenmiş önermeler dizisi olduğuyla ilgilenmiştir (2004). Bu çalışmada 5000 önermenin adıl ve boş artgönderim örüntüleri çözümlenmiştir. Bu analizde sadece 3. kişi göndergeleri için 317 adet boş artgönderim ve 144 adet açık adıl olduğu tespit edilmiştir. 3. kişi göndergelerde boş artgönderimlerin daha ağırlıklı olduğu kesindir. Öte yandan çalışmadaki ilginç bir bulgu ise 1. ve 2. kişi ile cansız göndergelerine yöneliktir. Bu göndergeler için tercih edilen artgönderimsel biçimin açık adıl olduğu belirtilmektedir. Öte yandan araya giren önermelerde söz konusu kişiye ilişkin bir açıklama yapılmış ise gönderge geriye-atlama yapısında adıl ya da boş artgönderim ile kodlanmaktadır.

2.4.3.7 Kılıçaslan (2004)

Kılıçaslan (2004), Türkçe'deki artgönderimsel ilişkilerin, cümle sınırları içinde tanımı doğrusal öncelik veya baskınlık kısıtlarıyla tam olarak ifade edilemeyeceği varsayımına dayanan bir kanıt sunmaktadır. Bununla birlikte, bu tip kısıtların etkinlikten yoksun olduğu bir cümle alanı da bularak bir adım öteye gitmiştir. Kılıçaslan, Banfield'ı (1982) takip ederek, cümlenin 'çekirdek' ve 'perifer'i arasında bir ayrım yapmaktadır. Bu açıklamaya göre, cümle odağı uygun bir soruya cevap olarak kabul edildiğinde, perifer alan solda cümle odağı, sağda ise yüklem ile sınırlanmıştır. Kılıçaslan, bir artgönderim ve öncülünün cümle periferinde iken, normalde bağlı oldukları kısıtlardan kurtulduğunu belirtmektedir. Artık, cümle yapısal olarak söylemden bütünüyle yalıtılmış bir birim olarak algılanamayacağından, Türkçe'de artgönderim çözümlemesini zorlaştıran bir diğer etken de bu olmaktadır.

BÖLÜM 3

BİLGİSAYARLI ARTGÖNDERİM ÇALIŞMALARI

Bu bölümde, otomatik artgönderim çözümleme işleminin bilgisayar ortamında hangi süreçlerden geçerek yapıldığı anlatılmıştır. Bilgisayarlı artgönderim çalışmalarının da kullandığı bu süreçler ayrıntılı bir şekilde incelenmiş ve daha sonra Türkçe ve diğer diller için yapılmış bilgisayarlı artgönderim çalışmalarından bahsedilmiştir.

3.1 Bilgisayarlı Artgönderim Çözümlemesi

Artgönderimin öncülünün otomatik bir şekilde tespit edilmesi işlemine otomatik artgönderim çözümlemesi denir. Bu bölümde otomatik artgönderim çözümlemesi için yapılması gereken temel adımlardan bahsedilecektir. Bu adımlar bugüne kadar artgönderimle ilgili yapılan çalışmalarda uygulanan temel adımlardır. Birçok çalışma benzer şekilde artgönderim için bilgisayar ortamında bir çözüm oluşturmaya çalışmıştır. Bilgisayarlı artgönderim çalışmalarında uygulanan otomatik artgönderim çözümlemesi genel olarak 3 adımda gerçekleşir:

1. Adım: Artgönderimin tespit edilmesi
2. Adım: Olası aday öncüllerin belirlenmesi
3. Adım: Olası aday öncüller arasından en iyi öncülün seçilmesi

Bu adımlara sırasıyla göz atalım.

3.1.1 Artgönderimin Tespit Edilmesi

Çözümleme işleminin başlayabilmesi için, ilk olarak öncülü bulunacak artgönderim tespit edilir. Sistemlerin büyük çoğunluğu, adsız artgönderimi ele alır. Çünkü, öncülü yüklem öbekleri, tümcecikler, tümceler yada tümce dizisi olan artgönderimlerin çözümlenmesi daha karmaşık bir işittir.

Bu çalışmada, üçüncü kişiye ait kişi adları, iyelik adları ve dönüşlü adlar olan adısal artgönderimler çözümlenmeye çalışılacağı için, bir söylem içerisinde geçen bu artgönderimler, çözümleme işleminin bu safhasında saptanırlar.

3.1.2 Olası Aday Öncüllerin Belirlenmesi

Artgönderim saptandıktan sonra, artgönderimin öncülünü bulabilmek için olası adaylar tespit edilir. Artgönderim çözümleme algoritmalarının çoğunluğu, öncülü isim öbeği olan artgönderimlerle ilgilenir. Bunun sebebi yüklem öbekleri, cümleler ve paragrafları öncül olarak kabul eden artgönderimlerin çözümlenmesinin oldukça karmaşık olmasıdır. Söylem parçası içinde herhangi bir artgönderimin öncesinde yer alan isim öbeklerinin tamamı, o artgönderim için aday öncül olarak kabul edilir. Genelde ise, artgönderimin öncülünün aranacağı bir kapsama alanı belirlenir. Bu kapsama alanı çoğunlukla artgönderimin bulunduğu cümle ve hemen öncesindeki cümleler olur.

Çocuklar pirinç tarlasında çalışan Ahmet'in babasına yardım için yola
koyuluyorlar. (3.1)

Onlar Ahmet'in iyi klavuz olduğunu düşünüyorlar. (3.2)

Yukarıdaki söylem parçasında, tümce 3.2'de yer alan *onlar* adının olası aday öncülleri, bir önceki tümcede yani tümce 3.1'de aranır ve olası aday öncüller *çocuklar*, *pirinç tarlası*, *Ahmet'in babası* ve *yardım* olarak belirlenir.

3.1.3 Olası Aday Öncüller Arasından En İyi Öncülün Seçilmesi

Artgönderimler saptandıktan sonra, program olası aday öncüller içinden onların öncüllerini seçerek çözümlmeye çalışır. Çözümleme işleminde kullanılan farklı bilgi kaynakları vardır. Tümce 3.2'deki *onlar* adının olası aday öncülleri belirlendikten sonra bu bilgi kaynaklarının kullanılmasıyla, *onlar* adının en iyi öncülü *çocuklar* olarak doğru bir şekilde bulunur. Şimdi bu en iyi öncülün bulunmasında rol oynayan bu bilgi kaynaklarını kısaca açıklayalım:

Biçimbilgisel ve Sözlüksel Bilgi: Artgönderime yol açan öğeler ve bunların gönderimde bulunabileceği varlıklara uygun kısıtlar belirlenmelidir. Bunların bir kısmı leksikon adı verilen teknik bir sözlük içerisinde belirlenirken, bir kısmı da gerçek işleme aşamasına geçmeden önce yapılan ve sözdizimden ayrı tutulan bir biçimbilgisel analiz ile belirlenmektedir.

Sözdizim Bilgisi: İlgilenilen gönderimsel varlıkların (ad öbekleri) cümle içerisinde konumları ve rolleri belirlenmelidir. Bunların belirlenmesi için ayrıştırıcı adı verilen yazılımlar kullanılabileceği gibi daha yüzeysel işlem yapan yazılımlarda kullanılabilir. Ayrıştırma aşamasının ortaya koyduğu kısıtlar ile birçok artgönderim referansı çözümlenebilmektedir.

Anlambilimsel Bilgisi: Biçimbilgisel, sözlüksel ve sözdizim bilgisinin koyduğu kısıtların yetersiz kaldığı durumlarda kullanılmak üzere anlamın getirdiği kısıtlamalar belirlenmelidir. Bunun için kullanılabilecek çeşitli anlambilimsel modeller olduğu gibi anlambilimsel bilginin bir kısmı teknik sözlük içerisinde de kodlanabilmektedir. Şu örneği inceleyelim.

Kedi çıktığı ağaçtan inmek istemiyordu. (3.3)

∅ Aşağıdakilere yalvarır gibi bakıyordu. (3.4)

Tümce 3.4'de, tümce 3.3'deki bir nesneye gönderimde bulunulmakta ve bu gönderim için iki aday (*kedi* ve *ağaç*) bulunmaktadır. Burada anlamsal kısıtlar çözümleme

sürecine katılır ve tümce 3.4'de gizli özne konumunda bulunan adılın ancak kediye gönderimde bulunabileceği, ağaçların bakamayacağı kısıtı kullanılarak bulunur.

Çözümleme işleminde kullanılan yukarıdaki bilgi kaynaklarına dayanan çözümleme kuralları, artgönderim çözümleme faktörleri olarak adlandırılırlar. Çözümleme işleminde kullanılan bu faktörler: cinsiyet ve sayı bakımından uyumluluk, sözdizimsel Bağlama Kuramı kısıtlamaları, anlambilimsel kısıtlamalar, sözdizimsel paralellik, anlambilimsel paralellik, belirginlik ve yakınlıktır. Bu faktörler, bazı ad öbeklerini olası aday kümesinden eleyebilir. Faktörler, bazı adaylara diğerleri üstünde öncelik vermek için öncelikli de kılınabilir: belirginlik, paralellik veya yakınlık gibi. Bilgisayarlı dilbilim literatürü, bu faktörler için birçok terminoloji kullanır. Örneğin, Carbonell ve Brown (1988), bu faktörler için kısıtlamalar ve öncelikler/tercihler terimlerini kullanmışlardır.

3.1.3.1 Kısıtlamalar

Kısıtlamalar, artgönderim ile onun öncülü arasındaki ilişkide sağlanması gereken zorunlu durumlar olarak kabul edilirler. Şimdi bunların neler olduğundan bahsedelim.

Cinsiyet ve sayı bakımından uyumluluk: Bu kısıtlama ile artgönderim ve öncülünün cinsiyet ve sayı bakımından uyumlu olmasına bakılır. Türkçe için cinsiyet bakımından uyumluluktan bahsedilemez. Çünkü İngilizcede cinsiyet kodlaması yapılabilmesine karşın Türkçede bu yapılamaz. Bu bilginin sözdizim yoluyla kodlanmaması artgönderim çözümleyici sistem veya kişi açısından karmaşıklık getirmektedir. Şu tümceleri inceleyelim:

Adam_i kadına_j bir kitap_k verdi. (3.5)

'The man gave the woman a book.'

Ø_i Çocuğun onu_{i/j/k} görmesini istemiyordu. (3.6)

'He did not want the child to see him / her / it.'

İki tümceden oluşan bu söylem parçası İngilizce olarak kodlanmış olsaydı, ikinci tümcedeki boş adılın, ilk tümcede söylem modeli içine alınan hangi nesneye gönderimde bulunabileceği sadece teknik sözlük bilgisi yardımıyla çözülebilirdi.

Doğum gününde Ali, Ayşe'ye bir *kitap* hediye etti. Ayşe *onu* henüz okumamış. (3.7)

Tümce 3.7'e bakıldığında Türkçe için sayı uyumluluğunun sağlanmasında herhangi bir sorun yoktur.

Oya_i evden çıkarken [*Ayşe ve Hasan*]_{*j*} *onu_i* uğurladılar. (3.8)

\emptyset_j Kapıyı kapattıktan sonra ağlamaya başladılar. (3.9)

Tümce 3.8'de *o* adılının yalnız *Oya*'ya, tümce 3.9'daki gizli adılın ise yalnız *Ayşe ve Hasan*'a gönderimde bulunabilmesi durumu sayı uyumunun bir sonucudur (Güner, 2008).

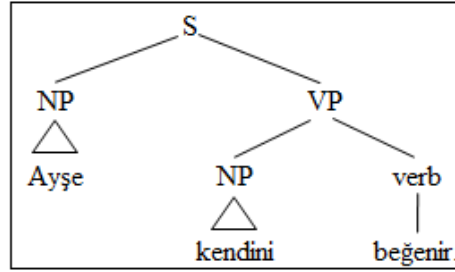
Sözdizimsel Bağlama Kuramı Kısıtlamaları: Bölüm 2.4.1'de detaylı bir şekilde sunulan Chomsky'nin Bağlama Kuramı (1981), artgönderimlerin nasıl gönderimde bulunacağına ilişkin önemli sözdizimsel kısıtları vurgular. Dönüşlü adılar, adılısılar ve gönderimsel ifadelerin yorumlanmasında Bağlama Kuramının sözdizimsel kısıtlamalarının rolü büyüktür ve her biri için farklılık göstermektedir. Şimdi bunları sırasıyla açıklayalım.

1. **Dönüşlü adıların yorumlanması;** dilbilgisi uyumluluğu, k-komutlandır ve yerel alan gibi bazı etkenlerle ilişkilendirilmiştir. Dönüşlü adıl ilk olarak, kişi ve sayı bakımından öncülüyle uyumlu olmak zorundadır. Bağlama Kuramı çerçevesinde tümce kavramı yerine kullanılan yerel alan kavramı, bağlanmanın sözdizimsel boyutunu belirleyen ve dönüşlü adıları direkt etkileyen bir koşul durumundadır. Diğer önemli bir koşul ise, dönüşlü adıların yorumlanmasını şu şekilde sınırlar: Dönüşlü bir adıl öncülü tarafından k-komutlandırılmış olmak zorundadır (Mitkov, 2002).

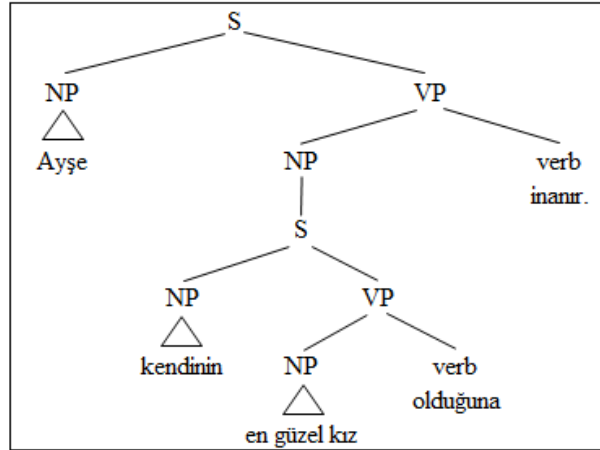
Ayşe_i kendini_i beğenir. (3.10)

Ayşe_i kendinin_i en güzel kız olduğuna inanır. (3.11)

3.10 ve 3.11’de kendi dönüşlü adılının öncülü, sayı ve kişi bakımından uyumlu, aynı yerel alanda bulunan ve kendi dönüşlü adılını her iki örnekte de k-komutlayan Ayşe’dir. Şekil 3.1 ve Şekil 3.2’de de görüldüğü gibi, Ayşe her iki sözdizimsel yapıda da kendi dönüşlü adılını k-komutlandırır. Bu sözdizimsel yapılarda geçen S(Sentence) tümce, NP(Noun Phrase) Ad Öbeği, VP(Verb Phrase) Eylem Öbeği ve verb de yüklem yerine kullanılmıştır. Bu sözdizimsel gösterimlerin Türkçe için detayları, bölüm 4.1.2’de detaylı bir şekilde açıklanmıştır.



Şekil 3.1 Tümce 3.10’un Sözdizimsel Yapısı



Şekil 3.2 Tümce 3.11’in Sözdizimsel Yapısı

2. *Adılların yorumlanması* ise; dönüşlü adılardan farklılık gösterir. 3.10 tümcesindeki *kendi* dönüşlü adılının bağlı olduğu ve *Ayşe*’ye gönderimde bulunduğu açıktır.

Ayşe_i onu_j beğenir. (3.12)

Tümce 3.12'deki *onu* adılısılı da 3.10'daki *kendi* dönüşlü adılı ile aynı sözdizimsel pozisyonda bulunmasına rağmen yerel alan da özgürdür ve bu alan dışına gönderimde bulunur. Adılısılar için uyulması gereken en önemli kısıtlama: Bir adıl, aynı yerel alan içindeki kendisini k-komutlandıran bir AÖ'e gönderimde bulunamaz. Bu kısıtlama, otomatik artgönderim çözümlemesinde sıkça kullanılır (Ingria ve Stallard, 1989).

3. **Gönderimsel ifade olarak tanımlanan sözlüksel ad öbekleri** ise; her yerde özgür olurlar. Yerel alanlarında veya yerel alanları dışında herhangi bir öncüle bağlı değildirler. Sözlüksel ad öbekleri için şöyle bir kısıtlama sözkonusudur: Sözlüksel ad öbeği, kendisini k-komutlayan bir ad öbeğine gönderimde bulunamaz (Ingria ve Stallard, 1989).

O Kerem'e hayrandır. (3.13)

O Kerem'in bir fotoğrafını çok sever. (3.14)

Kerem genç mankenin çok güzel bir kız olduğunu söyledi. (3.15)

Tümce 3.13 ve 3.14'de *o* adıları, *Kerem* sözlüksel ad öbeğini k-komutlandırır ancak eş dizimli değildirler. Tümce 3.15'de de *Kerem* ad öbeği, *genç mankeni* k-komutlandırır, ancak onlar da eş dizimli değildir.

Anlambilimsel kısıtlamalar : Bir öncülün uymak zorunda olduğu anlamsal bir model varsa buna gönderimde bulunan ifade de aynı anlamsal kısıtlara uymak zorundadır. Bu kısıtlama, artgönderime uygulandığı gibi öncüle de uygulanabilen bir kısıtlamadır. Örneğin,

Ali disketi bilgisayardan_i çıkardı ve onun_i iletişimi kesildi. (3.16)

Ali disketi_i bilgisayardan çıkardı ve sonra onu_i kopyaladı. (3.17)

Tümce 3.16'da *onun* iletişimi kopmuş bir *bilgisayar* iken, tümce 3.17'deki *onu* kopyalanan bir *diskettir*.

3.1.3.2. Öncelikler/Tercihler

Öncelikler/tercihler, kısıtlamalar gibi zorunlu durumlar değildir ve her zaman uygulanmazlar. Şimdi bunlara bir göz atalım.

Sözdizimsel paralellik: Sözdizimsel paralellik, diğer kısıtlamalar ve öncelikler/tercihler öncülü bulamadığı durumda yardımcı olabilir. Bu öncelik/tercih, artgönderimle aynı sözdizimsel işleve sahip olan isim öbeklerini verir.

Ahmet_i yeni arkadaşlarını_j kulübeye götürüyor. (3.18)

O_i orada onları_j ailesi ile tanıştıracak. (3.19)

Tümce 3.19’da özne konumundaki *o* adılının öncülü, sözdizimsel paralellik önceliği kullanıldığında, tümce 3.18’deki özne konumundaki *Ahmet* olarak bulunur. Yine tümce 3.19’da nesne konumundaki *onları* adılının öncülü de, sözdizimsel paralellik önceliğinin kullanılmasıyla, tümce 3.18’de nesne konumunda yer alan *arkadaşlar* olarak bulunur.

Anlamsal Paralellik: Anlamsal rolleri otomatik olarak tanıyabilen sistemler için önemli bir seçim unsuru olmaktadır. Artgönderim ile aynı anlamsal role sahip isim öbeğine öncelik tanınmasını sağlar.

Oya Ali’ye bir kitap vermişti.. (3.20)

Hasan da ona kalem vermişti. (3.21)

Yukarıdaki örnekte, *Ali* ve *ona* buldukları cümlelerde de olaydan etkilenen rolündeki öğelerdir, yani anlamsal rolleri aynıdır. Bu sebeple, *ona* adılı *Ali*’ye gönderimde bulunmaktadır (Güner, 2008).

Merkezleme: Sözdizimsel ve anlamsal ölçütler her ne kadar kuvvetli olsalar da muhtemel aday öncüller arasında tam bir ayırım yapamayabilirler. Böyle belirsizlikler oluştuğunda söylem içerisinde en dikkat çekici öğenin öncül olarak seçilmesi yoluna

gidilir. Bu en dikkat çekici öğeye hesaplamalı dilbiliminde, *odak* veya *merkez* denir (Grosz ve Sidner 1986; Grosz 1995). Aşağıdaki cümleyi inceleyelim.

Oya, bardağı tabağın içine koydu ve o kırıldı. (3.22)

Sadece bu cümleyi duyan bir insan veya bilgisayarın, *o* adının öncülünün *bardak* mı yoksa *tabak* mı olduğunu bulması imkansızdır. Fakat aşağıdaki gibi bir söylem parçası verildiğinde, artık artgönderimin çözümlenmesi için söylemin bağdaşıklığından faydalanılabilir.

Oya önceki gün markete alışverişe gitti ve orada bir bardak gördü. Onu almak istedi ancak yeterli parası yoktu. Dün tekrar markete gittiğinde bardağı alacak parası vardı. Bardağı alıp evine döndü. Mutfağa giden Oya, bardağı tabağın içine koydu ve o kırıldı...

Bu söylem parçası içinde geçen aynı artgönderim ilişkisi, bardak en dikkat çekici söylem varlığına dönüştüğü için rahatlıkla çözülebilmektedir.

İfadeler, Grosz ve Sidner'ın (1986) merkezleme kuramı kapsamında iki tip merkez ile ilişkilidir: ileriye ve geriye dönük merkezler. *İleriye dönük* merkezler, söylemin ileri kısımlarındaki göndergeler için potansiyel öncüllerden oluşa bir kümedir. Bunlar kendi içlerinde, önemlerine göre sıralanırlar. En önemli merkez, yani tercih edilen merkez, takip eden ifadeye *geriye dönük* merkez olacağı tahmin edilen varlıktır. Tek olan geriye dönük merkez, dikkat merkezindeki varlık olarak kabul edilir ve önceki en yakın ifadeye artgönderimsel bağlantıyı sağlar.

3.2 Bilgisayarlı Artgönderim Çalışmaları

Artgönderim çözümlenmesine yönelik bilgisayarlı dilbilim çalışmalarında geliştirilen algoritmalar, her ne kadar insanlar kadar başarılı olmasalar da, temel stratejileri açısından benzerdirler. Bu algoritmaların önemli bir kısmı dilbilim alanında ortaya konan bulguların bilgisayarlı ortamda gerçekleşmesine yöneliktir. Amaç, söylem içinde

olası öncülleri belirleyip, bunlar içinden en uygun olanını mümkün olan en yüksek başarıyla seçmektir. Algoritmalar uygulandıkları dil, hedefe ulaşmadaki yol, yöntemler ve kullandıkları özellikler bakımından farklılıklar göstermektedirler. Bu farklılıkları göz önüne aldığımızda son yıllarda kural tabanlı yaklaşımlardan öğrenme tabanlı yaklaşımlara doğru bir geçiş olduğu görülebilmektedir. Bunun yanında, teorik yaklaşımlardan, pratikte kullanıma sahip yaklaşımlara doğru da bir kayma vardır (Güner, 2008).

Bilgisayarlı dilbilim alanında yapılan artgönderim çözümleme çalışmaları kendi aralarında; biçimsel, sözdizimsel, anlamsal, sözlüksel ve söylem kaynaklı bilgiye dayanan *bilgi tabanlı yaklaşımlar*, makine öğrenmesi tekniklerini kullanan *öğrenme tabanlı yaklaşımlar* ve bilgi tabanlı yaklaşımlarla öğrenme tabanlı yaklaşımların birlikte kullanılmasıyla oluşan *melez yaklaşımlar* olarak üç temel sınıfa ayrılabilir.

Bu çalışmada, yukarıda verilen bu sınıflandırma yerine, bilgisayarlı artgönderim çalışmaları sözdizimsel yaklaşımlar ve sözdizimsel olmayan yaklaşımlar diye sınıflandırılıp incelenmiştir. Bu kapsamda çalışmamızın da temelini oluşturan Hobbs'un Naif Yaklaşımına ayrıca geniş bir şekilde yer verilip diğer bilgisayarlı artgönderim çalışmalarından ve Türkçe için de yapılmış bilgisayarlı artgönderim çalışmalarından kısaca bahsedilmiştir. Şimdi, artgönderim çözümlemesine ilişkin bilimsel yazında önemli bir yere sahip çalışmaların bazılarında yukarıda verdiğimiz sınıflandırma kapsamında kısaca bahsedilecektir. Daha sonra da Türkçe için yapılmış bilgisayarlı artgönderim çalışmalarından söz edilecektir.

3.2.1 Sözdizimsel Yaklaşımlar

Bu bölümde sözdizimsel tabanlı bazı yaklaşımlardan bahsedilecektir. Bu yaklaşımlar daha çok dilbilimsel teorilerden elde edilir ve öne sürülen model bu teorik arka plan üzerine oturur. Bu yaklaşımlardan en önemlisi bu çalışmanın da temelini oluşturan Hobbs'un Naif Yaklaşımıdır (Hobbs, 1976). Diğer sözdizimsel yaklaşımlar ise [Lappin ve Leass, 1994] tarafından öne sürülen RAP Algoritması ve RAP algoritmasının değiştirilmiş bir versiyonunu kullanan [Kennedy ve Borguraev, 1996] tarafından yapılan çalışmadır.

3.2.1.1 Hobbs'un Naif Yaklaşımı

Hobbs adil çözümlemesine dair iki yaklaşımda bulunmuştur: bir tanesi sözdizim tabanlı, diğeri ise anlambilim tabanlı yaklaşımlardır (Hobbs, 1976; 1977; 1978). Bu bölümde sözdizim tabanlı yaklaşımı üzerinde durulmuştur ve *Hobbs'un naif yaklaşımı* olarak bahsedilecektir ki bu yaklaşım, araştırma toplulukları içinde hatırı sayılır biçimde dikkat çekici özelliği ile hala en başarılı algoritmalarından birisi olarak geçer. Son karşılaştırmalar, modern çözümleme sistemlerinin büyük çoğunluğuyla başa baş olduğunu göstermektedir.

Hobbs'un algoritması, İngilizce için adil çözümlemesi yapabilen, sözdizim tabanlı bir algoritmadır. Bu algoritma, tümcelerın ayrıştırma ağacı üzerinde, adıldan başlayarak uygun olan öncülü buluncaya kadar, ağacın dalları arasında hareket etmeye dayalı bir prosedür içerir. Her bir gönderge için aday öncüllerin bulunması ve tespiti sözdizim ağacının metnin tersi yönüne doğru taranması ile yapılır. Tarama sonucunda uyum ve bağlama koşullarına uyan ilk aday gönderge ya da adil için olası aday öncül olarak önerilir.

Bu algoritma söz konusu adılın bulunduğu tümce ile önceki tümcelerın sözdizim ağaçlarını girdi olarak alıp bu ağaçlar üzerinde uygun bir öncül arar. Ağaç tarama algoritmasında belirlenmiş bir Söylem Modeli vardır. Bundan dolayı ağacı tarayacak algoritma bir dilbilgisi belirlemek zorundadır. Algoritma, bilgilerin yani tümcelerın, belli bir gramerin kuralları çerçevesinde bir ayrıştırma ağacı formatında gösterildiğini kabul eder. Çünkü sözdizimsel olarak ağacın yapısına ilişkin kabuller sonuçları etkileyecektir. Algoritmada, ağacın en uç noktasında bulunan adıldan başlayıp, yukarı yönde hareket edilir; S, NP veya VP noktalarına gelindiğinde durarak, bu noktaların alt kısmında kalan ağacın dalları arasında soldan sağa yönünde giderek uygun öncül olup olmadığını araştırır. Hobbs'un sözdizim ağaçlarını tarayan ve ilgili adılın olası aday öncüllerini bulan algoritması, aşağıdaki adımlardan oluşur.

1. S tümcesinin ayrıştırma ağacı üzerindeki, adılı içeren NP ad öbeği düğümünden başla.
2. Ağaçta ilk karşılaşılan NP veya S düğümüne çık. Bu düğümü X , bu düğüme ulaşmak için izlenen yolu da p diye adlandır.
3. X düğümünün altındaki p yolunun solunda kalan tüm dalları, soldan sağa yönde ve yayılma öncelikli arama biçimiyle dikkatle incele. Karşılaşılan herhangi bir NP düğümü ile X düğümü arasında bir NP veya S düğümü var ise, karşılaşılan ilk NP düğümü öncül olarak öner.
4. Eğer X düğümü cümlenin en üst S düğümü ise, en son gelen ilk olmak şartıyla metinde önceki cümlelerin ağaç yapılarını dikkatle incele; herbir ağacı soldan sağa yönde incele ve bir NP düğümüyle karşılaşıldığında, bu düğümü öncül olarak öner. Eğer X cümlenin en üst S düğümü değil ise, 5. adıma git.
5. X düğümünden, ağaç üzerinde yukarı yönde karşılaşılan ilk NP veya S düğümüne çık. Bu düğümü X , bu düğüme ulaşmak için izlenen yolu da p diye adlandır.
6. Eğer X düğümü bir NP düğümü ise ve eğer X 'e doğru olan p yolu X 'in direkt domine ettiği bir N-bar düğümü içinden geçmediyse, X 'i öncül olarak öner.
7. X düğümünün altındaki p yolunun solunda kalan tüm dalları, soldan sağa yönde dikkatle incele. Karşılaşılan herhangi bir NP düğümü öncül olarak öner.
8. Eğer X , S düğümü ise, X düğümünün bütün dalları p yolunun sağına doğru, soldan sağa yönde incele, ancak karşılaşılan herhangi bir NP veya S düğümü altına inme. Karşılaşılan NP düğümü öncül olarak öner.
9. 4. adıma git.

Hobbs'un algoritması ayrıştırma ağaç yapılarını, doğru cinsiyet ve sayıya sahip bir isim öbeği bulabilmek için, yukarıda ayrıntılı bir şekilde verilen belirli bir düzen içinde inceler. Algoritmanın 2. ve 3. adımları, dönüşlü adıl kullanılan yerlere dikkat eder. 5. ve 9. adımlar arasındaki döngü S ve NP düğümleri ile ilgilidir. 4. adım ise söylemdeki önceki tümcelere geçişi sağlar.

Hobbs, bir İngilizce parçanın yüzey yapılarını oluşturmak için aşağıda belirtilen bağlamdan bağımsız olan dilbilgisi kurallarını seçmiştir (parantezler seçimli anlamına, yıldız işareti de sıfır veya daha fazla sayıda oluş yani tekrar anlamına gelir):

S → NP VP
 NP → (Det) N-bar (PP/Rel)*
 NP → pronoun
 Det → article / NPs
 N-bar → noun (PP)*
 PP → preposition NP
 Rel → wh-word S
 VP → verb NP (PP)*

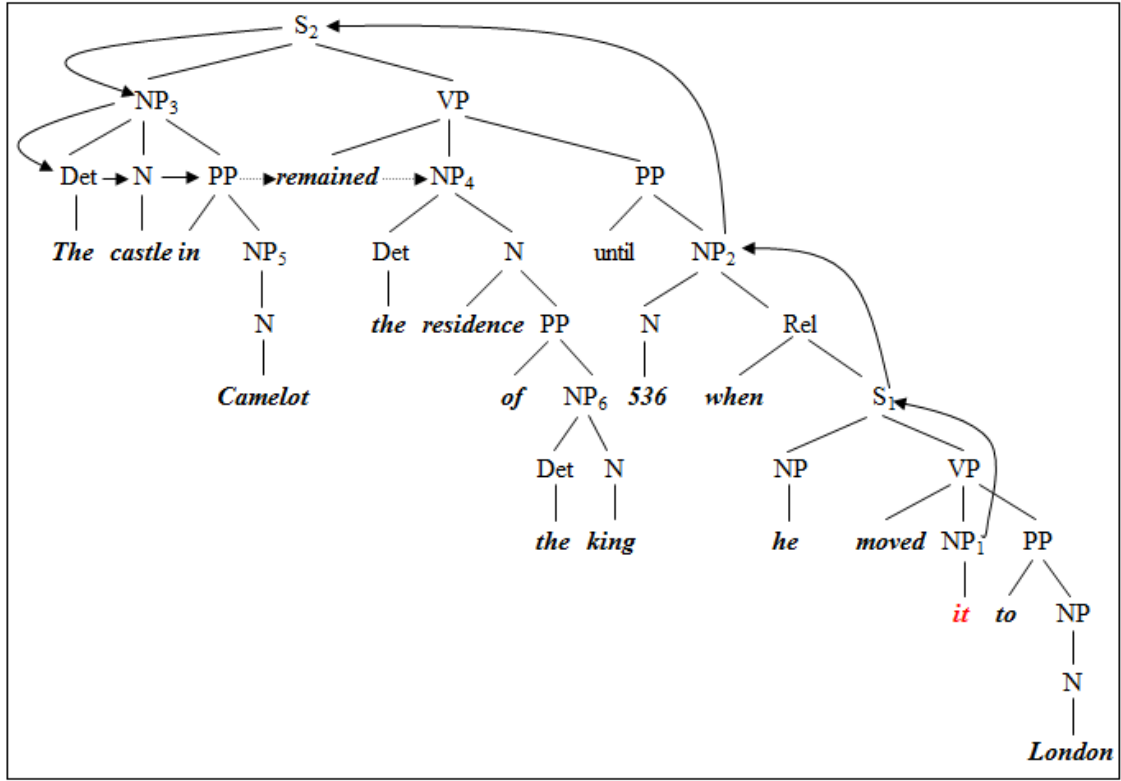
Aşağıdaki İngilizce örnekte ilk olarak *it* adının öncülleri algoritma yardımıyla bulunmaya çalışılmıştır. Daha sonra da *he* adının öncülleri benzer yolla bulunmuştur.

The castle in Camelot remained the residence of the king until 536 (3.24)
when he moved it to London.

Tümce 3.24'ün ağaç yapısının gösterildiği şekil 3.3'de NP₁ düğümü, algoritmanın 1.adımının başlangıç noktası olarak isimlendirilir. 2. adım bizi S₁ düğümüne götürür; bu düğüm X olarak adlandırılır ve X'e giden kesikli çizgili yola da *p* yolu denilir. 3. adım, X'in altında *p* yolunun sol tarafındaki dalları araştırır, ancak uygun bir NP düğümü bulamaz. 4. adım uygulanmaz. 5. adım ile NP₂'ye çıkılır. 6. adım, NP₂'yi öncül olarak önerir. Böylece, bu aşamada 536, öncül olarak önerilmiş olur.

'Tarihler taşınamaz, 'yerler taşınamaz' veya 'büyük nesnelere taşınamaz' gibi basit seçilebilir sınırlamalar, 536'yı öncül olarak yok saymaya yardımcı olurlar.

NP₂ elendikten sonra, 7. ve 8. adımlar bir şey getirmez ve kontrol 4. adıma geçer. 5. adım ile S₂'ye çıkılır, 6. adım uygulanmaz. 7. adımda, NP₃ (*the castle*) öncül olarak önerilir, ancak seçilebilir kısıtlamalara göre *the castle* elenir. Daha sonra arama NP₄'ün (*the residence*) doğru bir şekilde *it*'in öncülü olarak önerilmesiyle devam eder (Hobbs, 1976; 1977; 1978).



Şekil 3.3 Tümce 3.24'ün Ağaç Yapısı ve Algoritmanın Çalışmasının Gösterilimi

Eğer tümce 3.24'de *he*'nin öncülünü arıyorsak, algoritma ilk olarak NP₅'i (*Camelot*) cinsiyet uyumsuzluğundan dolayı eleyecektir ve sonunda NP₆ *the king* doğru bir şekilde bulunacaktır.

Hobbs, Langacker (1969) tarafından önerilmiş iki sözdizimsel sınırlamayı da kabul etmiştir. İlk sınırlama, dönüşlü olmayan bir adıl ve onun öncülü aynı cümle içinde bulunamayabilirler.

John likes him. (3.25)

John's portrait of him. (3.26)

John likes himself. (3.27)

3.25 ve 3.26 tümcelerindeki, *John* ve *him* eşgönderimli olamazlar. Tümce 3.27'deki *John* ve *himself* eşgönderimlidirler. Bu sınırlama, Hobbs'un algoritmasının 2. ve 3. adımları tarafından sağlanmıştır.

İkinci kural ise, Langacker (1969) tarafından önerilen, bir adılın öncülü, adıldan önce olmak zorundadır. NP₂ düğümünden önce yer alan bir NP₁ düğümü için denilebilir ki eğer ne NP₁ ne de NP₂ diğerini içermez ise ve NP₁'i çoğunlukla direkt olarak içeren S düğümü, NP₂'yi direkt içermez.

Hobbs, algoritmasında 3 farklı metinden aldığı 300 adılı (*he, she, it* ve *they*) değerlendirmiştir. Hobbs adılların ve öncüllerinin dağılımını incelemiştir. Bunun sonucunda, C₀, C₁, ..., C_N diye adlandırılan aday listeleri tanımlanmıştır.

C₀ = (a) Eğer adıl ana yüklemden önce geliyorsa, içinde bulunulan cümle ve bir önceki cümleye ait liste

(b) Eğer adıl ana yüklemden sonra geliyorsa, içinde bulunulan cümle ve bir önceki cümleye ait liste

C₁ = İçinde bulunulan cümle ve bir önceki cümleye ait liste

C_N = İçinde bulunulan cümle ve N cümle önceki cümleye ait liste

Hobbs, bütün öncüllerin %90'nını C₀ listesi içinde, %98'ini de C₁ listesi içinde bulmuştur. Algoritmanın orijinal halinde, Hobbs'un algoritması otomatik olmayan yolla test edilmiştir. Sonuç olarak, cümleler analiz edilmiş ve algoritma %88.3 durum için çalışmıştır. Seçilebilir kısıtlamaların kullanılmasıyla %91.7'lik bir başarı elde edilmiştir. Hobbs tarafından verilen %88.3 ve %91.7 başarı oranları ideal olarak kabul edilmiştir. Bir adıl çözümlene programı, yanlış sözdizimsel analiz veya sözcüksel izlemeden dolayı oluşan bazı hatalar içerebilir ve bundan dolayı da başarı oranı daha düşük çıkmış olabilir.

Hobbs'un yaklaşımı, bu alanda sözü en çok geçen çalışmalardan biridir ve sıkça olarak son çalışmaların (Baldwin 1997; Mitkov 1998a; Walker 1989) kıyaslama yapılabildiği klasik bir çalışma olarak kullanılmaktadır.

3.2.1.2 Lappin ve Leass'in Sözdizim Tabanlı Algoritması

[Lappin ve Leass, 1994] belirginlik temelli ve sözdizimsel kısıtlar ve kurallar içeren bir yöntem sunar. Bu yöntemde herhangi bir anlambilimsel veya edimsel kısıt bulunmaz. Sunulan yaklaşım sadece biçimbilimsel ve sözdizimsel kısıtlar uygular. Yöntemin temel yaklaşımı aday öncüllere biçimbilimsel ve sözdizimsel özelliklerine bakarak belli puanlar vermesidir. Her bir öncül topladığı puana göre sıralanır ve en çok puanı toplayan öncül verilen adılın ya da göndergenin öncülü olarak sisteme önerilir. Toplanan bu puan o öncülün söylem içindeki belirgenliğini ifade eder. Dolayısıyla bu teorinin sadece doğrusal bir yaklaşımı vardır.

Tüm söylem varlıklarına ya da gönderimsel ifadelerle bu belirginlik ağırlıklandırması uygulanır. Her söylem varlığının puanı hesaplanır. Ancak bu puan ele alınan adıla göre belirlenir. Her söylem varlığının adıldan bağımsız puanı olmaz. Ele alınan adıl ile olan uyumu en büyük etkendir. Örnek olarak adıl ile öncülün tümce yakınlığı önemli bir etkendir. Ancak yine de her ögenin ele alınan adıldan bağımsız özellikleri bulunur. Örneği ele alınan öncülün dilbilgisel görevi ağırlıklandırılmayı etkileyen güçlü bir bilgidir ve bu bilgi adıldan bağımsızdır.

RAP, Lappin ve Leass (1994) tarafından, geçişsiz üçüncü tekil kişi adıllarının çözümlenebilmesi için tanımlanmış bir algoritmadır. Bu algoritma, sözdizimsel önceliklere dayanan basit bir ağırlık tablosu kullanır. Şimdi hangi etken için ne gibi bir ağırlıklandırma yapılacağına bakılacak. [Lappin ve Leass, 1994] Tablo 3.1 içindeki ağırlık değerlerini kendi modellerinde uygulamıştır.

Tablo 3.1 Belirgenlik Türleri ve Ağırlıkları

Etkenin Türü	Ağırlık
Tümce Yakınlığı	100
Özne Olma Durumu	80
Varoluşsal Vurgusu	70
Belirtme Durumu	50
Dolaylı Nesnelere	40
Belirteç Öbeğinde Olmama Vurgusu	50
Baş Ad Öbeği	80
Rol Benzerliği	35
Öngönderim	-175

Gönderimi çözümlenecek adıl ile aday öncül arasındaki uzaklık en çok bir tümce kadarsa, öncülün belirginlik değerine 100 puan eklenir. Eğer öncül özne pozisyonundaysa 80 puan eklenir. Öncül varoluşsal pozisyonunda ise 70, belirtme durumuna sahipse 50, dolaylı nesne ise 40, baş ad öbeği ise 80 ve belirteç öbeğinde ise 50 puan eklenir. Bu ağırlıklandırma tablosu kullanılarak rastgele seçilen yaklaşık 360 adıl üzerinde algoritma sınıandı. Bu sınama sonucunda %86 oranında bir başarı kaydedildi.

Bu algoritmada, hiçbir anlambilimsel öncelik kullanılmaz. Genelde, algoritma tarafından iki tip işlem gerçekleştirilir:

1. Söylem Model Güncelleme
2. Adıl Çözümleme

İlk olarak, yeni bir ad öbeği (AÖ) geldiğinde, bu AÖ söylem modele aday öncül olarak eklenir ve belirginlik değeri hesaplanır. Belirginlik değeri, belirginlik faktörleri tarafından belirlenmiş ağırlıkların toplamına eşittir. En yüksek belirginlik değerine sahip AÖ, öncül olarak alınır. Ağırlıklar, yeni bir cümle işlendiğinde yarıya düşer. RAP Algoritması, şu adımlardan oluşur.

1. Aday öncülleri 4 cümle öncesine kadar biriktir.
2. Cinsiyet ve sayısal olarak uyuşmayan öncülleri çıkar.
3. Tümce-içi sözdizim kuralına uymayan öncüller de çıkar.
4. Aday öncüllerin belirginlik değerlerini, ağırlık tablosuna göre hesapla.
5. En yüksek belirginlik değerine sahip olan öncülü seç.

RAP Algoritması için aşağıdaki gibi bir örnek verebiliriz (Jurafsky ve Martin, 2000).

Nicola masanın üstüne bir kitap bıraktı. O onu Jane'den aldı. O onu okudu. (3.28)
Nicola kept a book on the table. She got it from Jane. She read it.

Söylem 3.28 insanlar için bile bulanık bir söylemdir. Üçüncü tümcedeki *o*, *Nicola*'nın mı yoksa *Jane*'inin mi artgönderimi? Tablo 3.2'de de görüldüğü üzere, ilk olarak, birinci tümcedeki aday öncüller alınır ve belirgenlik değerleri hesaplanır. Böylece Tablo 3.3'deki söylem modeli oluşmuş olur.

Tablo 3.2 Söylem 3.28'in İlk Tümcesindeki Aday Öncüllerin Belirgenlik Değerleri

Gönderge	Yakınlık	Özne	Varoluşsal	Nesne	Dolaylı Nesne	Belirteç Olmama	Baş AÖ	Toplam
Nicola	100	80				50	80	310
kitap	100			50		50	80	280
masa	100					50	80	230

Tablo 3.3 Söylem Modeli

Gönderge	AÖ	Değer
Nicola	{ Nicola }	155
Kitap	{ bir kitap }	140
Masa	{ masanın }	115

İkinci tümcedeki ilk AÖ, artgönderim *o*'dur. Algoritmanın 2. adımına uyularak, cinsiyet ve sayısal uyuşmaya bakılır. *Kitap* ve *masa*, çıkarılır. Kalan *Nicola*, *o*'nun öncülüdür.

Tablo 3.4 Söylem 3.28'in İkinci Tümcesindeki 'o' Adılıının Belirgenlik Değerleri

Gönderge	Yakınlık	Özne	Varoluşsal	Nesne	Dolaylı Nesne	Belirteç Olmama	Baş AÖ	Toplam
O	100	80				50	80	310

O ve *Nicola* aynı cümlede olmadıkları için, *o*'nun ağırlığı *Nicola*'nın ağırlığına eklenir.

Tablo 3.5 Söylem Modeli

Gönderge	AÖ	Değer
Nicola	{ Nicola, o ₁ }	465
Kitap	{ bir kitap }	140
Masa	{ masanın }	115

Diğer adıl *onu*, *bir kitap* ya da *masanın* artgönderimidir. İlk olarak *onu* için Tablo 3.6’da görülen belirgenlik değeri hesaplanır.

Tablo 3.6 Söylem 3.28’in İkinci Tümcesindeki ‘onu’ Adılıının Belirgenlik Değerleri

Gönderge	Yakınlık	Özne	Varoluşsal	Nesne	Dolaylı Nesne	Belirteç Olmama	Baş AÖ	Toplam
Onu	100			50		50	80	280

Dilbilgisel rol paralelliğine göre, *onu* ve *bir kitap* cümle içinde nesne konumundadır. Bu yüzden *bir kitap*ın belirgenlik değeri, +35 ile 175 olur. *Masanın* belirgenlik değeri 115 olarak kalır. Bu durumda, *bir kitap*, *onu* adılıının öncülüdür. *Onu* ve *bir kitap* ayrı cümlelerde olduğu için, *onu*’nun belirgenlik değeri *bir kitap*’ın belirgenlik değerine eklenerek, söylem model güncellenir.

Tablo 3.7 Söylem Model

Gönderge	AÖ	Değer
Nicola	{ Nicola, o ₁ }	465
Kitap	{ bir kitap, onu ₁ }	455
Masa	{ masanın }	115

Son olarak, ikinci cümlede yeni bir söylem göndergesi olarak ‘Jane’ gelir ve aşağıdaki Tablo 3.8’de görülmektedir.

Tablo 3.8 Söylem 3.28’in İkinci Tümcesindeki ‘Jane’ için Belirgenlik Değerleri

Gönderge	Yakınlık	Özne	Varoluşsal	Nesne	Dolaylı Nesne	Belirteç Olmama	Baş AÖ	Toplam
Jane	100				40	50	80	270

İkinci cümleden sonra söylem model şöyledir:

Tablo 3.9 Söylem Model

Gönderge	AÖ	Değer
Nicola	{ Nicola, o_1 }	465
Kitap	{ bir kitap, onu_1 }	455
Jane	{ Jane }	270
Masa	{ masanın }	115

Diğer cümleye geçerken, bu değerlerin yine yarısı alınarak, Tablo 3.10'daki yeni değerler oluşur.

Tablo 3.10 Söylem Model

Gönderge	AÖ	Değer
Nicola	{ Nicola, o_1 }	232.5
Kitap	{ bir kitap, onu_1 }	227.5
Jane	{ Jane }	135
Masa	{ masanın }	57.5

Son cümledeki *o* adılı, *Nicola* ya da *Jane*'in artgönderimi olabilir. *Nicola*'nın belirgenlik değeri daha yüksek olduğu için, *o* adılının öncülü *Nicola* olarak çözümlenir. *Onu*, *bir kitap* ya da *masanın* artgönderimi olabilir. *bir kitap*'ın belirgenlik değeri daha yüksek olduğu için, *onu* adılının öncülü de *bir kitap* olarak çözümlenir.

RAP algoritması, yaklaşık olarak 1.25 milyon kelime içeren bilgisayar elkitabında denenmiştir. Bu denemeden, üçüncü kişi adılı ve bunların öncülleriyle ilgili 360 tane oluşum seçilmiştir. RAP'ın 70 tümceler-arası artgönderim çözümlemedeki başarı oranı, % 86 bulunmuştur. RAP'ın 290 tümce-içi artgönderim çözümlemedeki başarı oranı, % 89 bulunmuştur. Lappin ve Leass'in bu çalışması, 1990'lı yıllarda artgönderim çözümleme için yapılan en etkili çalışmalardan birisidir. RAP, literatürdeki diğer yaklaşımların gelişmesinde sözdizim öncelikli bir temel oluşturmuştur.

3.2.1.3 Kennedy ve Borguraev'in Yöntemi

[Kennedy ve Boguraev, 1996] RAP algoritmasının değiştirilmiş bir versiyonunu sunmuşlardır. Bu çalışmada kullandıkları sistem, ayrıştırma ağacına ihtiyaç duymaz. Bu yöntemde dilbilgisel görevler yerine, sadece sözdizimsel bilgiye göre verilen bir adılın aday öncülleri puanlanır ve sıralanır. Bu özelliği ile RAP algoritmasından daha avantajlı bir durumu vardır. Çünkü gerçek sistemlerde mümkün olduğu kadar az bilgi kullanarak tasarlanan modeller her zaman çok daha etkili ve pratik olurlar. Bilgiye bağımlılığın en aza indirilmesi son zamanlardaki çalışmalar için önem verilen bir noktadır. Biçimbirimsel analiz yapan LINGSOFT isimli modelden gelen bilgileri kullanan bu yaklaşım, RAP algoritmasında olduğu gibi belirginlik ağırlandırması içerir. Magazin dergilerinden, ürün tanıtımlarından ve gazete haberlerinden oluşturulan derlem üzerinde test edilen algoritma 306 adıl için uygulanmıştır ve %75 gibi bir başarı elde edilmiştir.

3.2.2 Sözdizimsel Olmayan Yaklaşımlar

Bu alt bölümde geçen sözdizimsel olmayan yaklaşımlar, bölüm 3.2'de yapılan bilgi tabanlı, öğrenme tabanlı ve melez yaklaşımlar sınıflandırması kapsamına göre açıklanmıştır.

3.2.2.1 Bilgi Tabanlı Yaklaşımlar

Bu alt bölümde bilgi tabanlı yaklaşımlardan, Baldwin'in CogNIAC (1997) programından ve Mitkov'un MARS (1998b; 2002) adını verdiği sisteminden kısaca bahsedilmiştir.

3.2.2.1.1 Baldwin'in CogNIAC Programı

[Baldwin, 1997] az bilgili ve bir takım kurallardan oluşan CogNIAC isimli bir adıl çözümleme programı sunmuştur. Sırasıyla adılar üzerinde uygulanan bu kurallar, önemleri ve artgönderim çözümüyle olan ilişkilerine göre sıralanır. Adıla ilişkin uygun bir kural sağlandığında süreç tanımlanır ve sonuca varılır. Dolayısıyla kuralların sırası

çok önemlidir. Bu yöntemde uygulanan kuralları ve bu kuralların ne gibi bir başarı sağladığını aşağıdaki şekilde sıralayalım:

1. **Söylemde teklik:** Eğer sadece bir aday öncül varsa onu adılın öncülü olarak seç. Bu kural 8 kere doğru sonuç vermiştir.
2. **Dönüştürülebilirlik:** Eğer adıl dönüştürülebilir ise tümcedeki en yakın adayı öncül olarak kabul et. Bu kural 16 kere doğru, 1 kere yanlış çıkmıştır.
3. **Önceki ve mevcut tümcedeki teklik:** Eğer bir önceki tümce ya da mevcut tümcede tek bir adıl varsa onu öncül olarak seç. Bu kural 114 kere doğru, 2 kere yanlış sonuç vermiştir.
4. **İyelik durumundaki adılar:** Eğer adıl iyelik halinde ise (my, your, his, her vb.) ve bir önceki cümlede iyelik halinde olan bir öncül varsa onu adılın öncülü olarak kaydet. 4 kere doğru, 1 kere yanlış sonuç vermiştir.
5. **Mevcut tümcede teklik:** Eğer mevcut tümcede sadece 1 aday öncül varsa, onu adılın öncülü olarak kaydet. Bu kural 21 kere doğru, 1 kere yanlış sonuç vermiştir.
6. **Tekil Özne / Adılsız Özne:** Eğer bir önceki tümcenin özne ad öbeği sadece bir aday öncül içeriyorsa ve mevcut tümcenin öznesi adıl ise, bir önceki tümcenin ad öbeğini öncül olarak kabul et. Bu kural 11 kere doğru çalışmış ve hiç yanlış sonuç vermemiştir.

Sırasıyla uygulanan bu kurallar dışına çıkan durumlar için daha sonra iki ek kural daha eklenmiştir. Bu çalışma The Wall Street Dergisinden elde edilen derlem üzerinde 200 adıl için uygulanmıştır ve %92 kesinlik ve %64 anma oranında başarı elde etmiştir.

3.2.2.1.2 Mitkov'un Az-bilgili Yaklaşımı

Mitkov, teknik el kitaplarında geçen adıları çözümleyebilen sağlam ve az bilgiye dayanan bir yaklaşım sunar (Mitkov, 1998b). Algoritma, metni parçalara ayırıp işaretleyen bir işaretleyici tarafından ön işlemden geçirilmiş metni girdi olarak alır. Bu yaklaşımda, girdi metninde rastlanan bir adıl olduğu zaman, içinde bulunulan ve iki önceki cümlelerdeki isim öbekleri aday öncül olarak seçip çıkartılır. Bu adayların bazıları cinsiyet ve sayı uyumu filtreleri tarafından elenir. Mitkov, tercihleri belirtmek

için öncül göstergesi terimini kullanır. Uyum filtrelerinin uygulanmasından sonra, türe özel öncül göstergeleri kalan adaylara uygulanır ve en yüksek skoru alan aday isim öbeği öncül olarak bulunur.

Algoritmanın teknik el kitapları üzerindeki denenmiş başarı oranı %89.7'dir. Mitkov'un artgönderim çözümlene sisteminin tamamen otomatik versiyonu MARS diye bilinir (2002). Hiç bir işaret koyulmamış serbest metinler sistem tarafından analiz edilir ve içerdiği artgönderimsel ilişkiler otomatik olarak belirlenerek belli bir başarı oranıyla çözülür. MARS02 kodlu ilk versiyonda toplam beş aşamada işlemler tamamlanır. Birinci aşamada değerlendirilecek işlenmemiş metin sözdizimsel olarak ayrıştırılır. Bu aşamada ayrıştırılmış her bir parçanın biçimbilimsel, sözdizimsel, dilbilgisel analizi yapılır. Ad öbekleri bu süreçte belirlenmiş olur. İkinci aşamada artgönderimsel adılar, gönderimsel olmayan öğeler makine öğrenmesi algoritmaları aracılığı ile ayıklanır. Üçüncü aşamada her bir adıl için, adılın bulunduğu tümce ile iki tümce öncesine kadar olan metin parçasındaki olası adaylar belirlenir. Bu aşamada mevcut adaylar belli kısıtlar çerçevesinde yarışır. Yarışta sözdizimsel kısıtlar uygulanır. Öncelikle her bir aday öncül adıl ile sayı, kişi ve cinsiyet uyumu içinde olmalıdır. Dördüncü aşamada RAP algoritmasında olduğu gibi her bir aday öncülün puanı hesaplanır. Bunun için etkenlere bağlı olarak adaylar değerlendirilir ve puanlanır. Beşinci aşamada ise en yüksek puana sahip aday adılın öncülü olur. MARS'ın en son versiyonunda bazı güncellemeler yapılmıştır. MARS06 olarak nitelendirilen yeni versiyonda daha katı ve kesin uyum yöntemleri uygulanır. MARS sistemin en önemli özelliği dünyadaki en iyi sistem olması ve bu sistemin metin özetleme, bilgi çıkarımı ve metin sınıflandırılmasında kullanılmasıdır.

MARS bilgisayar kullanım klavuzlarından oluşan 8 farklı döküman üzerinde çalıştırılmıştır. Bu derlem 247.000 kelimededen ve 2.263 adıldan oluşur. MARS'ın tam otomatik versiyonundaki deneyler sonunda ortaya çıkan başarı oranı %61.55 olarak ölçülmüştür. Bu oran doğru çözümlenmiş adıl sayısının sistemde bulunan adıl sayısına oranıdır. Öte yandan yarı otomatik versiyonundaki sonuçlar %85-90 arasında çıkmıştır. Bu yöntemin Bulgarca dili adaptasyonu %72.6, Japonca dili adaptasyonu %75.8 başarı elde etmiştir (Yıldırım, 2008).

3.2.2.2 Öğrenme Tabanlı Yaklaşımlar

Öğrenme tabanlı yaklaşımlar, bilgi tabanlı yaklaşımlara bir alternatif oluşturan makine öğrenmesi gibi yapay zeka alanının en önemli alt alanlarından biridir. Öğrenme tabanlı algoritmalar, daha önceki deneyim durumlarından bir model oluşturur. Oluşturulan modeller, dayandığı veriyi iyi bir şekilde ifade etmeye çalışır ve bu verinin içindeki özellikle gizli olan kuralları çıkarmaya çalışır. Kısacası, veri madenciliğinin bir parçası olan makine öğrenme algoritmaları insanın bulamayacağı kuralları çıkarmak için kullanılır. En önemlisi ele aldığı veriyi analiz ederek gelecekte oluşacak durumlara tahminde bulunmasıdır. Şimdi bu yaklaşımlardan bazılarını kısaca bir göz atalım.

3.2.2.2.1 McCarthy ve Lehnert (1995)

McCarthy ve Lehnert (1995), RESOLVE adını verdikleri sistemleriyle, MUC-5 ticari metinler içeren derlemi üzerinde makine öğrenmesi yöntemi kullanarak eşgönderim çözümlemesi yapmaya çalışmışlardır. McCarthy ve Lehnert, aynı metni elle girilmiş kurallarla da çözümlemeye çalışarak, sonuçlarını karar ağacı algoritması kullandıkları yöntemleriyle karşılaştırmışlardır. Kullandıkları 1230 örnek durumun %26'sı olumlu ve %74'ü olumsuzdur. 8 özellik kullandıkları bu yöntemde, karar ağacı budanmadığında %87.6 duyarlılık sonucu elde ederlerken, budama yaparak sonuçlarını %92.4'e çıkarmışlardır. Öğrenmede kullandıkları özelliklerin tamamı anlamsal özellikler olup, sözdizimsel özelliklerle ilgilenmemişlerdir (Güner, 2008).

3.2.2.2.2 Aone ve Bennet (1996)

Aone ve Bennet'in (1996) çalışmasında otomatik olarak eğitilebilen bir artgönderim çözümleme sistemi geliştirilmiştir. Japonca gazetelerden oluşturdukları bir derlemede, artgönderimsel ilişkilerin işaretlenmesi ve bunlardan oluşturulan eğitim verileri üzerinde bir makine öğrenmesi yönteminin uygulanması yoluyla çözümleme işlemleri yapılmıştır. C4.5 karar ağacı algoritması (Quinlan, 1993) kullanılmıştır. Karar ağaçlarının, artgönderim ve olası öncül çiftleri üzerinden eğitilmesi için 66 özellik kullanılmıştır. Bu özellikler, sözlüksel bilgiler, sözdizimsel bilgiler, anlamsal bilgiler ve

adıl ile öncül arasındaki mesafe bilgisinden oluşur. Bunların bir kısmı sadece artgönderim veya öncüle ait ayrık özellikler olmakla birlikte, bir kısmı ikisi arasındaki ilişkiyi belirten özelliklerdir. Çalışmada iki olumlu örnek yaratma metodu kullanılmıştır: Geçişli ve geçişsiz. Geçişli olumlu örnekler, her bir artgönderim ile dahil olduğu artgönderimsel zincir üzerindeki onlardan önce gelen tüm ad öbekleriyle oluşturulan örneklerden oluşur. Geçişsiz örnekler ise, her bir artgönderim ile dahil oldukları artgönderimsel zincir üzerindeki onlardan önce gelen ilk ad öbekleriyle oluşturulan örneklerden oluşur. Olumsuz örnekler ise benzer bir şekilde her bir artgönderimsel ifade ile ondan önce gelen ve artgönderimsel zincir üzerinde olmayan gönderimsel ifade eşleşmelerinden yaratılır. Bu şekildeki eşleşmeler eğitim ve sağlama kümesini oluşturur. Bu deney [Quinlan, 1993] C4.5 karar ağacı algoritmasını kullanmış ve çalışmanın sonunda öğrenme tabanlı sistemin bilgi tabanlı sisteme oranla daha başarılı olduğu vurgulanmıştır. Bu çalışmadaki model 1971 adıl kapasiteli eğitim kümesi ile eğitilmiş 1359 adıl kapasiteli eğitim kümesi ile sınanmıştır. Bu test sonucunda alınan en iyi anma sonucu %70.20, kesinlik sonucu %88.55 ve f-ölçütü ise %77.27 olarak ölçülmüştür. Üretilen karar ağaçları sonucunda %90'a varan başarılarla ulaşılmıştır.

3.2.2.2.3 Soon, Ng ve Lim (2001)

Soon vd. (2001) çalışmalarında, C4.5 karar ağacı algoritması kullanarak artgönderim çözümlemesi yapmaya çalışmışlardır. Sistemde, bir sözlüksel özellik (eşsesli), sekiz dilbilgisel özellik (cinsiyet, sayı, ek, ad öbeği türü vb.), iki anlamsal özellik ve bir pozisyon (ad öbekleri ve adılar arasındaki uzaklık bilgisi) özelliklerinden oluşan 12 adet özellik kullanılmıştır. Oluşturulan sistem, 13.000 kapasiteli bir derlem üzerinde gerçekleştirilen deneylerden çıkan sonuçlar %52 anma, %68 kesinlik ve %58.9 f-ölçütü şeklindedir.

3.2.2.3 Melez Yaklaşımlar

3.2.2.3.1 Mitkov, Evans ve Orasan (2002)

Mitkov vd. (2002) tarafından ortaya konulan MARS (Mitkov's Anaphora Resolution System) adlı bu sistem tam otomatik olarak çalışmaktadır. İşaretlenmemiş metinleri alıp ayrıştırma ve artgönderim çözümlemesi işlemlerini yapar. Öncelikle sözdizimsel ayrıştırma yaparak isim öbeklerini belirler. Bu artgönderimsel öğelerin her biri için iki cümle öncesine kadar olan aday öncülleri toplayarak, bunlar üzerinde kişi/sayı ve cinsiyet uyumu filtrelerini uygular. Ayıklanmış aday öncüller kendi aralarında çeşitli etkenler kullanılarak sıralanır ve bir artgönderimin gerçek öncülü belirlenmiş olur. Bu tam otomatik sistem, gerçek uygulamalarda kullanılmış olup, teknik kullanım kılavuzlarından oluşan bir derlemede %61.6 gibi bir başarıya ulaşmıştır.

3.2.2.3.2 Preiss (2002)

Preiss (2002) çalışmasında, Kennedy ve Boguraev'in (1996) çalışmalarında kullandığı algoritmayı, makine öğrenmesi yöntemiyle gerçeklemiştir. Bu şekilde bellek-tabanlı bir yaklaşımın etkinliğini ölçmüştür. Preiss, artgönderim çözümlemesini, bir ikili sınıflandırma problemi olarak ele aldığı yaklaşımıyla, en az özgün kural-tabanlı yaklaşım kadar iyi bir başarı elde etmiştir.

3.2.3 Türkçe için Bilgisayarlı Artgönderim Çözümleme Çalışmaları

Türkçe için yapılan çok sayıda dilbilimsel artgönderim çalışması olmasına karşın, bilgisayarlı artgönderim çözümleme çalışması çok fazla değildir. Bunun yanında Türkçe için yapılan bilgisayarlı artgönderim çalışmaları arasında bizim bu çalışmada yaptığımız gibi Hobbs'un Naif Yaklaşımına dayanan sözdizim tabanlı adıl çözümlemesi yapabilen bir çalışma da bulunmamaktadır. Şimdi bu bölümde bizim çalışmamız dışındaki yapılan diğer bilgisayarlı çalışmalara kısaca bir göz atalım.

3.2.3.1 Tın ve Akman (1994)

Türkçe için yapılan ilk bilgisayarlı çalışmalardan biri olan bu çalışmada Tın ve Akman (1994), Türkçede artgönderim çözümlemesine dair Durum Anlambilimi çerçevesinde Durum Kuramının¹ bilgi-odaklı özelliklerinin, adılsıl artgönderim çözümlemesinde kullanılması için bir sistem tanımlamışlardır. Bilgisayarlı bir çerçeve içinde çözüm sunan bu sistem BABY-SIT olarak adlandırılmıştır. BABY-SIT, artgönderimsel bir anlam belirsizliğini çözebilmek için geri ve ileri bağlama kısıtlarıyla² ortaya çıkardığı ilişkili bağlam bilgisini kullanmaktadır. Teorik çerçevede oluşturulan bu bilgisayarlı model deneysel bir sonuç sunmamıştır. Dolayısıyla bir başarı oranı belirtilmemiştir.

Durum teorisine göre bir söylem yapısını bireyler, onların özellikleri, birbirleri arasındaki ilişkileri, konuları ve durumları gibi basit bileşenler belirler. Bir söylem durumu, anlatılan ifadeyi, onun konuşmacısını, sözcenin (utterance) konumunu ve adresleri içerir. Bir ifadenin sözcüsü, söylem durumları, bağlantılar ve tanımlanmış durum için hangi rollerin yer aldığına bağlı olarak, kelimeyi belli bir yola zorlar. Bir bağlam içindeki ifadenin sözcüsünün yorumlanmasında, kısmen ifade içindeki kodlanmış dilbilim biçiminden ve kısmen de sözce durumu tarafından bulunmuş bağlamsal faktörlerden oluşan bir bilgi akışı vardır. İfadenin anlamı ve buradan onun yorumu, vurgu, şekil ve tonlama gibi faktörler tarafından etkilenmiştir. Ancak, söylenen ifade içindeki durum ve bu sözce tarafından tanımlanan durumun en etkili rolleri oynadığı görülmektedir (Tın ve Akman, 1994).

1 Durum Kuramının özgün hali için bkz. Barwise ve Perry (1983), Kuramda günümüze kadar yapılan değişiklik ve gelişmeler için bkz. Devlin (2004).

2 Durum Kuramında tanımlandığı şekliyle kısıtlar, bir tipteki bir durumun başka tipteki bir durum hakkında bilgi taşıyabildiği, durum tipleri arasındaki ilişkilidir.

3.2.3.2 Turhan-Yöndem ve Şehitoğlu (1997)

Turhan-Yöndem ve Şehitoğlu tümce içi boş artgönderimler için öbek-yapı dilbilgisi çerçevesinde bir çalışma sundular. Bu çalışmada boş artgönderimlerin çözümünün tümce bileşenlerinin sırasına bağlı olduğu belirtilmiştir. HPSG (Head-Driven Phrase Structure Grammar) dilbilgisi tabanlı bir ayrıştırıcı kullanılmıştır. Çalışmada vurgulanan bir başka nokta ise çözümlemenin anlambilimsel bağlantılara bağlı olduğudur. Konuşucular ve dinleyicilerin artgönderimlerin çözümü için anlambilimsel bağlantıları kullandıkları belirtilmiştir.

3.2.3.3 Yüksel ve Bozşahin (2002)

Bu çalışmada, Türkçe için içeriğe uygun adıl üretimi Yüksel ve Bozşahin tarafından yapılandırılmıştır. Türkçe hem görünür hem de boş adillara sahiptir. Bu çalışma Merkezleme Kuramı ve Bağlama Kuramı ilkelerini uyarlayarak oluşturulmuş ve sınıflandırılmış kurallarla yerel ve yerel olmayan gönderimleri modellemeye çalışmıştır. Yerel gönderimler, Bağlama Kuramı'nın temel kurallarına göre yapılandırılmıştır. Bağlama kuramına göre, dönüşlü adılların öncülleri yerel etki alanı; kişi adıllarının öncülleri de yerel etki alanı dışında olmalıdır (Mitkov, 2002). Sistemdeki yerel olmayan gönderimler ise Bağlama Kuramı ve Merkezleme Kuramı kurallarının etkileşimi tarafından planlanmıştır. Çalışmanın içerdiği bir diğer konu ise Türkçe'nin adıl düşürme yapısı ile ilgilidir. Bu başlık altında adılların nasıl düştüğü ya da nasıl boş artgönderime dönüştüğü tartışılmıştır.

Sistem, doğal bir dilin üretim dizgesinde kullanılabilir, aynı zamanda bir makine çeviri dizgesinin üretim aşamasında da yer alabilir. Sistemin başarı oranı %70 bulunmuştur.

3.2.3.4 Küçük ve Turhan-Yöndem (2007)

Dilek ve Turhan-Yöndem, üçüncü kişi adlarını ile dönüşlülük adlarını çözümleyen, Türkçe için az bilgili bir adıl çözümleme sistemi sunmuşlardır (2007). Sistem, diğer diller için geliştirilen sistemlerde olduğu gibi, mümkün olduğunca sözdizimsel ve anlamsal bilgi kullanmamayı tercih etmektedir. Sistemin çözümlendiği adıl türleri özel isimlere işaret eden üçüncü tekil ve çoğul (düşmüş ve düşmemiş) kişi adlarını ile dönüşlülük adlarıdır.

Önerilen sistemin temelinde az bilgi gerektiren kısıt ve tercihler bulunmaktadır. İngilizce, İspanyolca gibi diğer diller için geçerli olan kısıt ve tercihlerin yanı sıra Türkçe'ye özel olanları da ilk kez önerilmiş ve bu sistemde kullanılmıştır. Türkçe'ye özgü olan kısıt ve tercihlerin belirlenmesi amacıyla sistem gerçekleştirilmeden önce örnek bir Türkçe metin üzerinde analiz yapılmıştır. Diğer birçok az-bilgili adıl çözümleme sisteminde olduğu gibi bu sistemde de kısıtlar adılın işaret ettiği nesnelere elenmesinde, tercihler ise bunların sıralanmasında kullanılmaktadır. Her bir tercihin bir makine öğrenimi yolu olan yapay sinir ağı (neural network) yöntemi kullanılarak sistem tarafından öğrenmesi sağlanılmıştır. Sistemin kullandığı kısıtlar, sayı uyumu, dönüşlülük adlı ve kişi adlı kısıtlarıdır. Sistemde kullanılan tercihler ise, tırnaklı/tırnaksız metin tercihi, yakınlık tercihi, yalın hal tercihi, ilk isim tümceciği tercihi, ek-fiil tercihi, tekrar tercihi, noktalama tercihi ve düşmüş kişi adlı tercihidir.

Sistemin değerlendirilmesi aşamasında karşılaştırma yapabilmek için iki farklı algoritma daha gerçekleştirilmiştir. Bunlardan ilki sadece kısıtları kullanmakta, diğeri ise kısıtların kullanılmasının ardından adıla en yakın olan adayı adılın işaret ettiği nesne olarak vermektedir. Sistemin bu iki algoritmaya karşı iki farklı türde metin üzerinde deneme sonuçları, sistemin bu iki algoritmadan daha başarılı olduğunu göstermiştir (Küçük, 2005).

3.2.3.5 Yıldırım ve Kılıçaslan (2007)

[Yıldırım ve Kılıçaslan, 2007] tarafından yapılan çalışmada Türkçe artgönderim problemi çeşitli yaklaşımlar ışığında ele alınmıştır. Bilgi tabanlı tümdengelimli yöntemlerle, öğrenme tabanlı tümevarımlı yöntemler birbirleriyle çeşitli açılardan karşılaştırılmıştır. Belirginlik tabanlı yaklaşım ile Merkezleme Kuramı, bilgi tabanlı yöntemleri değerlendirmek için ele alınmıştır. Belirginlik tabanlı yöntem derlem üzerinde denenmiş ve %52 oranında bir başarı elde edilmiştir. Bir diğer bilgi tabanlı yaklaşım olan Merkezleme Algoritması söylemin merkezinin davranışına ilişkin çeşitli kurallar içerir. Merkezdeki ögenin kendini merkezde tutma eğilimi prensibi üzerine kurulmuş bu yöntem Türkçe derlem üzerinde sınanıldığında ise %71 gibi bir sonuç elde edilmiştir. Öte yandan adılısı ögelerin çok daha belirgin olması ihtimaliyle bu algoritma güncellenmiş ve %72 gibi bir başarı oranı elde edilmiştir. Bu noktada ise adılısı ögelerin belirginliğinin diğer ögelere göre çok daha fazla olduğu gözlemlenmiştir. Bilgi tabanlı yaklaşımın bu iki önemli algoritmasından merkezleme yönteminin belirginlik tabanlı yöntemle göre daha başarılı olduğu gözlemlenmiştir. Bu deneyden Türkçe metinlerin merkezleme kuramının vurguladığı kurallara yatkın olması sonucu çıkarılmış ve bu bulgu çeşitli örneklerle güçlendirilmiştir.

Makine öğrenme algoritmalarından sadece beş tanesi çeşitli ayarlarla denenmiştir: Karar Ağaçları, K-en Yakın Komşu Algoritması, Destek Vektör Makineleri, Naïve Bayes Algoritması, ve Algılayıcılar artgönderim problemine uygulanmıştır. Bu algoritmalarından ilk ikisinin ifade gücü arttıkça, belli bir noktadan sonra aşırı uyum sorunu ürettiği gözlemlenmiştir. Bununla birlikte ifade gücü azaldıkça yetersiz uyum sorunu yarattığı da gözlemlenmiştir. Diğer üç algoritma ise doğrusal ayarlarla kullanıldıklarında derlem üzerinde başarılı olamamıştır. Bu ayarlar doğrusal olmayacak şekilde düzenlendiğinde ise başarı oranları yüksek çıkmıştır. Bu açıdan üzerinde çalışılan Türkçe derlemin doğrusal sınıflandırıcılarla ifade edilemeyeceği ve sınıflandırılmayacağı sonucu çıkarılmıştır. Doğrusal ayara sahip bir algılayıcının performansının belirginlik tabanlı yaklaşımla aynı çıkmasının yöntemlerin yapısal benzerliklerinden kaynaklandığı ve bu yapısal benzerliğin Türkçe artgönderim problemi için yeterli bir çözüm yöntemi oluşturamadığı gösterilmiştir.

Makine öğrenmesindeki elde edilen başarılar genelde anma, kesinlik ve f-ölçütü metrikleriyle değerlendirilir. Ancak makine öğrenmesi modelinin ürettiği sonuçların artgönderim çözümlemesi için bir aşamadan daha geçirilmesi gerekmiştir. Çünkü makine öğrenme algoritmaları söyleme veya artgönderime özel olan kuralları bilemeyeceğinden bazı düzenlemeler yapılmıştır. Öğrenme sonrası süreç olarak tanımlanan bu aşamada öğrenme için kullanılan anma, kesinlik, f-ölçütü metriklerinden başarı oranı metriğine geçilmiştir. Bu geçiş sırasında “ifade gücünün çok olması aşırı uyuma yol açar” biçimindeki olumsuz etkinin yok olduğu gözlemlenmiştir. Muğlaklık durumunda başvurulan *en yakın adayın öncül olarak seçilmesi*, bu etkiyi azaltmıştır.

Bazı Türkçe metinlerdeki artgönderimsel ilişkiler bilgisayarlı bu modellerle az bilgi kullanılarak çözümlenmiş ve deney sonuçlarına dayanarak birçok bulgu sunulmuştur.

BÖLÜM 4

SİSTEM TASARIMI VE UYGULAMASI

Bu çalışmada, Türkçe'deki üçüncü kişi adılarını çözümleyen sözdizim tabanlı bir adıl çözümleme sistemi tasarlanmıştır. Sistem, Hobbs'un Naif Yaklaşımı (Hobbs, 1976, 1978) ve Chomsky'nin Bağlama Kuramı (Chomsky, 1981, 1995) ilkelerini uyarlayarak, Türkçe için oluşturulmuş kurallarla üçüncü kişi adılarının öncüllerini bulmayı amaçlamaktadır. Türkçe için bilgisayarlı artgönderim çözümlemesi yapabilen bu çalışma, şimdiye kadarki Türkçe için yapılmış çalışmalar içinde Hobbs'un yaklaşımına dayanan ve onun algoritmasına benzer sözdizimsel bir algoritma kullanan ilk çalışmadır.

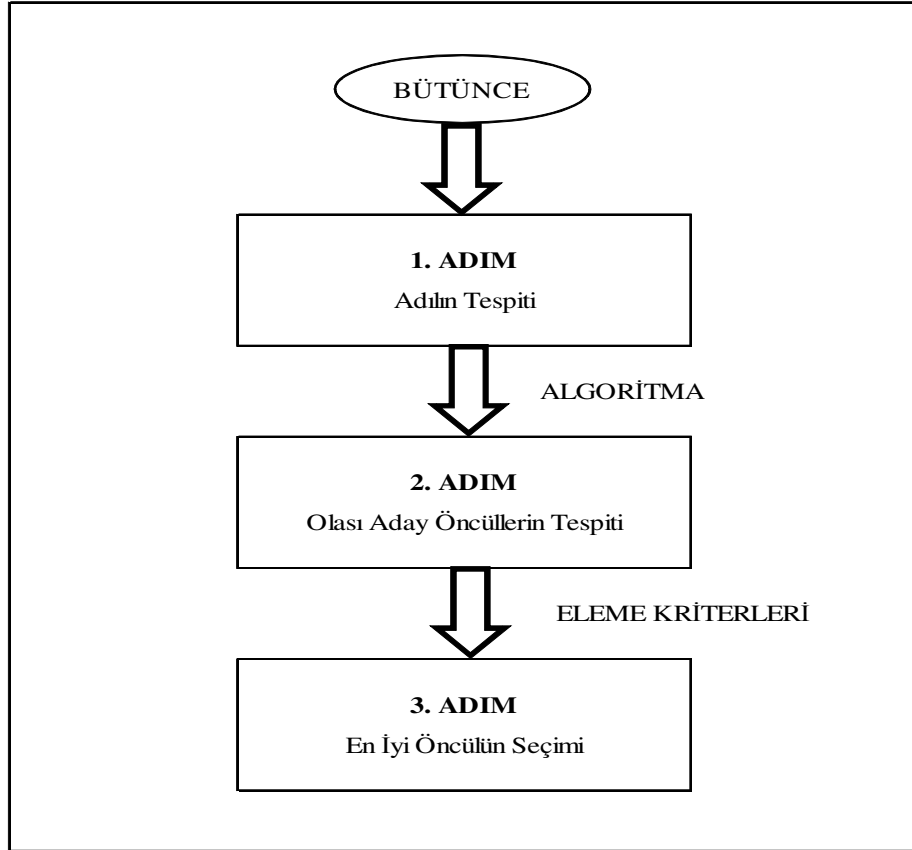
Hobbs'un algoritması bölüm 3.2.1.1'de de detaylı bir şekilde anlatıldığı gibi, İngilizce için adıl çözümlemesi yapabilen, sözdizim tabanlı bir algoritmadır. Bu algoritma, tümcelerın ayrıştırma ağacı üzerinde, adıldan başlayarak uygun olan öncüle kadar ağacın dalları arasında hareket etmeye dayalı bir prosedür içerir. Her bir adıl için aday öncüllerin bulunması ve tespiti; sözdizim ağacının metnin tersi yönüne doğru taranması ile yapılır. Tarama sonucunda, bölüm 2.4.1'de de detaylı bir şekilde incelenen Chomsky'nin Bağlama Kuramı'ndaki bağlama koşullarına uyan ad öbekleri, adıl için olası aday öncül olarak önerilir.

Chomsky'nin Bağlama Kuramı'na göre, adılısıllar, göndergeler ve gönderimsel ifadeler bağlı oldukları tümce alanına göre farklılıklar gösterir. Göndergeler tümce içinde yani yerel alanda bağlı olurlar; adılısıllar başka tümceciklerdeki ya da tümce dışındaki yani yerel alan dışındaki AÖ'lerine bağlıdırlar. Bağlama Kuramı, bağlamanın yapılabileceği ve yapılamayacağı alanları, bağlama koşullarıyla belirlemeye çalışır. Bağlama koşulları, bölüm 2.4.1'de detaylı bir şekilde açıklanmış olan k-buyurma ve

eşdizimli olma durumlarını içerir. Bu iki yaklaşım ve kuramın kurallarının Türkçe için uyarlanması sonucunda sistemimizin ana kısmını oluşturan algoritma elde edilmiştir.

4.1 Sistem Tanıtımı

Bölüm 3.1’de de bahsedildiği gibi, otomatik artgönderim çözümlemesine dair, bilgisayarlı artgönderim çalışmaları tarafından uygulanan üç temel adım vardır. Bu adımlar bugüne kadar artgönderimle ilgili yapılan çalışmalarda da uygulanan temel adımlardır. Genelde bir çok çalışma Şekil 4.1’deki benzer aşamalardan geçerek artgönderim için çözüm oluşturmaya çalışmıştır.



Şekil 4.1 Otomatik Artgönderim Çözümleme

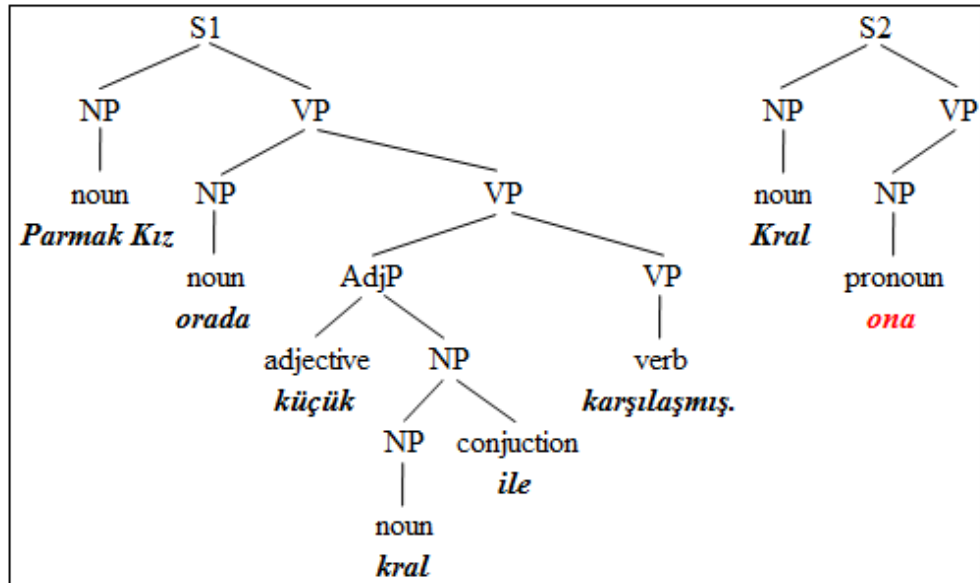
Şimdi bu adımların bizim çalışmamızda nasıl uygulandığına kısaca bir göz atalım.

1. Adım: İlk adımda, öncülü bulunacak adıl tespit edilir. Bunun için ilk olarak birkaç Türkçe metinden otomatik metin işleme yolu ile içinde üçüncü kişi adılı ‘o’ olan tümce ve bu tümcelerden bir önceki tümceler seçilerek alınır. Bu çalışmada, bir adılın öncülünü ararken, olası aday öncüllerin bulunması için seçilen alanlar adılın içinde bulunduğu tümce ve bir önceki tümce olarak seçilmiştir. Bu şekilde adıl çözümlemesi, her bir söylem için iki tümceyle sınırlandırılmıştır. Seçilen bu tümceler bütüncüye elle girilir. Daha sonra sistem bu tümceleri programda giriş verisi olarak kullanır ve program bu tümceleri ağaç yapıları şeklinde ayrıştırır. Ayrıştırma işlemi sırasında adıla rastlandığında, öncülü bulunacak adıl da tespit edilmiş olur. Program bütüncüden aldığı verileri Şekil 4.2’deki gibi ağaç yapılarına dönüştürür. Bu dönüştürme işlemi sırasında adıla rastlandığında, program ağaç yapısı oluşturma işlemini yarıda keser ve sistem bir sonraki adım için hazır hale gelir. İki tümceden oluşan söylem 4.1’i inceleyelim.

Parmak Kız orada küçük kral ile karşılaşmış. (4.1)

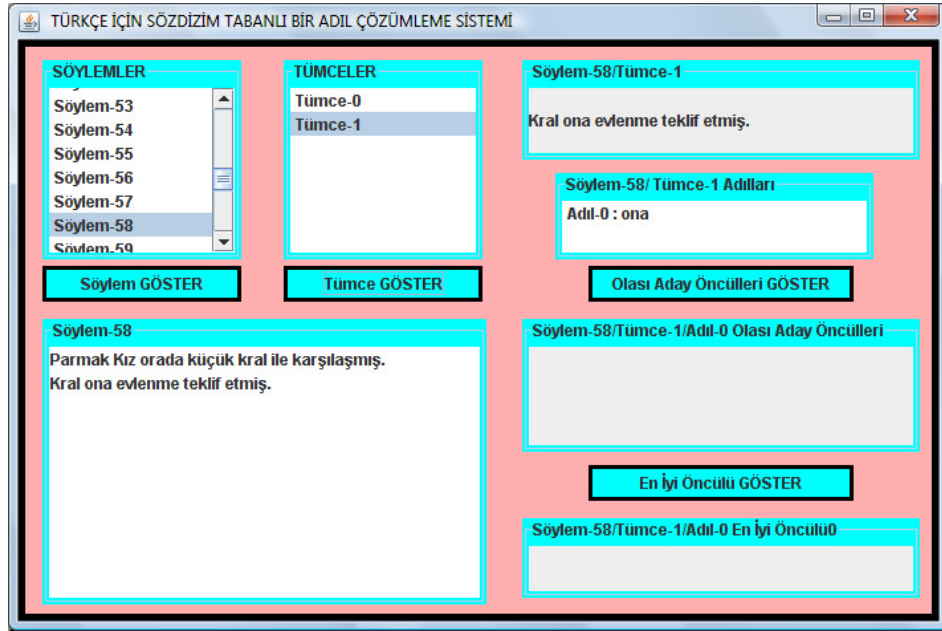
Kral ona evlenme teklif etmiş.

Aşağıda da gördüğümüz gibi ikinci tümcedeki ‘ona’ adılına kadar, tümcelerin ağaç yapıları oluşturulmuştur. Adıla rastlandığı andan itibaren ağaç yapısı oluşturma işlemi durmuş ve ikinci adıma geçilmiştir.



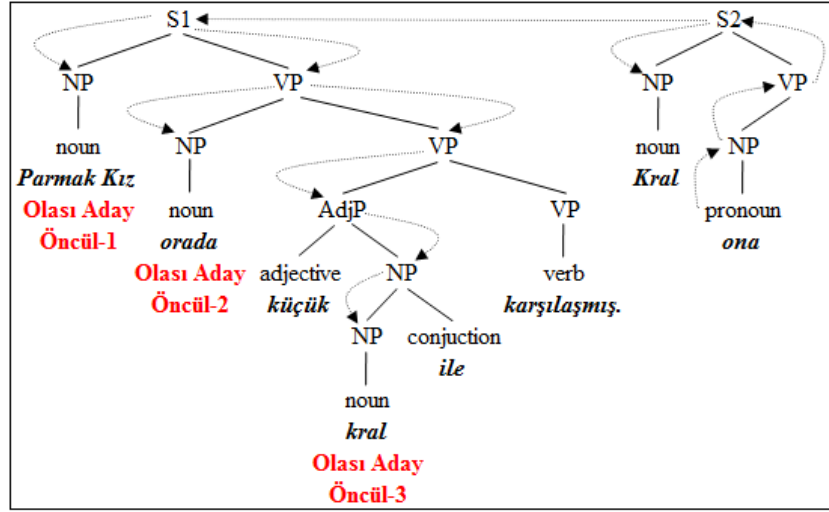
Şekil 4.2 Adılın Tespit Edilmesi

Bu işlemlerin program tarafından nasıl adım adım yapıldığını Şekil 4.3'deki programın ön yüzü sayesinde görsel olarak anlayabiliriz. Örnek 4.1, sistemde Söylem-58'e karşılık gelir. Ön yüzde Söylem-58'i seçip 'Söylem GÖSTER' düğmesini tıklarsak, bu söylemi tamamını 'Söylem-58' başlıklı kutucuğun içinde görebiliriz. 'Tümceler' başlıklı kutucuğun içinde ise tümceleri numaralandırılmış biçimde görürüz. Hangi tümce üzerinde işlem yapılmasını istiyorsak, o tümceyi seçip 'Tümce GÖSTER' düğmesine tıkladığımızda, sağ üst köşede 'Söylem-58/Tümce-1' başlıklı kutucuğun içinde seçilen tümceyi görürüz. Aynı anda bu tümcenin içindeki adılar da 'Söylem-58/Tümce-1 Adılları' başlıklı kutucuğun içinde sistem tarafından gösterilir. Buraya kadar ilk adımda adılın tespit edilmesi işlemi, program tarafından bu şekilde gerçekleştirilir.



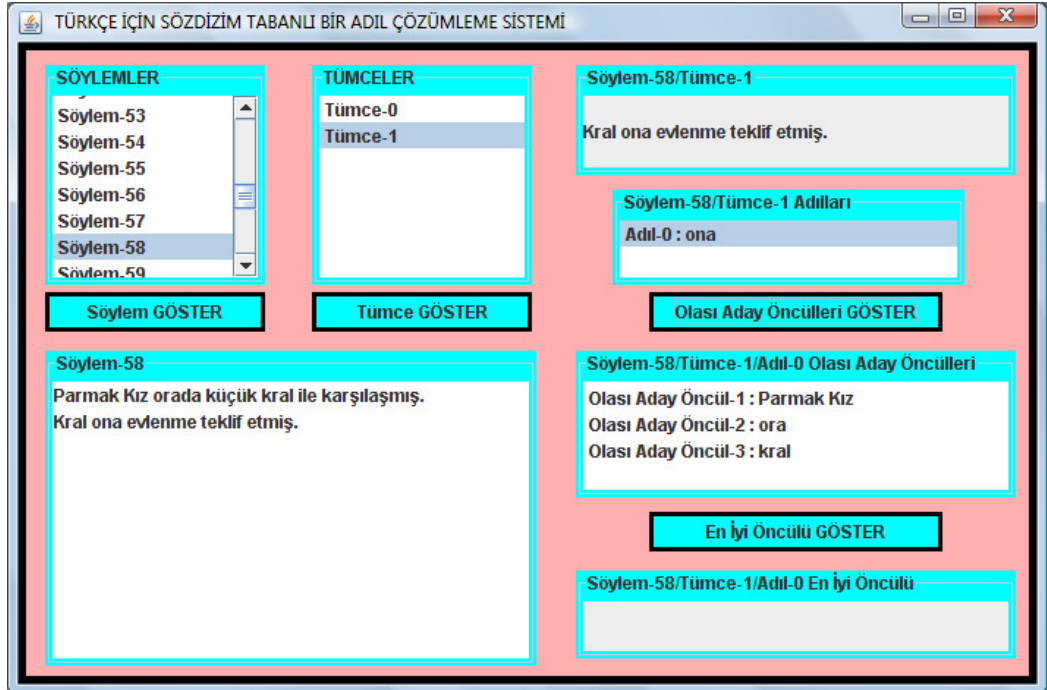
Şekil 4.3 Program Tarafından Yapılan Adıl Tespiti

2. Adım: İkinci adımda ise, adılın tespit edilmesiyle birlikte çözümleme işlemi de başlar ve bu aşamada algoritma uygulanır. Hobbs'un algoritmasına benzer bir şekilde, içinde adıl bulunan tümcenin ağaç yapısı üzerinde ilk adımda tespit edilen adıldan başlayarak yukarı yönde hareket edilir. S, NP ve VP noktalarına gelindiğinde durup, bu noktaların alt kısmında kalan ağacın dalları arasında soldan sağa doğru hareket edilerek uygun öncül olup olmadığı araştırılır ve bulunan AÖ'ler, *Parmak Kız*, *orada* ve *kral* aşağıdaki şekil 4.4'de de görüldüğü gibi olası aday öncül olarak sistem tarafından önerilir.



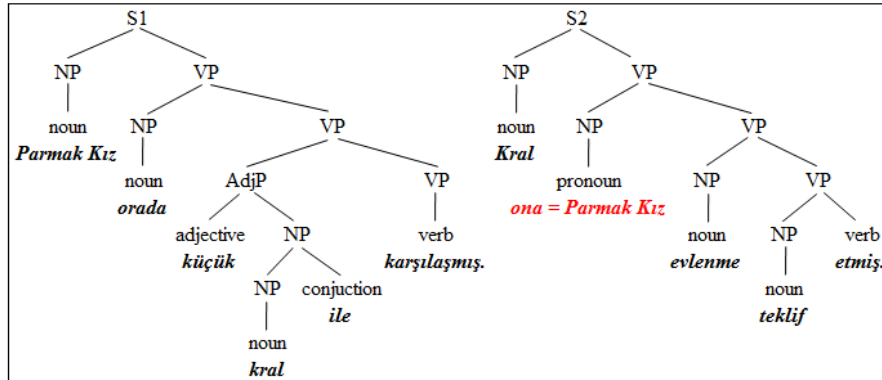
Şekil 4.4 Olası Aday Öncüllerin Tespiti

Programın ön yüzünü gösteren Şekil 4.5’de, ‘Olası Aday Öncülleri GÖSTER’ düğmesine tıklandığında algoritma yukarıdaki ağacı tarama işlemine başlar ve OAÖ’leri bulur. Program da bulunan bu OAÖ’leri, ‘Söylem-58/Tümce-1/Adıl-0 Olası Aday Öncülleri’ başlıklı kutucuk içinde gösterir.



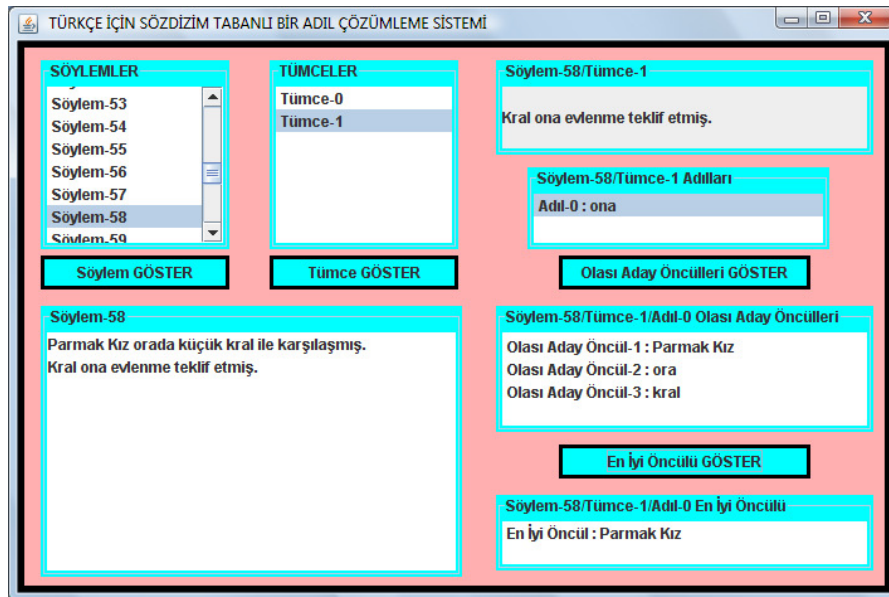
Şekil 4.5 Program Tarafından Olası Aday Öncüllerin Gösterilmesi

3. Adım: Son adımda da bulunan bu olası aday öncüllere bizim eleme kriteri olarak belirlediğimiz bir takım kısıt ve tercihler sırasıyla uygulanır. Her bir uygulamadan sonra bazı öncüller elenecektir. En son kalan olası aday öncülün, en iyi öncül olarak kabul edilmesiyle söz konusu adılın öncülü otomatik bir şekilde bilgisayar ortamında bulunmuş olur. Şekil 4.6’da gösterildiği gibi ‘ona’ adılının yerine bulunan öncülü ‘Parmak Kız’ konur ve program ağaç yapısını oluşturmaya kaldığı yerden devam eder. Daha sonra program, tümcede başka adıl olup olmadığını araştırır. Eğer bulamazsa, işlemi sonlandırır.



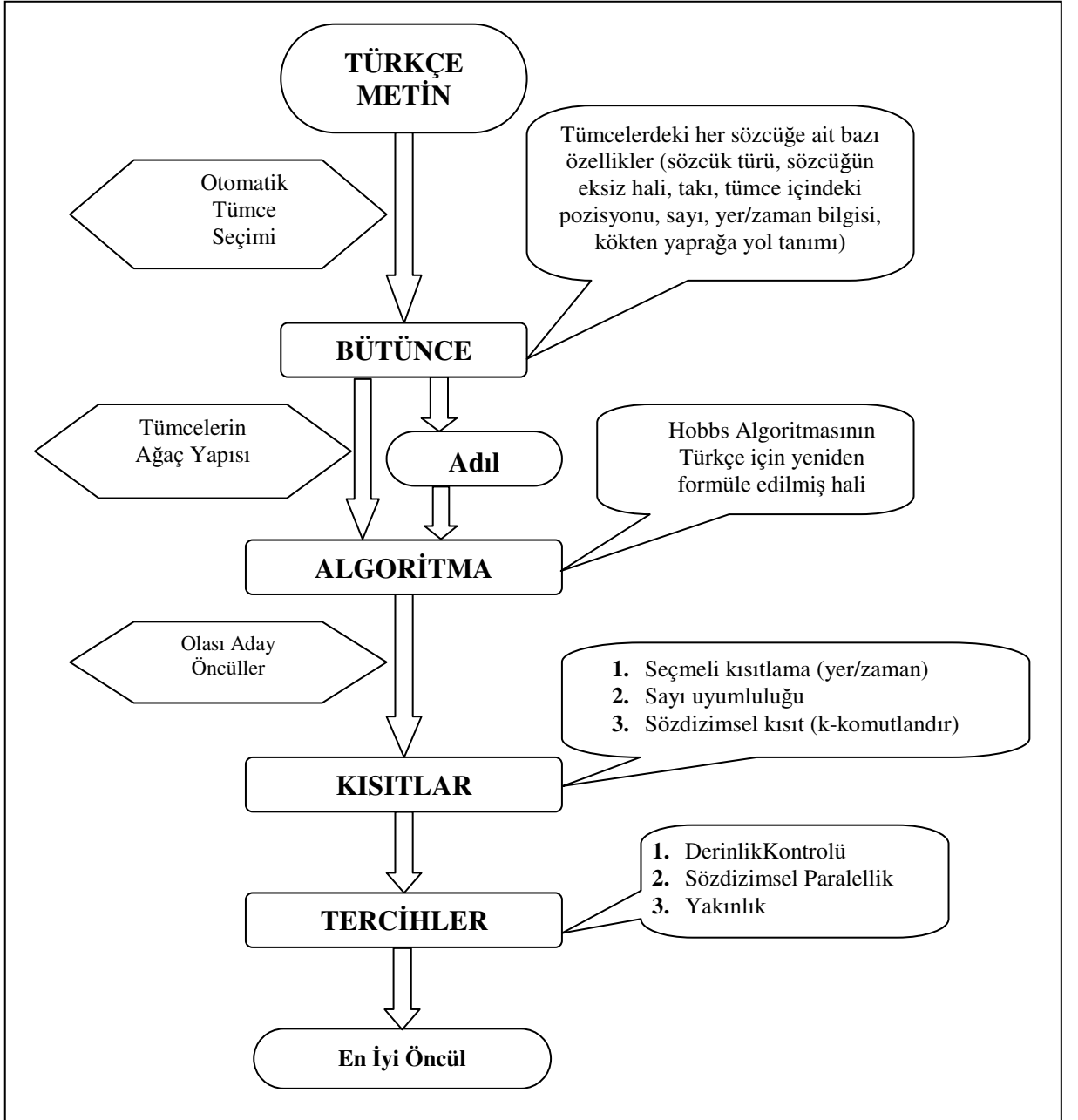
Şekil 4.6 En İyi Öncülün Seçimi

Programın ön yüzünü gösteren Şekil 4.7’de, ‘En İyi Öncülü GÖSTER’ düğmesine tıklandığında, programın arka planında kısıtlar ve tercihlere göre bulmuş olduğu en iyi öncül, ‘Söylem-58/Tümce-1/Adıl-0 En İyi Öncülü’ başlıklı kutucuk içinde gösterilir.



Şekil 4.7 Program Tarafından En İyi Öncülün Gösterilmesi

Biz de bu çalışmada, adılın öncülünü otomatik bir şekilde bilgisayar ortamında buldururken Şekil 4.1'deki temel adımları sistemimize uygulayıp, sistemimizi bu adımlara göre Şekil 4.8'deki gibi yapılandırmaya çalıştık.

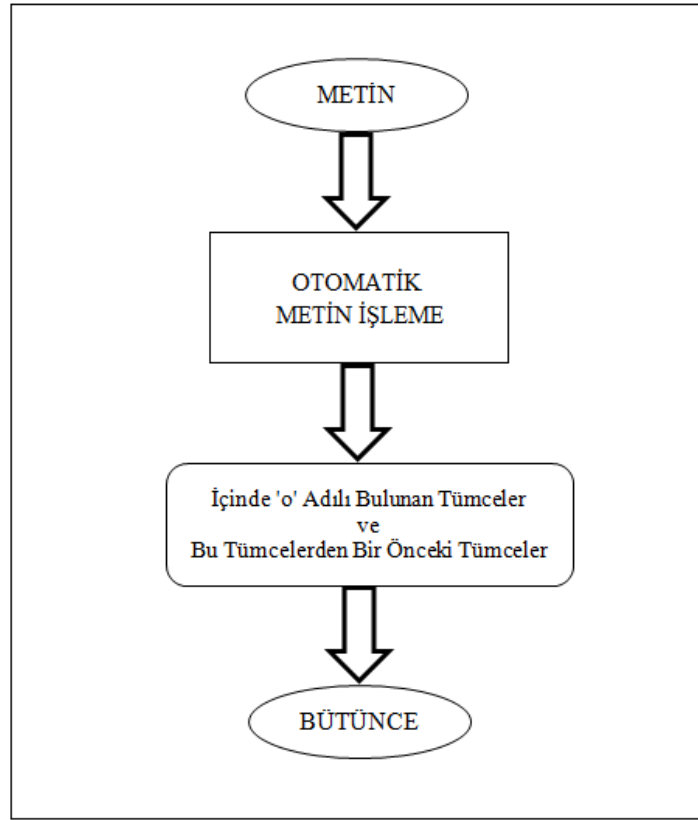


Şekil 4.8 Türkçe için Sözdizim Tabanlı Adıl Çözümleme Sistemi Akış Diyagramı

Şimdi Şekil 4.8'de gösterilen sistemimizin akış diyagramını ayrıntılı bir şekilde sırasıyla inceleyelim.

4.1.1 Bütünce

Bütünceye girilmeden önce, seçtiğimiz Tükçe metin ilk olarak ön bir işlemde geçirilir. Bu işlem, metnin otomatik bir şekilde taranıp, içinde sadece üçüncü kişi adlı 'o' olan tümcelerin ve bu tümcelerden bir önceki tümcelerin seçilmesi işlemidir. Daha sonra seçilen bu tümceler sistemimizde kullanılmak üzere bütünceye girilir. Her bir ikili tümce grubu numaralandırılır ve her biri birer söylem kabul edilip, bütünceye elle girilir. Aşağıdaki Şekil 4.9 bize metne uygulanan ön işlem aşamalarını göstermektedir.



Şekil 4.9 Seçilen Metine Uygulanan Ön İşlem

Bütünceye girilen her bir söylem yukarıda da belirtildiği gibi iki tümceden oluşur. Bu tümceler, söz konusu adılın olduğu tümce ve bu tümceden bir önceki tümcedir. Bu tümcelerdeki her bir sözcüğe ait bazı özellikler elle girilir. Bu özellikler; sözcük türü, sözcüğün eksiz hali, takı, tümce içindeki pozisyonu, sayı, yer/zaman bilgisi, kökten yaprağa yol tanımıdır. Bunları Tablo 4.1'de açık bir şekilde bütünce içeriği olarak görebiliriz.

Tablo 4.1 Bütünce İçeriği

	Söylem No	Tümce No	Sözcük No	Sözcük Türü	Ad Kökü	Takı	Dilbilgisel Görev	Tekil/Çoğul	Yer	Zaman	Ağaç Yapısı				
Fil	4	0	0	noun	fil	nom	sub	sg	-1	-1	S	NP	noun		
çok	4	0	1	-1	-1	-1	-1	-1	-1	-1	S	VP	AdjP	AdvP	adverb
iri	4	0	2	-1	-1	-1	-1	-1	-1	-1	S	VP	AdjP	AdjP	adjective
bir hayvandır.	4	0	3	-1	-1	-1	-1	-1	-1	-1	S	VP	verb	verb	
O	4	1	0	pronoun	-1	nom	sub	sg	-1	-1	S	NP	pronoun		
kocaman	4	1	1	-1	-1	-1	-1	-1	-1	-1	S	VP	AdjP	adjective	
bir	4	1	2	-1	-1	-1	-1	-1	-1	-1	S	VP	AdjP	NP	det
ağacı	4	1	3	noun	ağaç	acc	obj	sg	-1	-1	S	VP	AdjP	NP	noun
devirebilir.	4	1	4	-1	-1	-1	-1	-1	-1	-1	S	VP	VP	verb	
O	4	2	0	pronoun	-1	nom	sub	sg	-1	-1	S	NP	pronoun		
sonra	4	2	1	-1	-1	-1	-1	-1	-1	-1	S	VP	AdvP	adverb	
onu	4	2	2	pronoun	-1	acc	obj	sg	-1	-1	S	VP	VP	NP	pronoun
taşıyabilir.	4	2	3	-1	-1	-1	-1	-1	-1	-1	S	VP	VP	VP	verb

Tablo 4.1’de her sözcük için girilen özellikler aşağıda tanımlanmıştır:

- 1. Sözcük türü:** Sadece ad ve adıl olan sözcüklerin türleri belirtilmiştir.
- 2. Ad kökü:** Türü ad olan sözcüklerin, çekim eki almamış yalın halleri belirtilmiştir.
- 3. Takı:** Sözcük türü ad olan sözcüklerin aldığı takılar belirtilmiştir.
- 4. Dilbilgisel Görev:** Sözcük türü ad olan sözcüklerin tümce içinde özne ya da nesne olması durumu belirtilmiştir.
- 5. Tekil/Çoğul Durumu:** Sözcük türü ad olan sözcüklerin tekil/çoğul olma durumu belirtilmiştir.
- 6. Taşınmaz Bir Yer Olma Özelliği:** Sözcük türü ad olan sözcüklerin taşınmaz bir yer olma durumu belirtilmiştir. “Bir ada”, “oralarda”, “orayı” gibi isimler bir yeri belirten ifadelerdir.
- 7. Zamanla İlgili Olma Özelliği:** Sözcük türü ad olan sözcüklerin, zamanla ilgili bir kavram olma durumu belirtilmiştir. “Akşamlarda”, “bir gün”, “saat 10”, “bu sabah” gibi adlar zaman veya tarihle ilgili ifadelerdir.
- 8. Ağaç Yapısı:** Tümce içinde geçen tüm sözcük türleri için ağaç yapısını oluşturmada kullanılan her bir sözcüğe ait kökten (S), yaprağa kadar izlenen yol bilgisi belirtilmiştir.

Tablo 4.1’de görüldüğü gibi bütüncüye sadece sözcük türü ad olan sözcüklerin, isim kökü, takı, dilbilgisel görev, tekil/çoğul durumu, yer ve zaman olma bilgileri girilmiştir. Sözcük türü ad olmayan sözcükler için ise bu bilgiler, kullanılmayacağı için -1 ile ifade

edilmiştir. Bu özellikler dışında kalan, ağaç yapısının tanımlandığı kökten yaprağa yol tanımı bilgisi, bütün sözcükler için ayrı ayrı girilmiştir.

Bu çalışmada Tablo 4.2’de içeriği verilen iki farklı bütünce kullanılmıştır. Bütünce-1, 96 söylemden oluşmaktadır. Bu söylemler, 266 tümce ve 1.535 sözcük içerir. Bunların 202’si kişi adlıdır. Bütünce-2 ise 205 adıl ve 11.315 sözcükten oluşmaktadır.

Tablo 4.2 Sistemin Bütünceleri

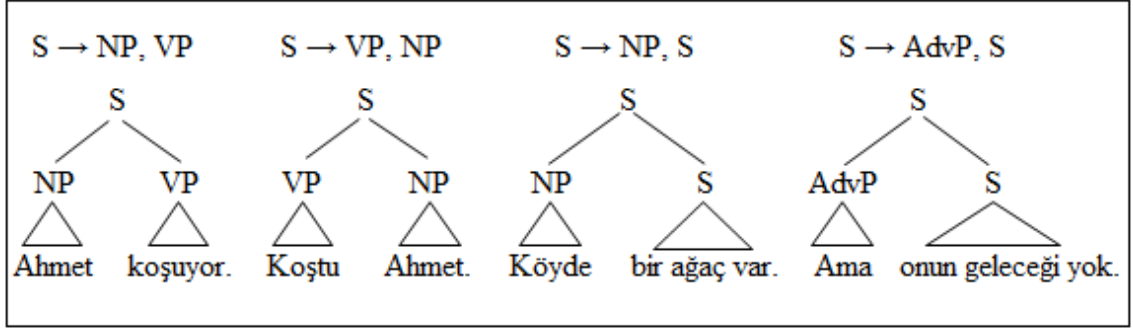
Sistem Bütünceleri	Adıllar (Özel İsmeye Gönderimde Bulunan)		Adıllar (Özel İsim Dışındaki İsimlere Gönderimde Bulunan)		Toplam
	Adıl	Dönüşlü	Adıl	Dönüşlü	
Sözdizimsel Tabanlı Sistemin Eğitim Bütüncesi	89	0	113	0	202
Az Bilgili Sistemin Eğitim Bütüncesi	190	15	0	0	205

4.1.2 Algoritma

Hobbs’un algoritmasını Türkçe için yeniden düzenleyerek adıl çözümlemesi yapabilen yeni bir algoritma oluşturduk. Chomsky’nin Bağlama Kuramı’nın Türkçe’ye olan etkilerini de göz önüne alarak oluşturduğumuz yeni algoritma, söz konusu adılın bulunduğu tümce ile kendinden önceki tümcenin sözdizim ağaçlarını girdi olarak alıp, Hobbs’un Algoritması’nda olduğu gibi tümcelerin ayrıştırma ağaçlarını tarayarak adıllarla eşgönderimli olan olası aday öncülleri bulmaktadır. Ağaç üzerinde ilerlerken, Türkçe’de geçerli olan bağlama özelliklerinden k-komutlandır özelliğini kullanan bir yol izlenir.

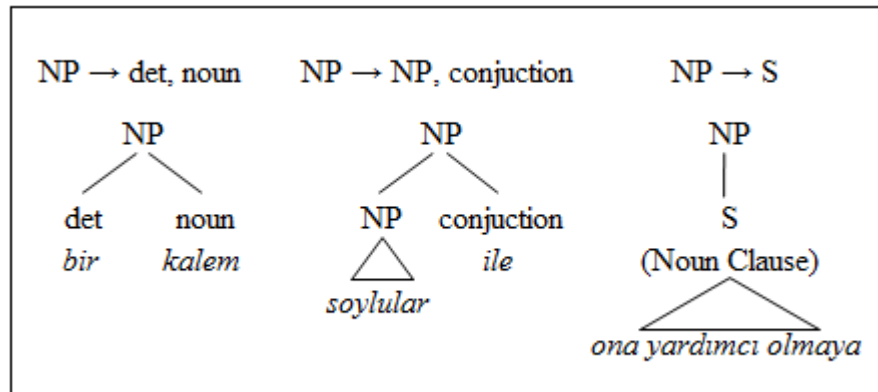
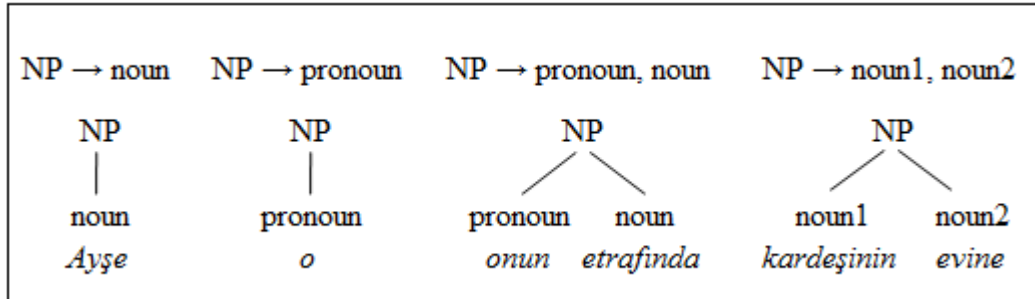
Algoritma, tümcelerin belli dilbilgisi kuralları çerçevesinde bir ayrıştırma ağacı formatında gösterildiğini kabul eder. Bundan dolayı ağacı tarayacak olan algoritma bir dilbilgisi belirlemek zorundadır. Çünkü sözdizimsel olarak ağacın yapısına ilişkin kabuller sonuçları etkileyecektir. Bütüncedeki tümcelerin ağaç yapıları, Türkçe için aşağıda kuralları belirtilmiş olan bağlamdan bağımsız dilbilgisine göre oluşturulmuştur. Ağaç yapılarının oluşturulmasında, ikili ağaç yapısı modeli benimsenmiştir.

Aşağıdaki Şekil 4.10'da görüldüğü üzere S (Sentence) tümceye karşılık gelir. Algoritmamızda kullandığımız S'e ait dilbilgisi kuralları ikili ağaç yapılarıyla birlikte belirtilmiştir.



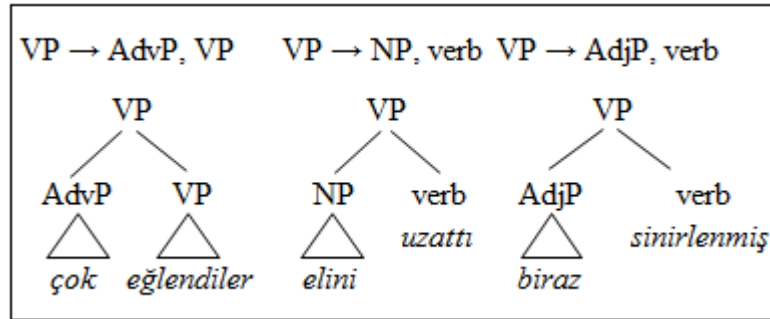
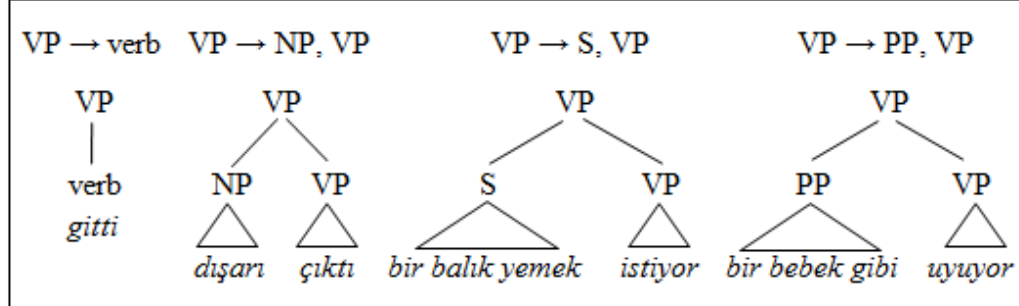
Şekil 4.10 'S' Tümce için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Şekil 4.11'de NP (Noun Phrase) Ad öbeği için algoritmada kullandığımız dilbilgisi kurallarının ağaç yapılarıyla birlikte gösterimine yer verilmiştir.



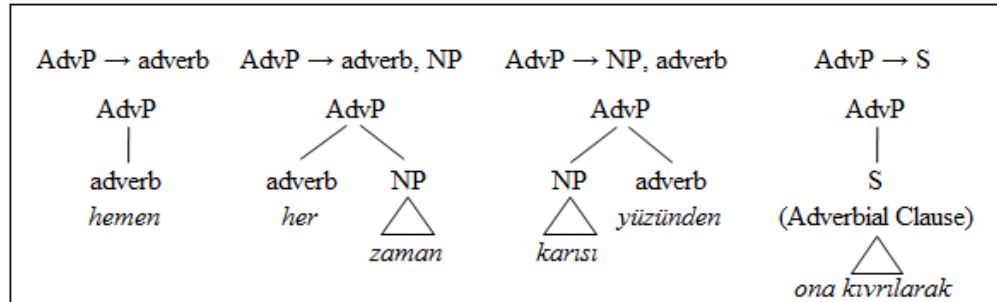
Şekil 4.11 'NP' Ad Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Şekil 4.12’de VP (Verb Phrase) Eylem öbeği için algoritmada kullandığımız dilbilgisi kuralları ağaç yapılarıyla birlikte gösterilmiştir.



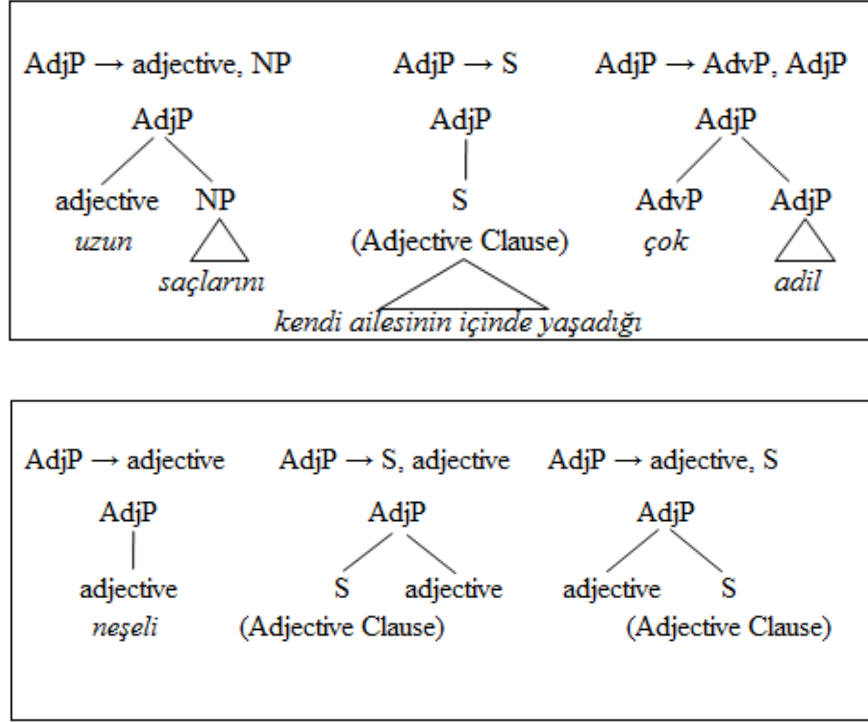
Şekil 4.12 ‘VP’ Eylem Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Şekil 4.13’de AdvP (Adverb Phrase) Belirteç öbeği için algoritmada kullanılan dilbilgisi kurallarının ağaç yapılarıyla birlikte gösterimine yer verilmiştir.



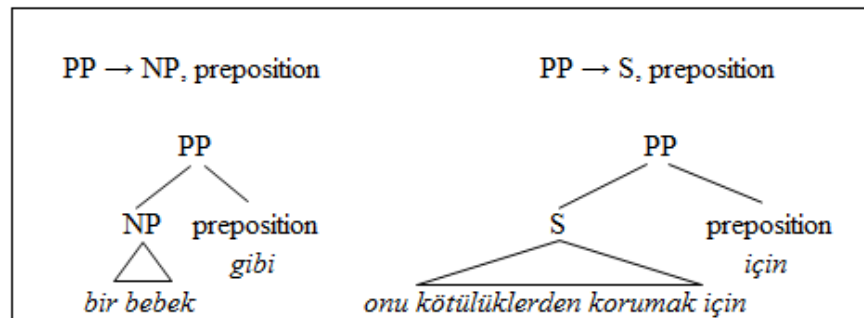
Şekil 4.13 ‘AdvP’ Belirteç Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Şekil 4.14’de AdjP (Adjective Phrase) Sıfat öbeği için algoritmada kullandığımız dilbilgisi kurallarının ağaç yapılarıyla birlikte gösterimine yer verilmiştir.



Şekil 4.14 'AdjP' Sıfat Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Şekil 4.15’de PP (Preposition Phrase) İlgeç öbeği için algoritmada kullandığımız dilbilgisi kurallarının ağaç yapılarıyla birlikte gösterimine yer verilmiştir.



Şekil 4.15 'PP' İlgeç Öbeği için Türkçe Dilbilgisi Kuralları ve Ağaç Yapıları

Algoritma; cümlelerin ayrıştırma ağaçları üzerinde en uç noktadaki kişi adlından başlayıp üst noktalara doğru hareket ederek olası öncüllere ulaşmaya çalışmaktadır. S, NP veya VP noktalarına gelindiğinde durup, bu noktaların alt kısmında kalan ağacın dalları arasında soldan sağa doğru giderek uygun öncül olup olmadığını araştırır. Adılın olası öncülleri, Chomsky'nin Bağlama Kuramı'nın ikinci prensibine uyularak, adılın içinde bulunduğu tümcede değil de bir önceki tümcede aranır.

Hobbs Algoritmasının Türkçe için yeniden formüle edilmiş hali şöyledir:

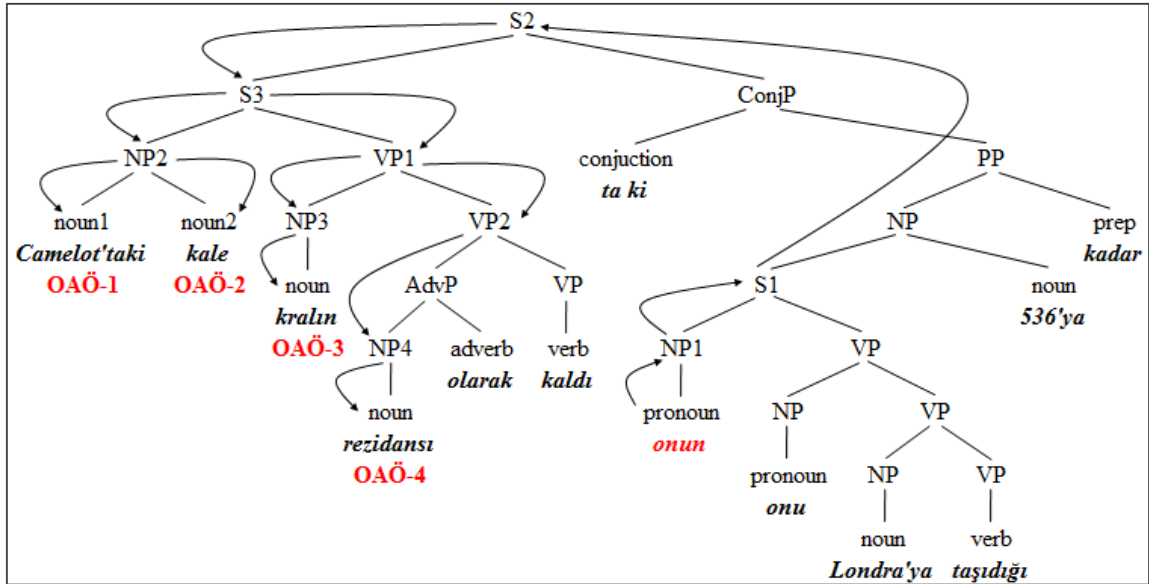
1. S tümcesinin ağaç yapısı üzerinde, adılı doğrudan içeren NP düğümünden başla.
2. Ağaçta ilk karşılaşılan VP veya S düğümüne çık. Bu düğümü X , bu düğümüne ulaşmak için izlenen yolu da p diye adlandır.
3. X düğümünün altındaki p yolunun solunda kalan tüm dalları, soldan sağa yönde dikkatle incele. Karşılaşılan herhangi bir NP düğümü ile X düğümü arasında bir S düğümü var ise, karşılaşılan ilk NP düğümünü öncül olarak öner.
 - (i) Eğer X düğümünün altında herhangi bir S düğümü var ise ve X düğümü ile S düğümü arasında kalan bir NP düğümü var ise, bu düğümü eleme faktörlerinden sözdizimsel kısıtta kullanmak üzere sakla. Ama bu NP düğümü bu kritere göre öncül olamaz.
 - (ii) Eğer X düğümünün altında herhangi bir S düğümü yok ise ve X düğümünün altında bir NP düğümü var ise, bu düğüm eleme faktörlerinden sözdizimsel kısıtta kullanmak üzere sakla. Ama bu NP düğüm bu kritere göre öncül olamaz.
4. Eğer X düğümü tümcenin en üst S düğümü ise, en son gelen ilk olmak şartıyla metinde önceki tümcelerin ağaç yapılarını dikkatle incele; her bir ağacı soldan sağa yönde incele ve bir NP düğümüyle karşılaşıldığında, bu düğümü öncül olarak öner. Eğer X tümcenin en üst S düğümü değil ise, 5. adıma git.

5. X düğümünden, ağaç üzerinde yukarı yönde karşılaşılan ilk VP ya da S düğümüne çık. Bu düğümü X, bu düğüme ulaşmak için izlenen yolu da p diye adlandır. 3. adıma git.

Tümce 4.1’de, iki tane kişi adlı bulunmaktadır. Algoritma birinci adılın çözümlenmesi işlemiyle başlar, daha sonra da ikinci adılın çözümlenmesi işlemine geçer. Birinci adılın öncülü bulunduktan sonra, birinci adıl yerine ağaç üzerinde öncülü konur ve ikinci adılın çözümlenmesi işlemine devam edilir.

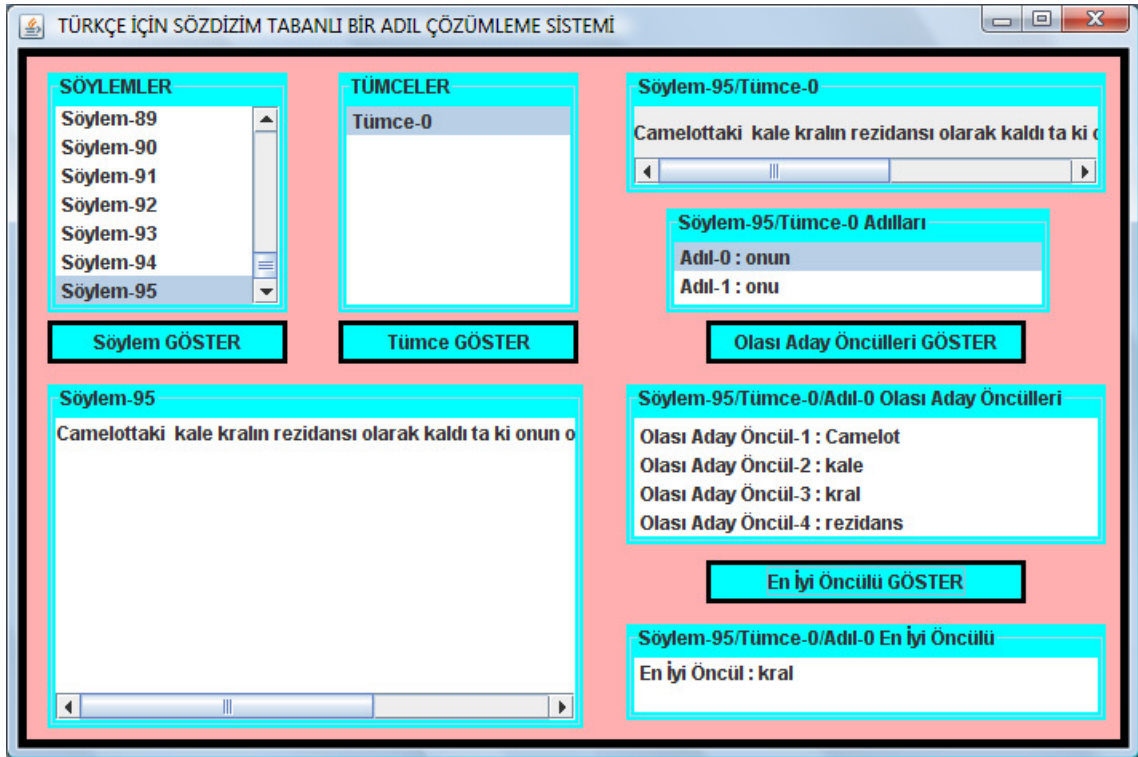
Camelot-ta-ki kale kral-ın rezidans-ı olarak kaldı ta ki o-nun
 Camelot-LOC-REL castle king-GEN residence-ACC as remain-PAST when s/he-GEN-3SG
 o-nu Londra-ya taşı-dı-ğı 536-ya kadar. (4.2)
 it-ACC Londra-DAT move-PAST-ACC 536-DAT until.

Şekil 4.16, algoritmanın tümce 4.2’deki ‘onun’ adılı için nasıl çalıştığını göstermektedir.



Şekil 4.16 Tümce 4.2’nin Ağaç Yapısı ve ‘onun’ Adılı için Algoritmanın İşleyişi

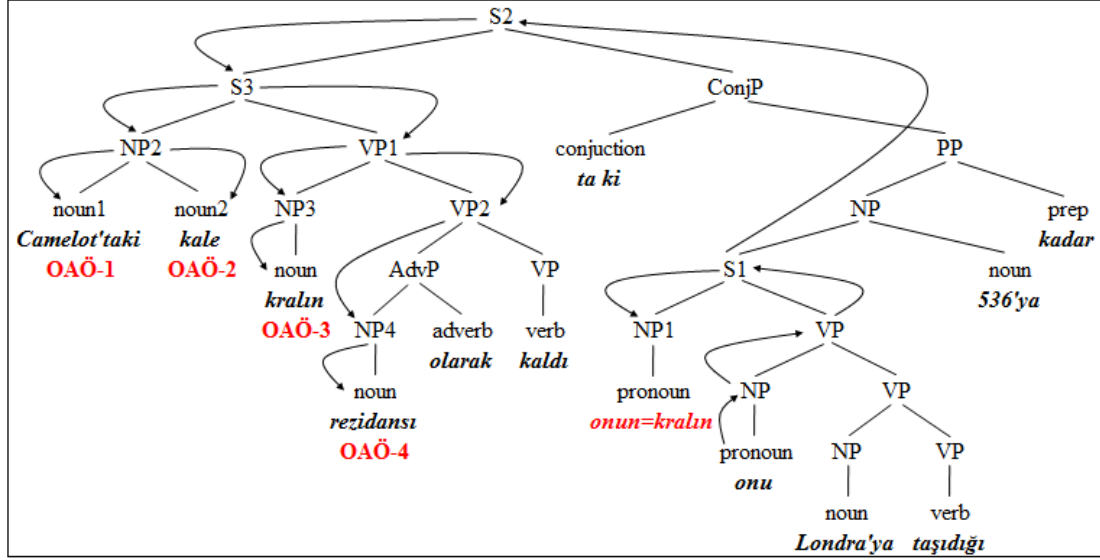
Şekil 4.16'daki algoritmanın birinci adımı, 'onun' adını doğrudan içeren NP₁ düğümünü bularak başlar. 2. Adım ile S₁ düğümüne çıkılır. Bu düğümüne X ve bu düğümüne ulaşmak için izlenen yola p₁ denilir. 3. Adım uygulanmaz, çünkü X düğümünün altında p₁ yolunun solunda kalan herhangi bir dal bulunmamaktadır. 4. Adım da uygulanmaz, çünkü S₁ düğümü, cümledeki en üst noktadaki S düğümü değildir. Böylece, 2. adımdan 5. adıma geçilir. 5. Adım ile S₂ düğümüne çıkılır. Bu düğümüne X ve bu düğümüne ulaşmak için izlenen yola p₂ denilir ve algoritma 3. adımla devam eder. 3. Adım, X düğümünün altında p₂ yolunun solunda kalan dalları araştırır ve S₃ düğümüne sahip üç tane NP düğümü bulur. Şekil 4.17'de de görüldüğü gibi Adıl Çözümleme Algoritması adımları verilen programın tümce 4.2 için çalışması sonucunda, 'onun' adlı için 'Camelot', 'kale', 'kral' ve 'rezidans' olası aday öncülleri (OAÖ) bulunur.



Şekil 4.17 Tümce 4.2 'Onun' Adılı için Program Çıktısı

Daha sonra bu olası aday öncüllere, kısıtların uygulanmasıyla 'Camelot' ve 'kale' olası aday öncülleri elenir. Kalan olası aday öncüllerin sayısı 1 olmadığı için tercihler uygulanır. Sözdizimsel paralellik tercihiyle, 'rezidansı' olası aday öncülü elenir ve 'kral' aday öncülü, 'onun' adının en iyi öncülü olarak doğru bir şekilde bulunur.

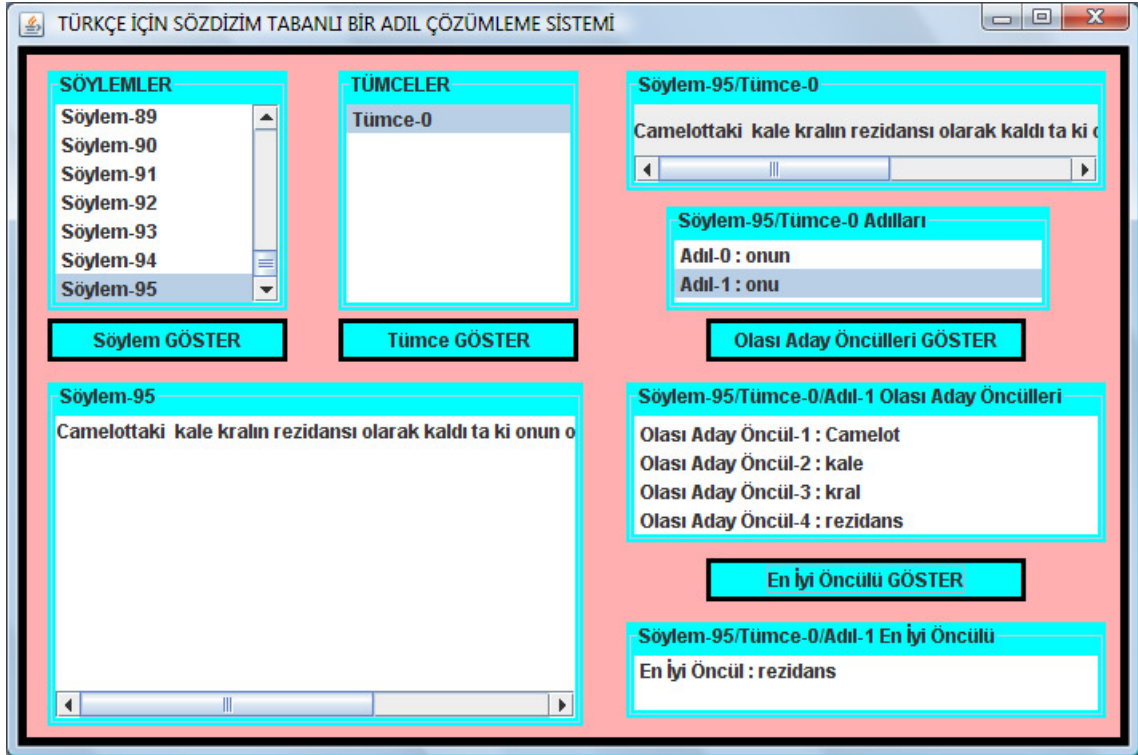
Tümcenin ağaç yapısı üzerinde, ‘onun’ adılıının yerine, öncülü konur ve ikinci adıl ‘onu’ için algoritma Şekil 4.18’de görüldüğü gibi çalışmasına devam eder.



Şekil 4.18 Tümce 4.2'nin Ağaç Yapısı ve ‘onu’ Adılı için Algoritmanın İşleyişi

Şekil 4.18’de, algoritmanın birinci adımı, ‘onu’ adılıını doğrudan içeren NP₁ düğümünü bularak başlar. 2. Adım ile VP₁ düğümüne çıkılır. Bu düğüme X ve bu düğüme ulaşmak için izlenen yola p₁ denilir. 3. Adım uygulanmaz, çünkü X düğümünün altında p₁ yolunun solunda kalan herhangi bir dal bulunmamaktadır. 4. Adım uygulanmaz, çünkü X düğümü, cümlenin en üst noktasındaki S düğümü değildir. Böylece, 2. adımdan 5. adıma geçilir. 5. Adım ile S₁ düğümüne çıkılır. Bu düğüme X ve bu düğüme ulaşmak için izlenen yola p₂ denilir. Ardından algoritma 3. adımla devam eder. 3. Adım, X düğümünün altında p₂ yolunun solunda kalan dalları araştırır, ve NP₂ düğümüne rastlanır. Bu NP düğümü en iyi öncülün seçilmesi sırasında kullanılacak olan sözdizimsel kısıtta kullanılmak üzere saklanır. Sözdizimsel kısıta göre NP₂ aday öncül kabul edilemez. 4. Adım uygulanmaz ve 3. adımdan 5. adıma geçilir. 5. Adımda S₂ düğümüne çıkılır. Bu düğüme X ve bu düğüme ulaşmak için izlenen yola p₃ denilir. Algoritma 3. adımla devam eder. 3. Adım, X düğümünün altında p₃ yolunun solunda kalan dalları araştırır, ve S₃ düğümüne sahip üç tane NP düğümü bulur: ‘Camelot’taki’, ‘kale’, ‘kralın’ ve ‘rezidansı’ aday öncüller olarak kabul edilir. Şekil 4.19’daki

programın çalışmasını gösteren ekran çıktısında da ‘onu’ adlı için sistemin bulduğu olası aday öncülleri görülmektedir.



Şekil 4.19 Tümce 4.2 ‘onu’ Adılı için Program Çıktısı

Daha sonra bu aday öncüller, kısıtların uygulanmasıyla ‘Camelot’taki’, ‘kale’ ve ‘kralın’ aday öncülleri elenir. Geriye bir aday öncül kaldığı için, ‘rezidansı’, ‘onu’ adılının öncülü olarak doğru bir şekilde bulunur. Bu durumda tercihlerin uygulanmasına gerek kalmaz.

4.1.3 Kısıtlar

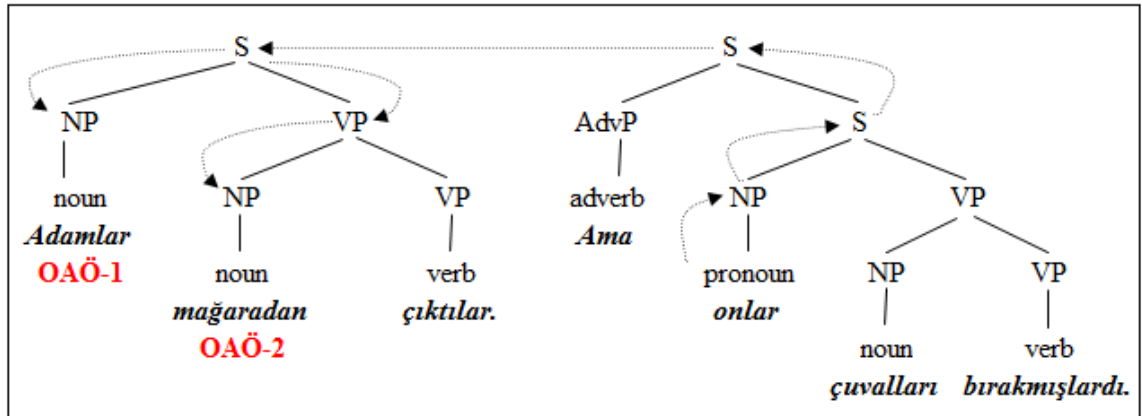
En iyi öncülün seçilebilmesi için olası aday öncüllere bazı kriterler uygulanır. Bunlar kısıtlar ve tercihlerdir. Kısıtlar, olası aday öncüllere uygulanan zorunlu kriterlerdir. Bu kriterlere uymayan olası aday öncüller elenir. Kısıtlar olası aday öncüllere uygulandıktan sonra, eğer olası aday öncül sayısı 1’e düşmüşse tercihlere gerek kalmaz ve uygulanmazlar. Bu çalışmada, ilgili adılın en iyi öncülünü bulmak için algoritma

tarafından bulunan OAÖ'lere sırasıyla uygulanan kısıtlar, aşağıda detaylı bir şekilde örneklerle açıklanmıştır.

1. Seçmeli Kısıtlama: Algoritmanın bulduğu OAÖ'lere sistem tarafından uygulanan ilk kısıtlama, bizim belirdeğimiz bir kısıtlama olan seçmeli kısıtlamadır. Bu kısıtlama, adıl ve öncülü için aynı semantik özelliklerin taşınması gerektiğini şart koşar. Bu çalışmada, her bir ad öbeğine ait bir yer ya da zaman olma özelliğini belirten bilgiler bütüncüye elle girilmiştir. Böylece, basit anlamda semantik bilgisi kullanmış olmaktadır. OAÖ'lerin bu semantik özelliklerine bakıldığında eğer yer ya da zaman olma özelliği taşıyorsa bu OAÖ'ler elenir. Çünkü, kişi adlarının yer ya da zaman olmayla ilgili semantik özellikleri yoktur.

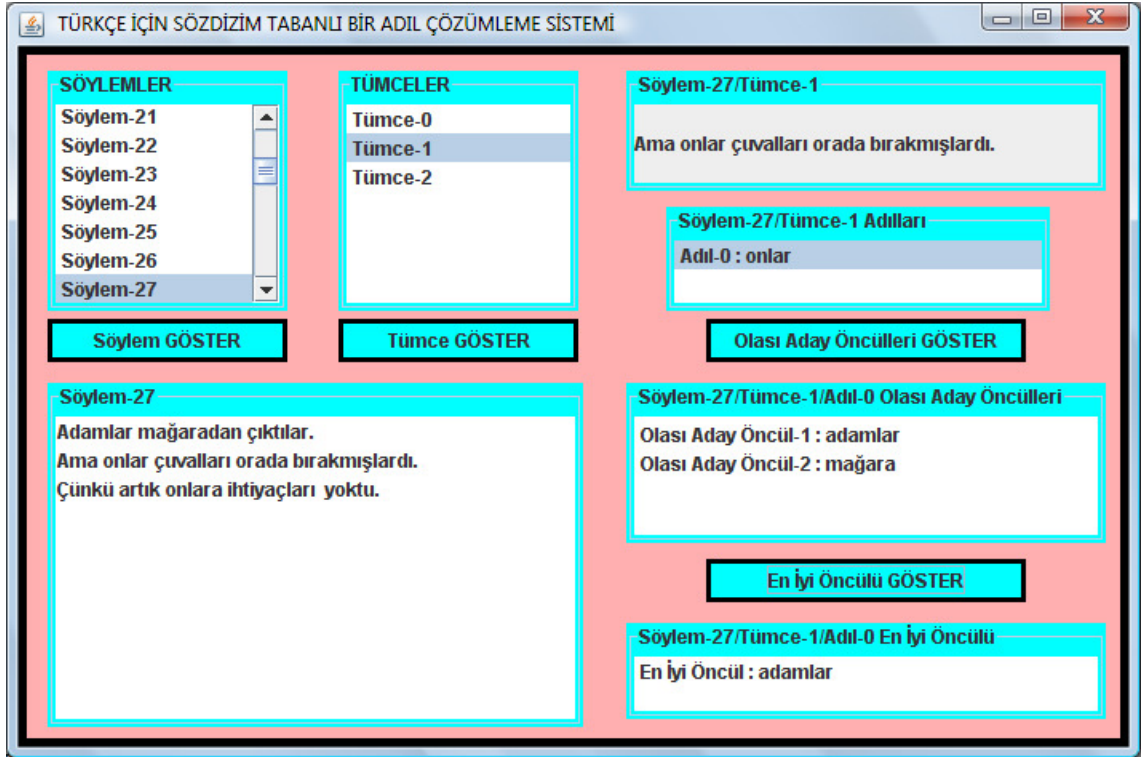
Söylem 4.3'de ikinci tümcede geçen *onlar* adını çözümlmeye çalıştığımızda Şekil 4.20'den de anlaşılacağı üzere algoritma, ağaç yapılarını belirlediğimiz kurallara göre tarayarak, OAÖ'leri *adamlar* ve *mağara* olarak bulur. *Mağara* OAÖ'ü, semantik olarak bir yeri belirten ad öbeğidir. Bu sebeple elenir ve kalan OAÖ *adamlar*, *onlar* adılı için en iyi öncül olarak bulunmuş olur.

Adamlar mağaradan çıktı. Ama onlar çuvalları orada bırakmıştı. (4.3)



Şekil 4.20 Söylem 4.3'nin Ağaç Yapısı ve 'onlar' Adılı için Algoritmanın İşleyişi

Şekil 4.21'de de bu işlemin program tarafından çözümünü gösteren ekran çıktısı yer almaktadır.

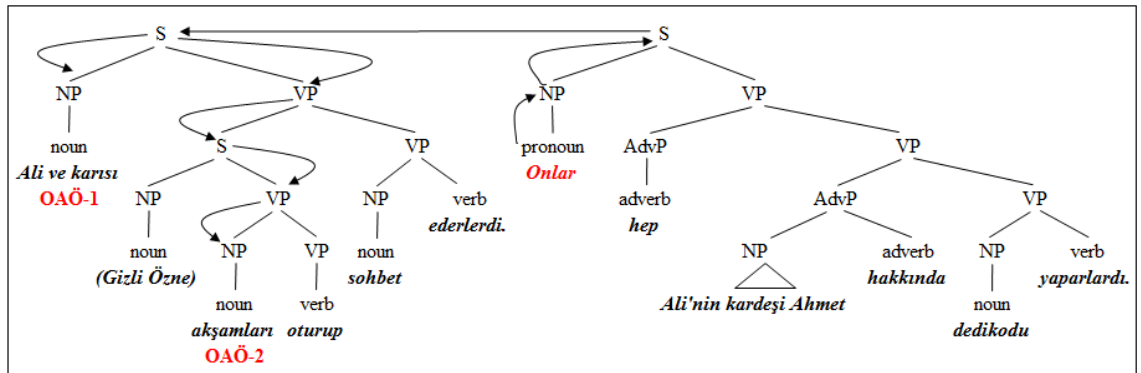


Şekil 4.21 Söylem 4.3'ün 'onlar' Adılı İçin Program Çıktısı

Söylem 4.4'ün ikinci tuncesindeki *onlar* adılı için de Şekil 4.22'de olası aday öncüllerin nasıl bulunduğu gösterilmektedir. Bulunan olası aday öncüller, *Ali ve karısı* ile *akşamlar* ad öbekleridir. *Akşamlar* olası aday öncülü zamana ait bir özelliği belirttiği için elenir ve *onlar* adılıının öncülü *Ali ve karısı* olarak doğru bir şekilde bulunmuş olur.

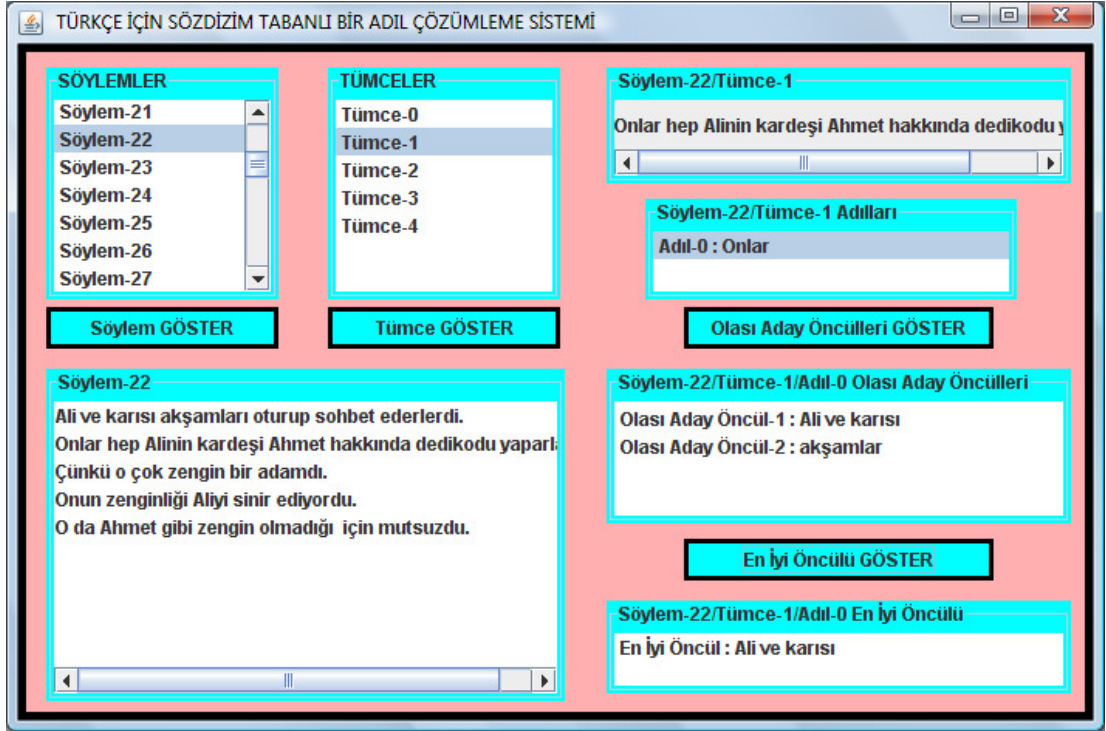
Ali ve karısı akşamları oturup sohbet ederlerdi. (4.4)

Onlar hep Ali'nin kardeşi Ahmet hakkında dedikodu yaparlardı.



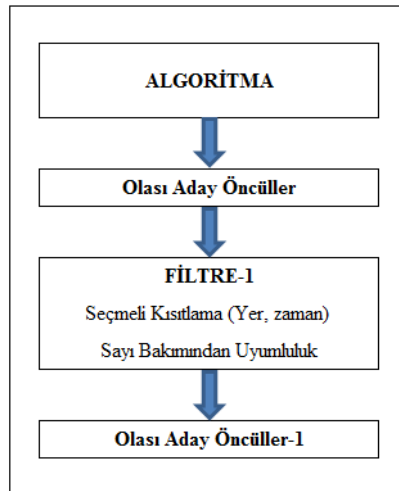
Şekil 4.22 Söylem 4.4'ün Ağaç Yapısı ve 'onlar' Adılı için Algoritmanın İşleyişi

Şekil 4.23’de de program tarafından elde edilen bu sonuç görsel olarak ifade edilmiştir.



Şekil 4.23 Söylem 4.4’ün ‘onlar’ Adılı için Program Çıktısı

2. Sayı Bakımından Uyumluluk Kısıtı: Adıl ve öncülü, kişi ve sayı olarak uyumlu olmak zorundadır. Algoritmanın belirlediği OAÖ’lerin seçmeli kısıtlamaya göre (yer ve zaman özelliklerinin kontrolü) ve adıl ile sayı bakımından uyumluluğunu kontrol eden, programda Filtre-1 diye tanımladığımız bir eleme süzgeci vardır. Şekil 4.24’den de anlaşılacağı üzere bütün OAÖ’ler bu süzgeçten geçerler.

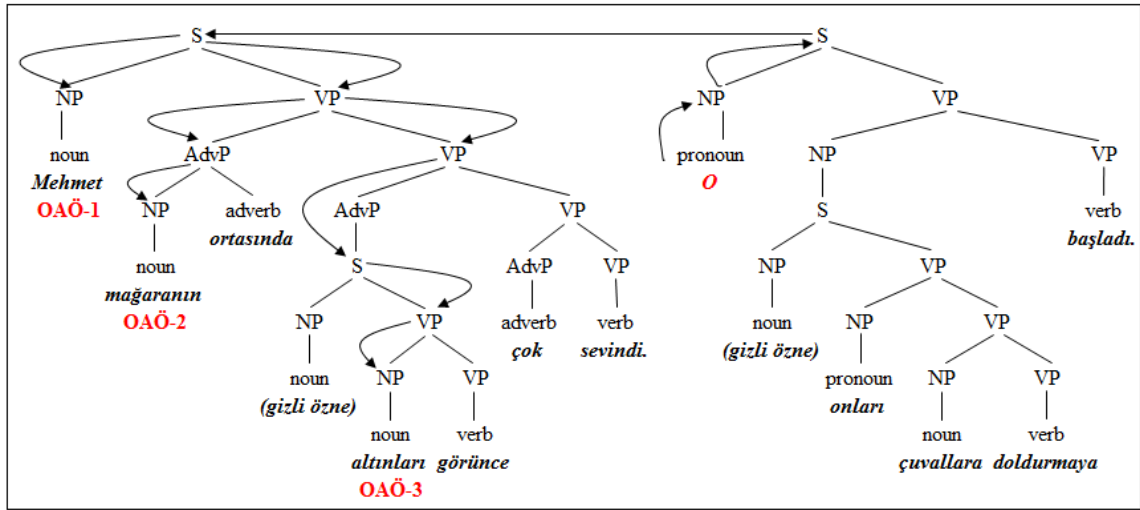


Şekil 4.24 Filtre-1

Söylem 4.5’de ikinci tümcede iki tane adıl bulunmaktadır. Şekil 4.25’de de görüldüğü gibi sistem önce ilk adıl olan *o* adılı için çözümleme işlemini yapar ve OAÖ’leri *Mehmet*, *mağara* ve *altınlar* olarak bulur.

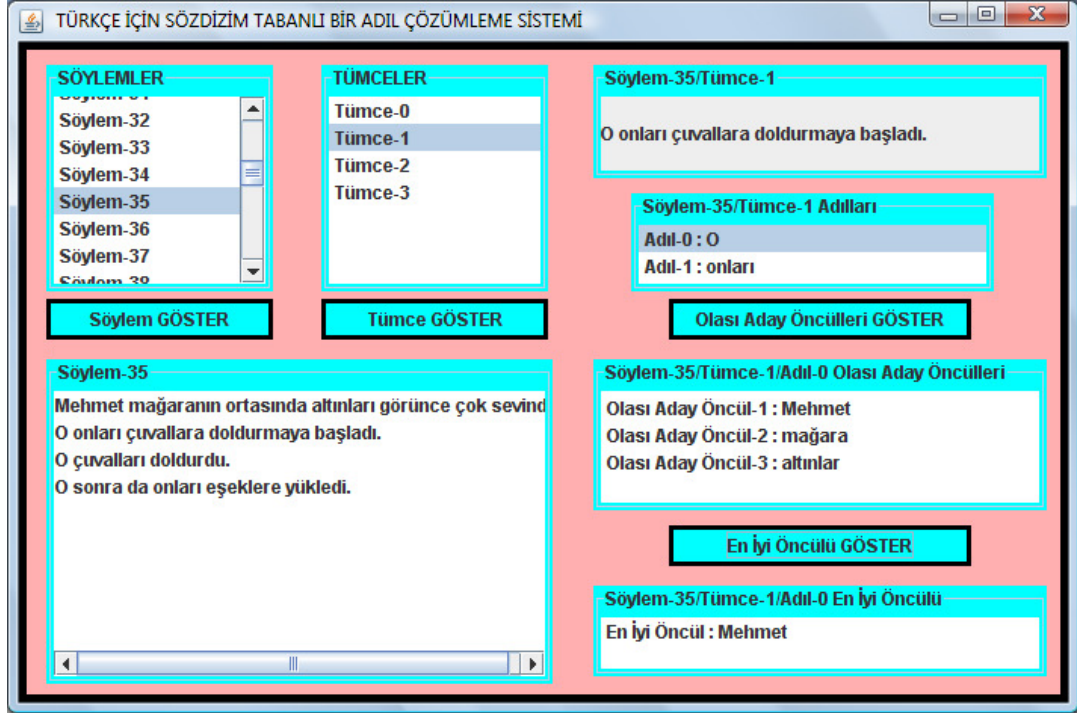
Mehmet mağaranın ortasında altınları görünce çok sevindi. (4.5)

O onları çuvallara doldurmaya başladı.



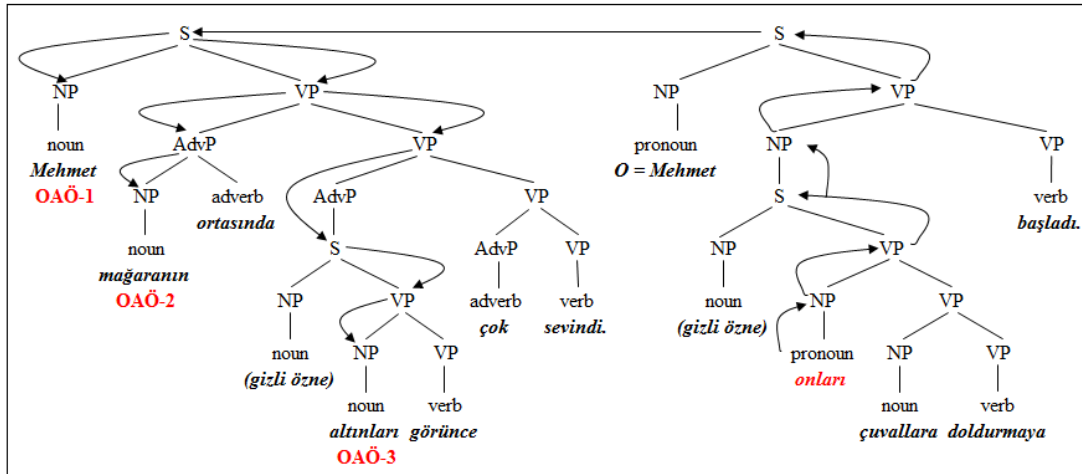
Şekil 4.25 Söylem 4.5'in Ağaç Yapısı ve 'o' Adılı için Algoritmanın İşleyişi

Bulunan bu OAÖ'ler, programda Filtre-1'den geçirilerek, ilk olarak yer ve zaman olma özelliklerine bakılır ve *mağara* OAÖ'ü, yer olma özelliği taşıdığından elenir. Daha sonra kalan *Mehmet* ve *altınlar* OAÖ'lerinin, *o* adılı ile sayı uyumluluğuna bakılır burada da *altınlar* OAÖ'ü elenir. Bunun sonucunda *o* adılının öncülü *Mehmet* olarak doğru bir şekilde bulunmuş olur. Programın ön yüzünde, 35 nolu bu söylem seçildiğinde, 'o' adılı için bulunan OAÖ'leri ve en iyi öncülü gösteren Şekil 4.26'daki ekran çıktısı elde edilir.



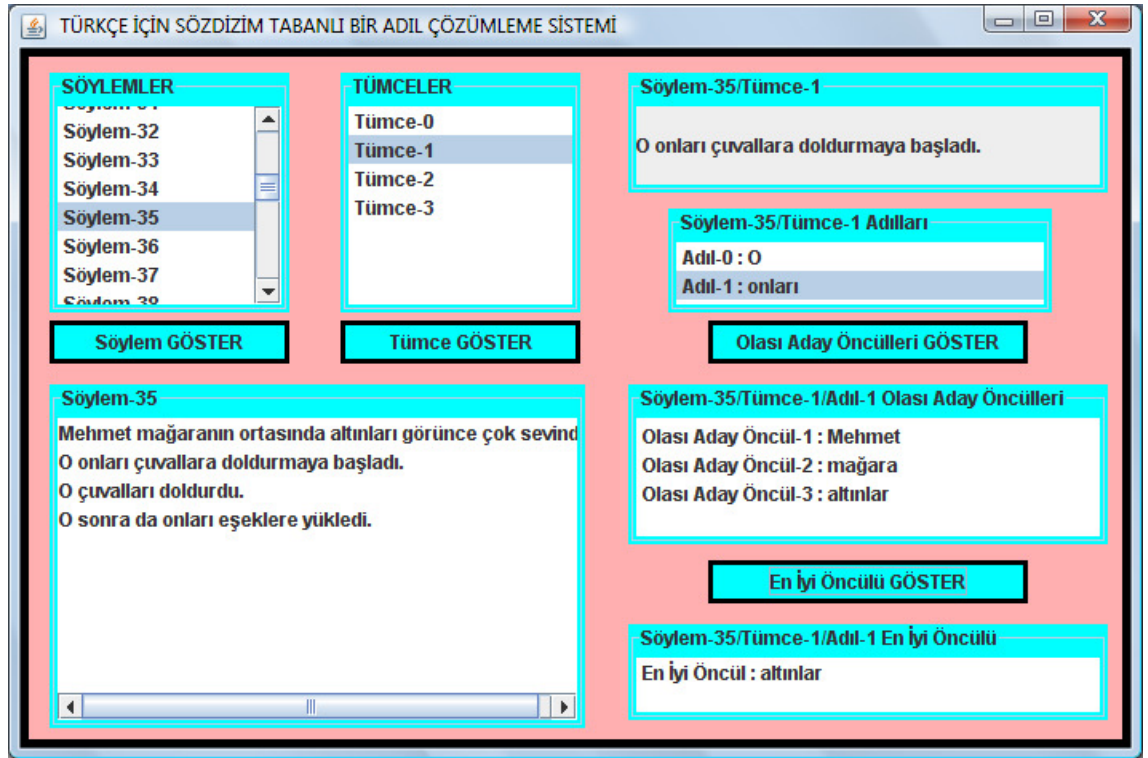
Şekil 4.26 Söylem 4.5'in 'o' Adılı İçin Program Çıktısı

Daha sonra Şekil 4.27'de de görüldüğü gibi *o* adının yerine bulunan öncülü *Mehmet* konularak, ağaç üzerinde adıl araştırma işlemine devam edilir ve çözümlenecek ikinci adıl olan *onlar* adılına rastlanır. *O* adılı için bulunan OAÖ'lerin aynıları *onlar* adılı için de bulunur ve aynı şekilde programda Filtre-1'den geçirilir. Önce *mağara* OAÖ'ü elenir, daha sonra da *Mehmet* OAÖ'ü sayı uyumluluğu bakımından elenir. *Onlar* adılı için öncülü *altınlar* olarak doğru bir şekilde bulunmuş olur.



Şekil 4.27 Söylem 4.5'in Ağaç Yapısı ve 'Onları' Adılı İçin Algoritmanın İşleyişi

Şekil 4.28’de de program tarafından elde edilen bu sonuç görsel olarak ifade edilmiştir



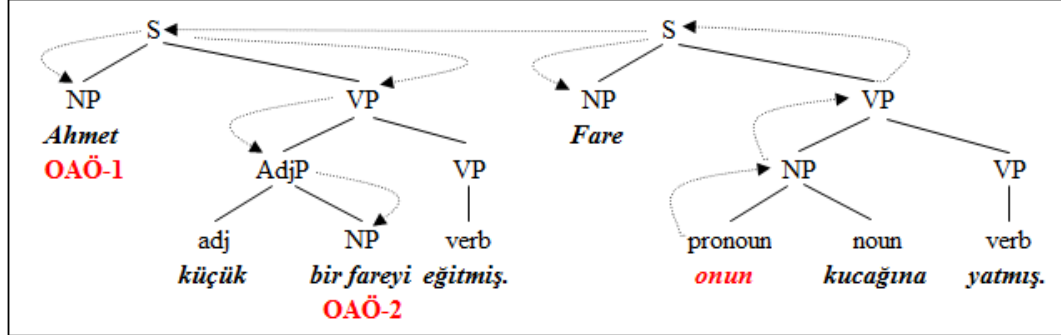
Şekil 4.28 Söylem 4.5’in ‘onları’ Adılı için Program Çıktısı

3. Sözdizimsel Bağlama Kuramı Kısıtlamaları: Bölüm 3.1.3.1’de de açıklanan bu kısıtlama ile Chomsky’nin Bağlama Kuramının (1981) artgönderimlerin çözümlenmesinde ne kadar büyük bir rol oynadığı anlaşılmaktadır. Bağlama Kuramının dönüşlü adılar, adılısıllar ve gönderimsel ifadeler için ayrı ayrı yorumladığı sözdizimsel kısıtlamalardan, biz bu çalışmada sadece adılısıllar için yapılmış olan kısıtlamalarla ilgileneceğiz. Chomsky tarafından yapılan Bağlama Kuramının ilkelerinden olan bu kısıtlamaya göre; bir adıl, aynı yerel alan içindeki kendisini k-komutlandıran bir AÖ’e gönderimde bulunamaz. Diğer bir deyişle, bir adıl yerel alanı içinde özgür olmalıdır, yani gönderimi yerel alan dışında aranmalıdır. Buna göre aşağıdaki örneğe bir göz atalım.

Ahmet küçük bir fareyi eğitmiş. Fare onun kucağına yatmış.

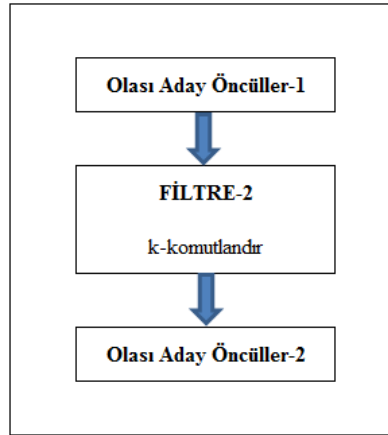
(4.6)

Söylem 4.6'nın ikinci tükmesindeki *onun* adılı için, Şekil 4.29 ile algoritmanın OAÖ'leri nasıl bulduđu gösterilmiştir.



Şekil 4.29 Söylem 4.6'nın Ağaç Yapısı ve 'onun' Adılı için Algoritmanın İşleyişi

Algoritma ikinci tükcede *onun* adılından önce geçen ve adılı k-komutlandırılan bir AÖ var mı diye araştırır ve *fare* ad öbeğinin *onun* adılı k-komutlandığını tespit eder ve bu AÖ'yü daha sonra eleme işleminde kullanmak üzere saklar. Algoritma tarafından bulunan OAÖ'ler programda önce Filtre-1'den geçirilir. Eğer OAÖ'lerin sayısı bire düşmediyse, program bu kısıtlamanın uygulandığı, Şekil 4.30'daki Filtre-2 ile OAÖ'leri sorgulama işlemine devam eder.



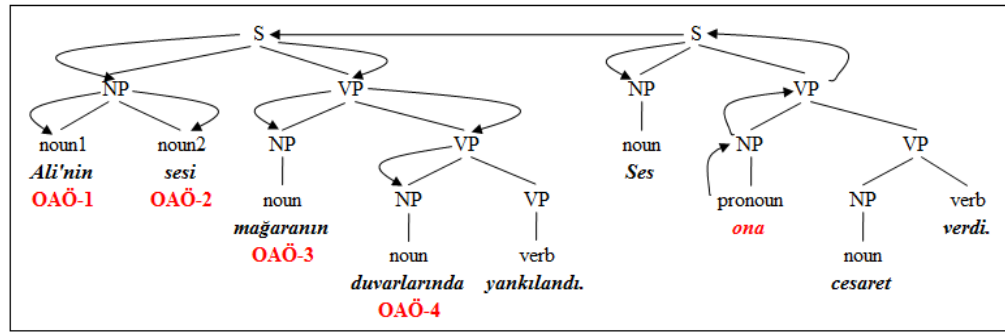
Şekil 4.30 Filtre-2

Şekil 4.30'da algoritmanın bulduđu OAÖ'ler Filtre-1'den geçirilir ve elenen olmaz. *Ahmet* ve *fare* OAÖ'leri Filtre-2'ye uygulanır. Bağlama Kuramı'nın ilkelerine göre yukarıda da belirtildiği gibi bir adıl kendisini k-komutlandırılan bir AÖ'e gönderimde

bulunamayacağı için, Filtre-1'den gelen *Ahmet* ve *fare* OAÖ'leri içinden *onun* adını k-komutlandırmış olan *fare* olası aday öncülü çıkartılır. Böylece *Ahmet* olası aday öncülü, *onun* adılıının öncülü olarak doğru bir şekilde bulunmuş olur.

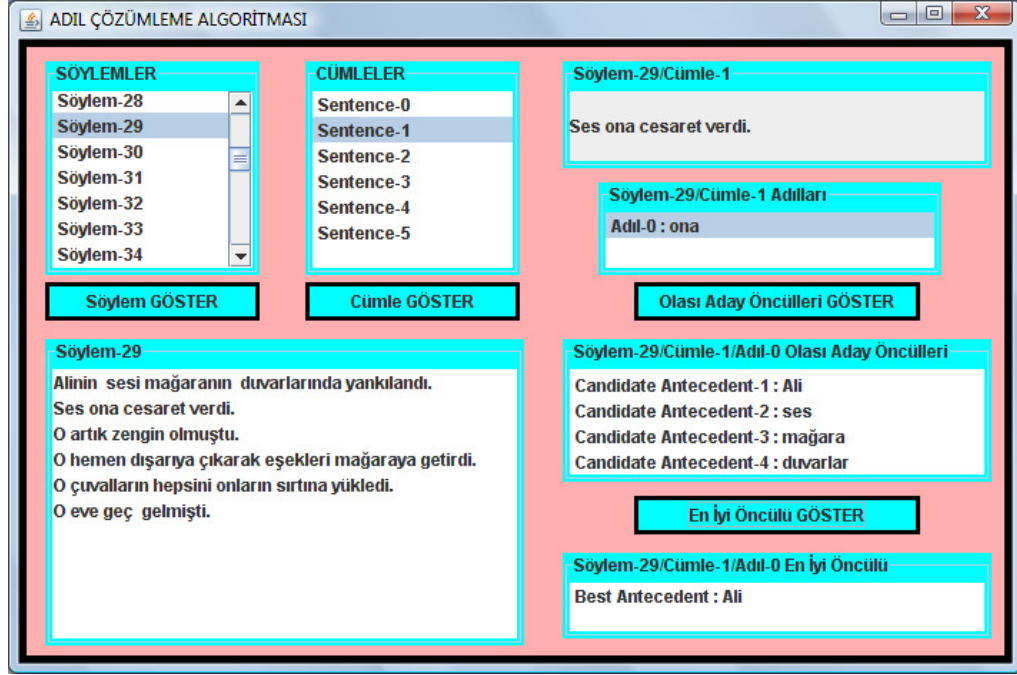
Ali'nin sesi mağaranın duvarlarında yankılandı. Ses ona cesaret verdi. (4.7)

Söylem 4.7'deki *ona* adılı için algoritma Şekil 4.31'deki gibi çalışır. İlk olarak *ona* adılı, algoritma tarafından tespit edilir ve adılıın içinde bulunduğu tümcede adıldan önce adılı k-komutlayan bir AÖ olup olmadığı araştırılır. Bunun sonucunda *ses* AÖ'nin *ona* adılıını k-komutlandığı tespit edilir. Daha sonra bir önceki tümcenin ağaç yapısına geçilerek, OAÖ'ler *Ali*, *ses*, *mağara* ve *duvarlar* olarak bulunur.



Şekil 4.31 Söylem 4.7'nin Ağaç Yapısı ve 'ona' Adılı için Algoritmanın İşleyişi

Daha sonra bu OAÖ'ler programda Filtre-1'den geçirilir. Filtre-1'de ilk olarak uygulanan seçmeli kısıtlama ile *mağara* ve *duvarlar* OAÖ'leri yer olma özelliği taşıdıkları için elenirler. Filtre-1 çıkışında OAÖ sayısı bire düşmediği için kalan *Ali* ve *ses* OAÖ'leri Filtre-2'ye gelirler. Filtre-2'de *ses* OAÖ'ü, *ona* adılıını kendi yerel alanı içinde k-komutlandığı için elenir ve *Ali* OAÖ'ü, *ona* adılıının öncülü olarak doğru bir şekilde bulunur. Bu sonuç, Şekil 4.32'de program çıktısı olarak da verilmiştir.

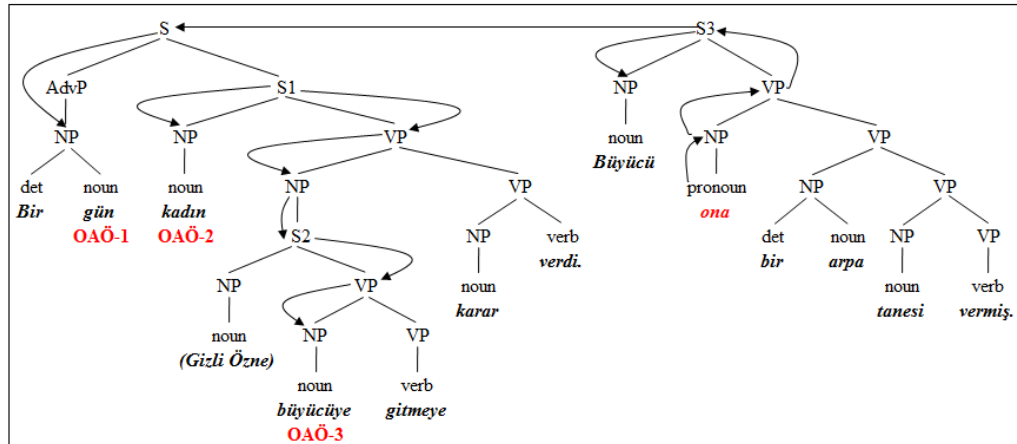


Şekil 4.32 Söylem 4.7'nin 'ona' Adılı için Program Çıktısı

Söylem 4.8'deki *ona* adılı için algoritma Şekil 4.33'deki gibi çalışır. İlk olarak *ona* adılı, algoritma tarafından tespit edilir ve adılın içinde bulunduğu tümcede adıldan önce adılı k-komutlayan bir AÖ olup olmadığı araştırılır. Bunun sonucunda *büyücü* AÖ'nin *ona* adılını k-komutlandığı tespit edilir. Daha sonra bir önceki tümcenin ağaç yapısına geçilerek, OAÖ'ler *gün*, *kadın* ve *büyücü* olarak bulunur.

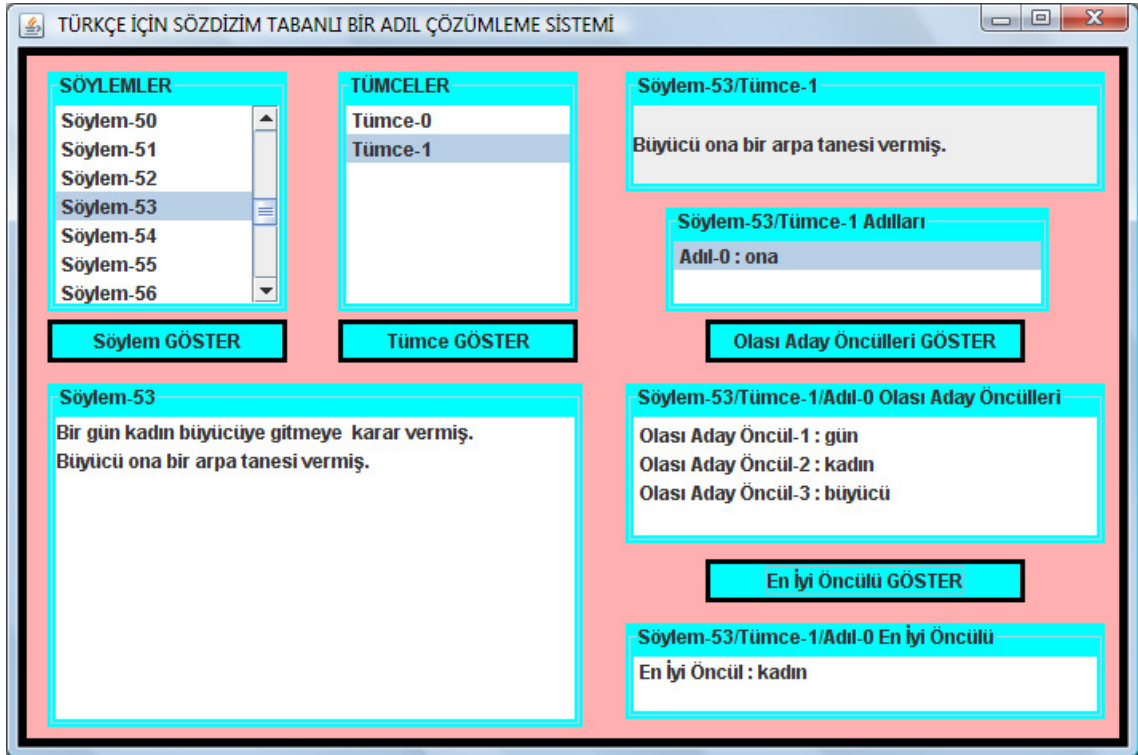
Bir gün kadın büyücüye gitmeye karar verdi. (4.8)

Büyücü ona bir arpa tanesi vermiş.



Şekil 4.33 Söylem 4.8'in Ağaç Yapısı ve 'ona' Adılı için Algoritmanın İşleyişi

Daha sonra bu OAÖ'ler programda Filtre-1'den geçirilir. Filtre-1'de ilk olarak uygulanan seçmeli kısıtlama ile *gün* OAÖ'ü zamanla ilgili bir sözcük olduğu için elenir. Filtre-1 çıkışında OAÖ sayısı bire düşmediği için kalan *kadın* ve *büyücü* OAÖ'leri Filtre-2'ye gelirler. Filtre-2'de *büyücü* OAÖ'ü, *ona* adını kendi yerel alanı içinde k-komutlandırdığı için elenir ve *kadın* OAÖ'ü, *ona* adının öncülü olarak doğru bir şekilde bulunur. Bu sonuç, Şekil 4.34'de program çıktısı olarak da verilmiştir.



Şekil 4.34 Söylem 4.8'in 'ona' Adılı için Program Çıktısı

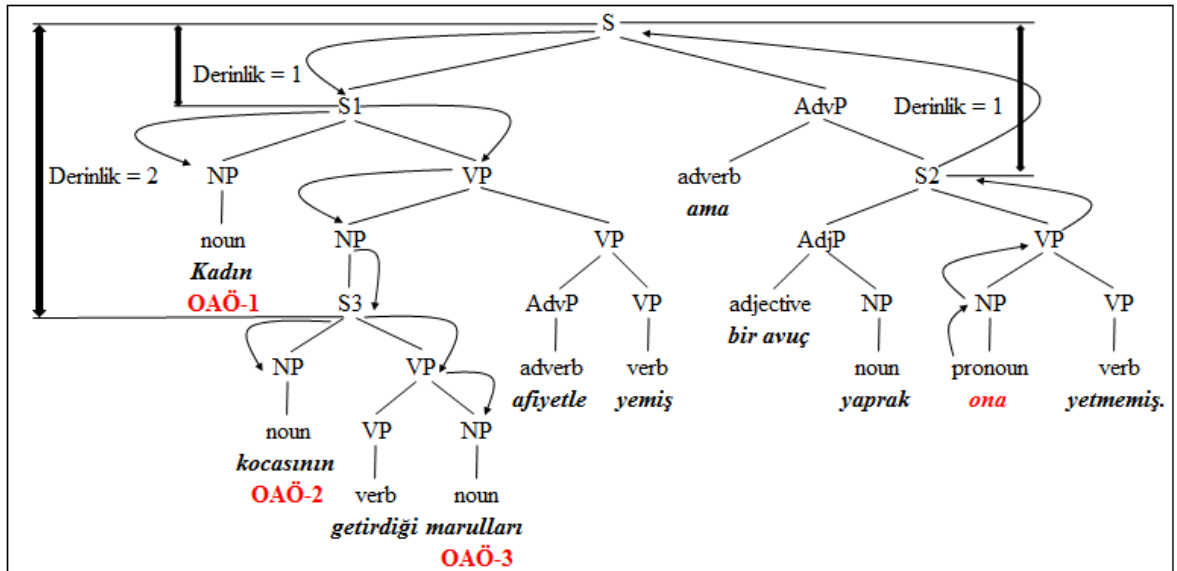
4.1.4 Tercihler

Tercihler, olası aday öncüllere uygulanan zorunlu olmayan kriterlerdir. Kısıtlar olası aday öncüllere uygulandıktan sonra, eğer olası aday öncül sayısı bire düşmüşse tercihlere gerek kalmaz ve uygulanmazlar. Eğer kısıtlar uygulandıktan sonra olası aday öncül sayısı birden fazla sayıda ise tercihler uygulanır. Bu çalışmada bizim belirlediğimiz Türkçe için olan tercihler sırasıyla aşağıda belirtilmiştir:

1. Derinlik Kontrolü Tercih: Bu aşamada tümcelerin ağaç yapıları üzerinden derinliğine, yani dikey yönde içiçe gömülü tümceciklerin sayısına bakılır. İçiçe gömülü tümceciklerden oluşan bir tümcede, derinlik algoritma açısından önemlidir. Adıl ile aynı derinlik bilgisine sahip AÖ'lerin en iyi öncül olma olasılığı, derinliği farklı olan AÖ'lere göre daha yüksektir. Bunu aşağıdaki örneği inceleyerek daha iyi anlayabiliriz.

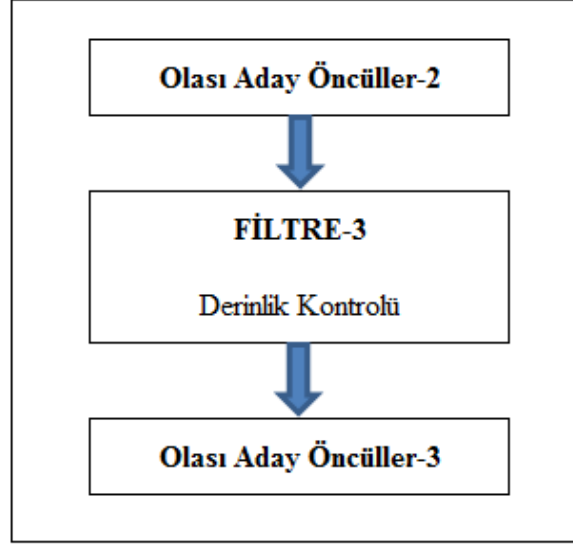
Kadın, kocasının getirdiği marulları afiyetle yemiş ama bir avuç yaprak ona yetmemiş. (4.9)

Söylem 4.9, Şekil 4.35'de de görüldüğü gibi yatay yönde yan yana S1 ve S2 diye iki tümcecikten oluşan bir tümce görünümündedir. Birinci tümcecik S1, dikey yönde gömülü, ağaç yapısında S3 diye adlandırılmış, bir tümcecik daha içerir. Bizi ilgilendiren dikey yöndeki tümceciklerin kaçınıcı tümcecik olduğu bilgisidir. Aynı derinlik bilgisine sahip, söz konusu adıldan farklı tümceciklerde yer alan OAÖ'lerin, adılın en iyi öncülü olma olasılığı daha yüksektir. Bu örnekte, *ona* adılının OAÖ'leri algoritma tarafından *kadın*, *kocasısı* ve *marulları* olarak tespit edilmiştir.



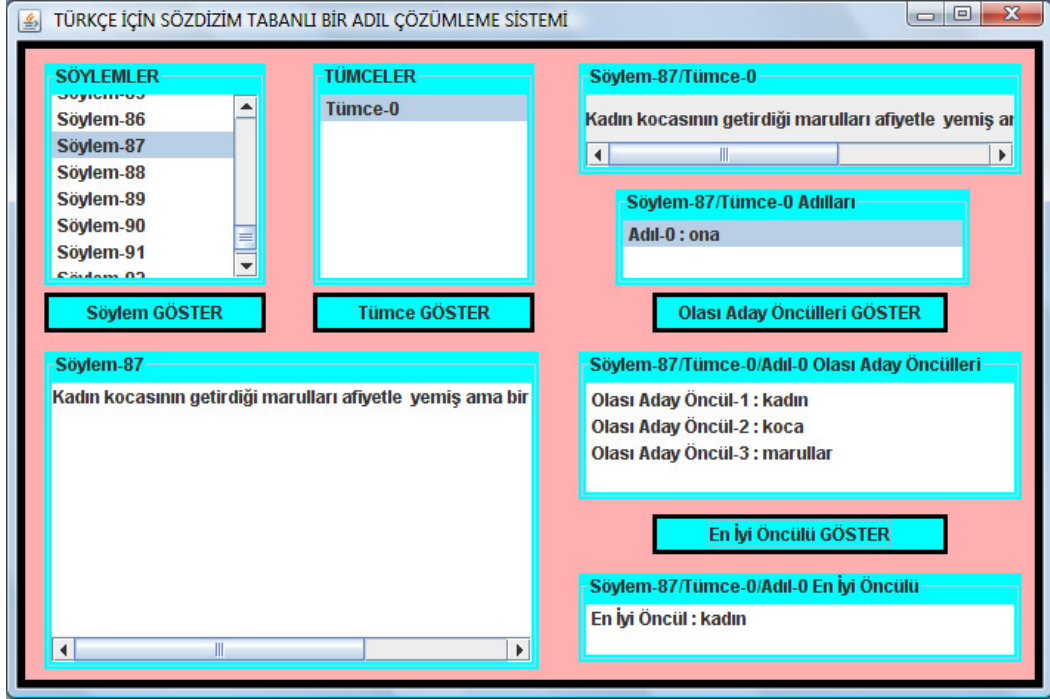
Şekil 4.35 Söylem 4.9'un Ağaç Yapısı ve 'ona' Adılı için Algoritmanın İşleyişi

Daha sonra tespit edilen bu OAÖ'ler, programda Filtre-1'den geçer ve *marullar* OAÖ'ü sayı bakımından *ona* adılı ile uyumlu olmadığı için elenir. Kalan OAÖ'lerin sayısına bakılır, bir olmadığı için kalan OAÖ'lere Filtre-2 uygulanır. Ancak, Filtre-2 çıkışında da OAÖ'lerin sayısında bir değişim olmaz. Bu aşamada, OAÖ'ler, derinlik kısıtlaması uygulayan Şekil 4.36'daki Filtre-3'e gelirler.



Şekil 4.36 Filtre-3

Filtre-3'de, *kadın* OAÖ'ü için dikey yöndeki derinlik indisinin bir olduğu, yani adılınkiyle aynı olduğu, *kocası* OAÖ'ü için derinlik indisinin ise ikiye eşit olduğu, yani adılınkinden farklı olduğu tespit edilir. Bu kısıtlama ile her bir OAÖ'nün derinlik indislerine bakılır ve adıl ile aynı derinlik indisine sahip olmayan *kocası* OAÖ'ü elenir. Böylece, dikeyde *ona* adılı ile aynı derinlik indisine sahip ama farklı tümcecikte yer alan *kadın* OAÖ'ü, *ona* adılının en iyi öncülü olarak doğru bir şekilde bulunur. Şekil 4.37 de bu sonuçlar program çıktısı olarak görülmektedir.

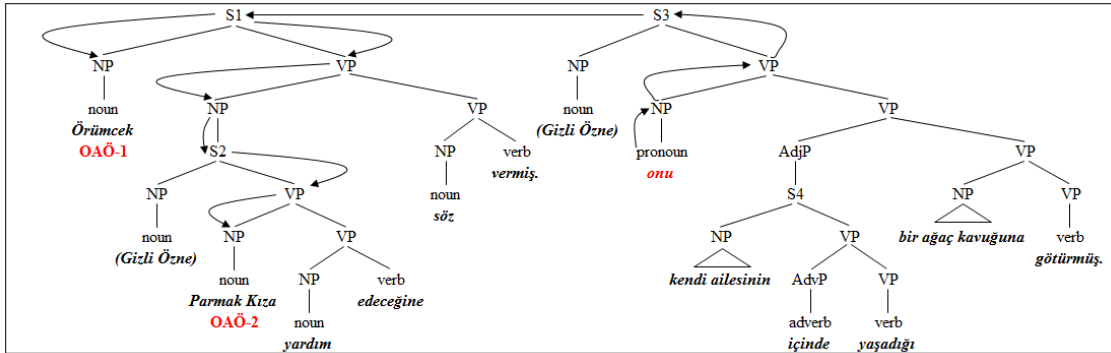


Şekil 4.37 Söylem 4.9'un 'ona' Adılı için Program Çıktısı

Şekil 4.38'de de görüldüğü gibi söylem 4.10'un ilk tümcesi, S1 ve S2 olmak üzere içiçe iki tümcecikten, söylemin ikinci tümcesi de S3 ve S4 diye adlandırılan içiçe iki tümcecikten meydana gelmektedir. İkinci tümcedeki *onu* adılıının OAÖ'leri algoritma tarafından *örümcek* ve *Parmak Kızı* olarak tespit edilmiştir.

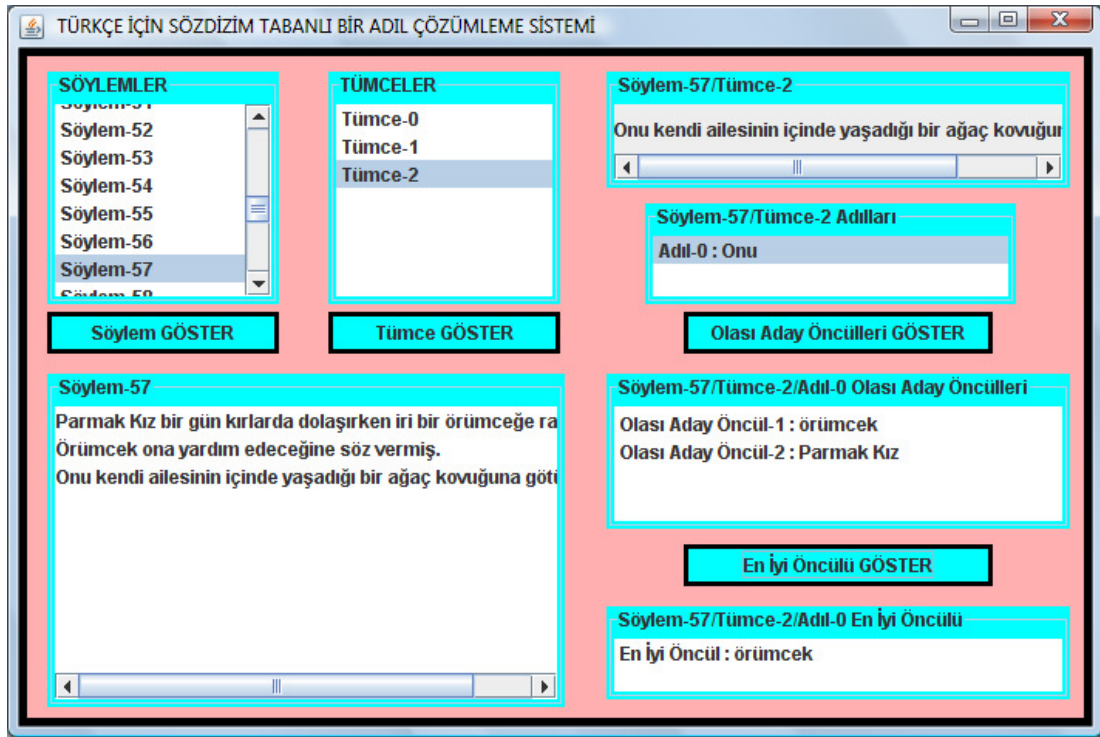
Örümcek, Parmak Kızıya yardım edeceğine söz vermiş. (4.10)

Onu kendi ailesinin yaşadığı bir ağaç kavuğuna götürmüş.



Şekil 4.38 Söylem 4.10'un Ağaç Yapısı ve 'onu' Adılı için Algoritmanın İşleyişi

Daha sonra tespit edilen bu OAÖ'ler, programda sırasıyla Filtre-1 ve Filtre-2'den geçerler ve elenen herhangi bir OAÖ olmaz. Derinlik kontrolü yapan Filtre-3'e gelindiğinde, OAÖ'lerin derinlik indisleri, adılın derinlik indisleriyle kıyaslanır. Bunun sonucunda, *onu* adılıyla aynı derinlik indisine sahip *örümcek* OAÖ'ü, en iyi öncül olarak seçilir. Program tarafından bu sonuç, Şekil 4.39'da da ekran çıktısı olarak görülmektedir.



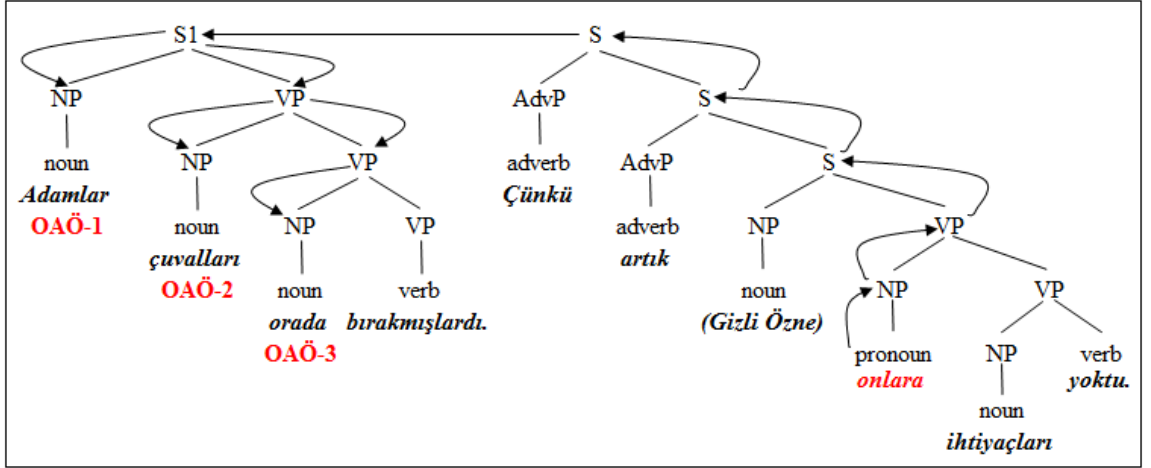
Şekil 4.39 Söylem 4.10'un 'onu' Adılı için Program Çıktısı

2. Sözdizimsel Paralellik Tercihi: Sözdizimsel paralellik bir tercihtir. Adılın sözdizimsel rolü (özne veya nesne) ile aynı sözdizimsel role sahip olan OAÖ'ler tercih edilir. Aynı sözdizimsel role sahip olmayan OAÖ'ler elenir.

Adamlar çuvalları orada bırakmışlardı. (4.11)

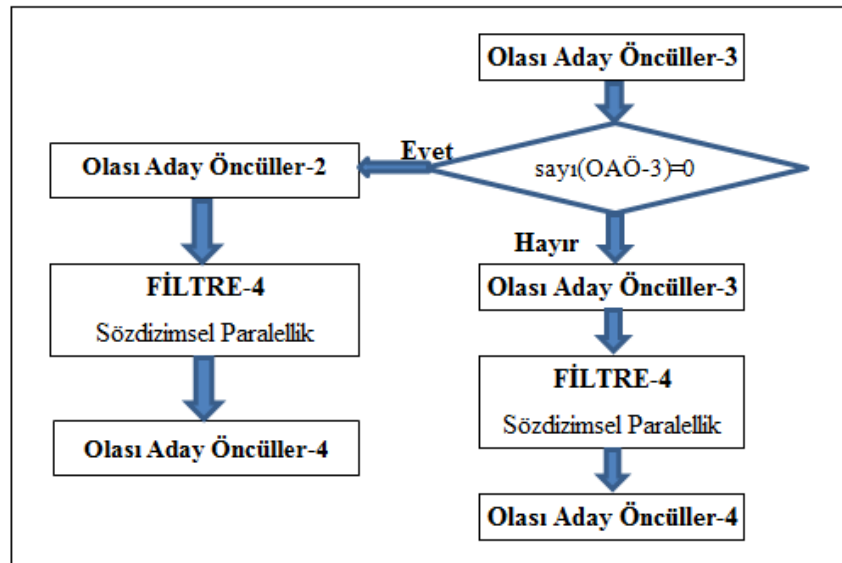
Çünkü artık onlara ihtiyaçları yoktu.

Söylem 4.11'deki ikinci tümcede yer alan *onlara* adılı çözümlenmek istendiğinde, Şekil 4.40'da görüldüğü gibi algoritma ağaç yapılarını tarayarak, *adamlar*, *çuvallar* ve *orada* AÖ'lerini OAÖ olarak bulur.



Şekil 4.40 Söylem 4.11'in Ağaç Yapısı ve 'onlara' Adılı için Algoritmanın İşleyişi

Algoritma tarafından bulunan OAÖ'lere ilk olarak programda Filtre-1 uygulanır ve seçmeli kısıtlamaya göre *orada* OAÖ'ü yer olma özelliği taşıdığı için elenir. Geriye kalan OAÖ'lerin her ikisi de sayı bakımından *onlara* adılı ile uyumludur. Filtre-1 çıkışında OAÖ sayısı bire düşmediği için, kalan OAÖ'lere Filtre-2 uygulanır. Filtre-2 çıkışında da elenen OAÖ olmamıştır. OAÖ'lere Filtre-3 uygulanır ve aynı derinlik indisine sahip olduklarından yine elenen OAÖ olmamıştır. Filtre-3 çıkışında OAÖ'lerin sayısı bir kez daha kontrol edilir. Bunun sonucuna göre, Şekil 4.41'de de görüldüğü gibi Filtre-4'e gönderilen OAÖ'ler değişir.



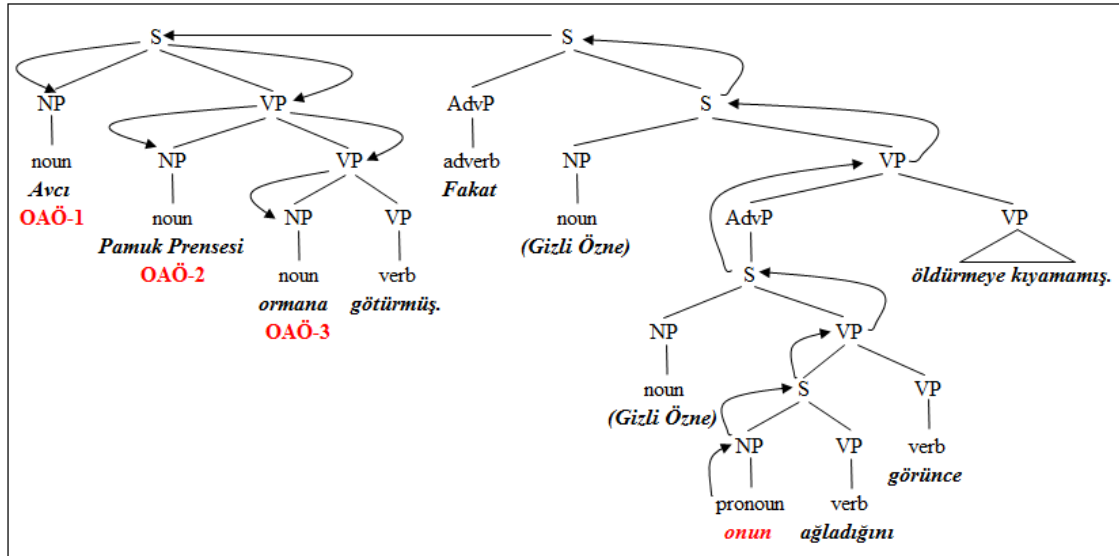
Şekil 4.41 Filtre-4

Bu durumda, sözdizimsel paralellığe bakılır. *Onlara* adının tümce içindeki görevi, nesne konumundadır. Olası aday öncüllerin tümce içindeki sözdizimsel rollerine bakıldığında, *adamlar* OAÖ'sü özne konumundayken, *çuvallar* OAÖ'sü nesne konumundadır. Buna göre, özne konumundaki *adamlar* OAÖ'sü elenir ve en son kalan *çuvallar* OAÖ'sü, *onlara* adının öncülü olarak seçilir.

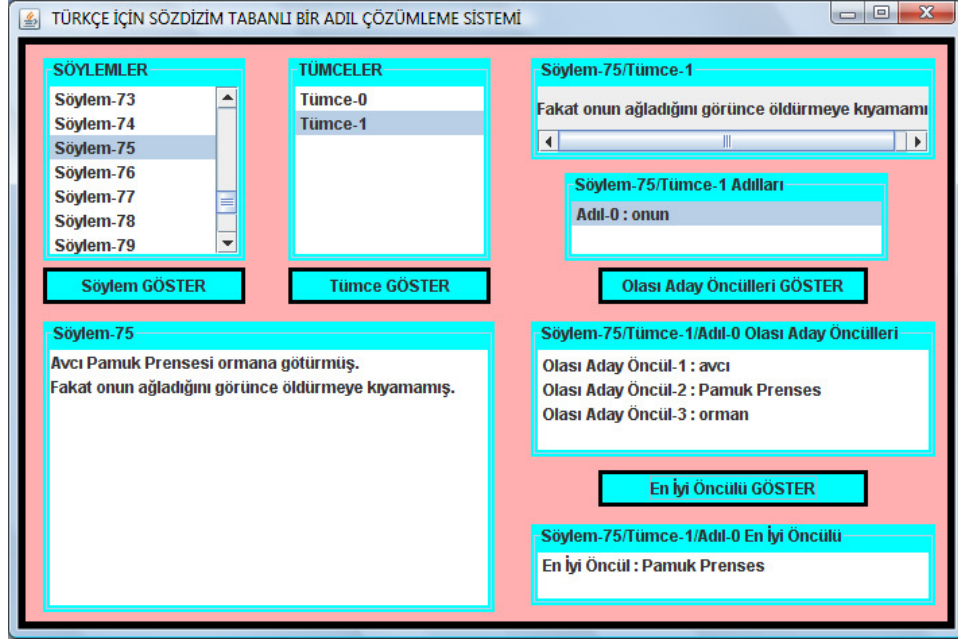
Söylem 4.12'nin ikinci tümcesindeki *onu* adlı için Şekil 4.42'deki gibi OAÖ'ler *avcı*, *Pamuk Prenses* ve *orman* olarak bulunur. Bu OAÖ'lere Filtre-1 uygulanır ve seçmeli kısıtlamaya göre *orman* OAÖ'ü elenir ve Filtre-1 çıkışında geriye kalan *avcı* ve *Pamuk Prenses* OAÖ'lerine Filtre-2 uygulanır. Filtre-2 sonucunda da elenen OAÖ olmamıştır. Aynı şekilde OAÖ'lere Filtre-3 uygulandığında da elenen OAÖ olmamıştır. Filtre-4'e gelindiğinde, *onun* adlı ile aynı sözdizimsel role sahip olan OAÖ, nesne konumundaki *Pamuk Prenses* OAÖ'dür. Böylece *onun* adının öncülü *Pamuk Prenses* olarak doğru bir şekilde bulunmuş olur. Bu sonuç, program çıktısı olarak Şekil 4.43'de görülmektedir.

Avcı Pamuk Prensesi ormana götürmüş. (4.12)

Fakat onun ağladığını görünce öldürmeye kıyamamış.

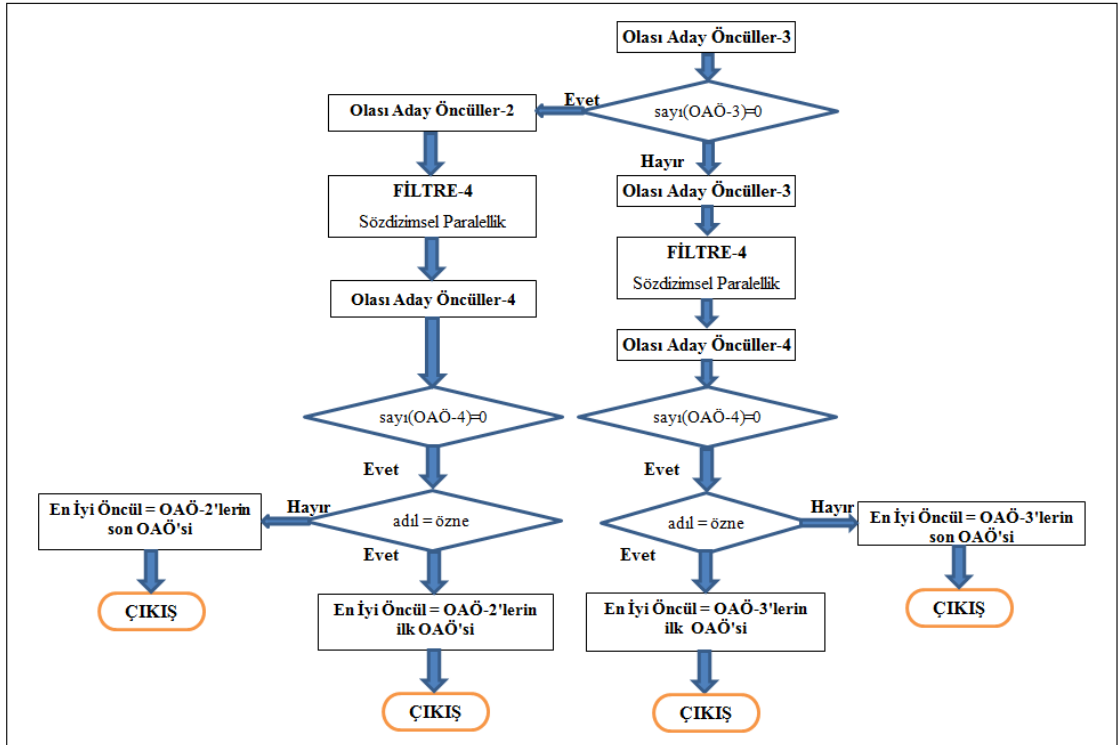


Şekil 4.42 Söylem 4.12'nin Ağaç Yapısı ve 'onun' Adılı için Algoritmanın İşleyişi



Şekil 4.43 Söylem 4.12'nin 'onun' Adılı için Program Çıktısı

Programda Filtre-4 ile sözdizimsel paralellik bakıldıktan sonra, çıkışında OAÖ'lerin sayısı kontrol edilir. Eğer bütün OAÖ'ler elenmişse, adılın sözdizimsel rolüne bakılır, özne ya da nesne olmasına göre, Şekil 4.44'de de görüldüğü gibi seçilen en iyi öncül farklılık gösterir.



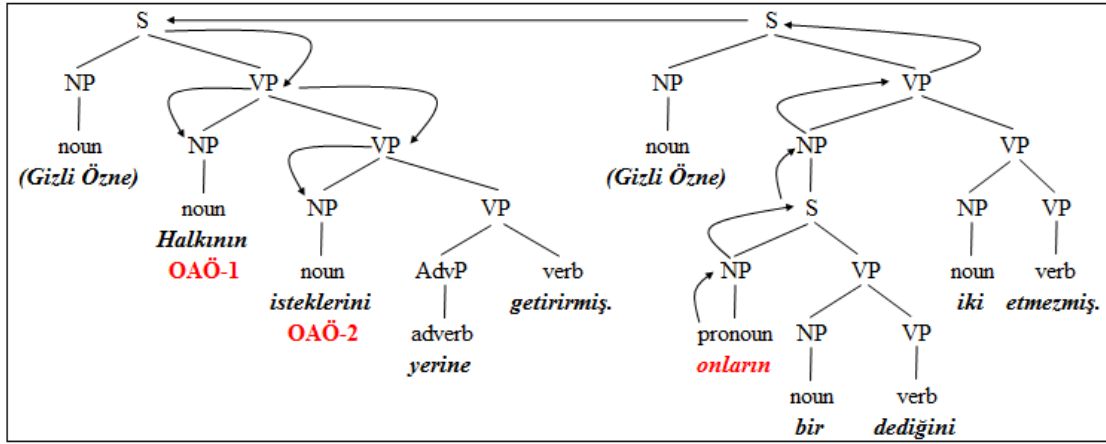
Şekil 4.44 Filtre-4 Çıkışı

Eğer adılın sözdizimsel rolü özne ise, Filtre-4 uygulanmadan önceki OAÖ'lerden en soldaki yani özne konumuna en yakın olan öncül, en iyi öncül olarak seçilir. Bunu bir örnekle açıklayalım.

Söylem 4.13'de ikinci tümcede geçen *onların* adılının Şekil 4.45'de görüldüğü gibi OAÖ'leri, *halk* ve *istekler* olarak bulunur.

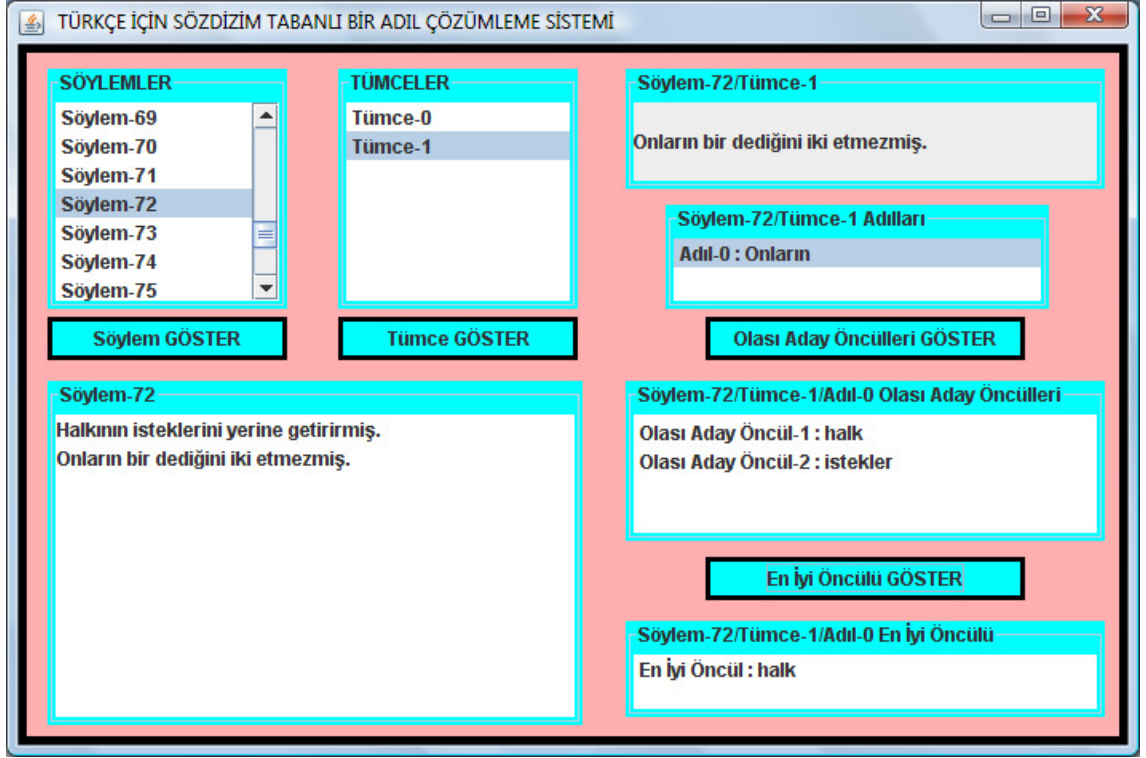
Halkın isteklerini yerine getirmiş. (4.13)

Onların bir dediğini iki etmezmiş.



Şekil 4.45 Söylem 4.13'ün Ağaç Yapısı ve 'onların' Adılı için Algoritmanın İşleyişi

Bu OAÖ'lere önce Filtre-1, daha sonra da Filtre-2 uygulanır. Her ikisinin çıkışında da elenen OAÖ olmamıştır. Derinlik kısıtlaması olan Filtre-3 uygulandığında ise her iki OAÖ'de elenecektir. Şekil 4.39'daki akış takip edilirse, her iki OAÖ de Filtre-4 tarafından sözdizimsel paralelliğe göre kontrol edilir ve OAÖ'rin her ikisi de elenir. Çünkü OAÖ'lerin sözdizimsel rolü nesne, adılın sözdizimsel rolü ise öznedir. Bu durumda, sözdizimsel rolü özne olan *onların* adılının en iyi öncülü, ağaç üzerinde en soldaki yani özne konumuna en yakın olan OAÖ yani *halk* seçilerek doğru bir şekilde bulunur ve bu sonuç program tarafından da Şekil 4.46'da olduğu gibi görsel olarak ifade edilmektedir.

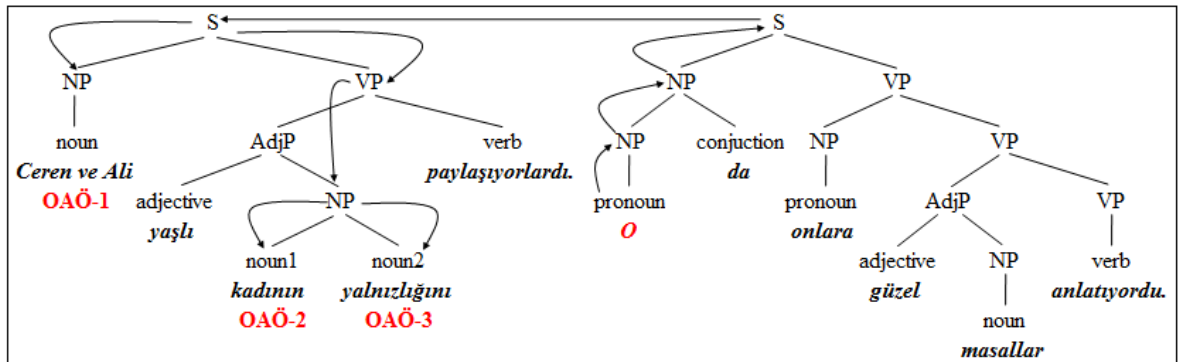


Şekil 4.46 Söylem 4.13'ün 'onların' Adılı için Program Çıktısı

Söylem 4.14'deki ilk adılı çözümlenmeye çalışırsak, Şekil 4.47'de görüldüğü gibi OAÖ'ler *Ceren ve Ali*, *kadın ve yalnızlık* olarak bulunur.

Ceren ve Ali yaşlı kadının yalnızlığını paylaşıyorlardı. (4.14)

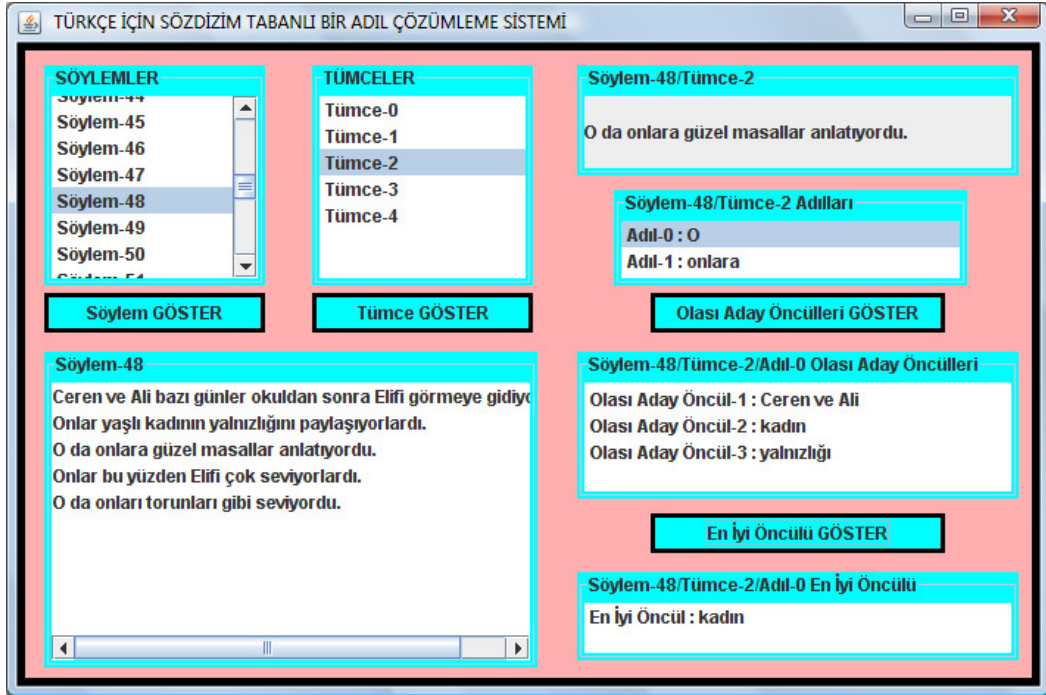
O da onlara güzel masallar anlatıyordu.



Şekil 4.47 Söylem 4.14'ün Ağaç Yapısı ve 'o' Adılı için Algoritmanın İşleyişi

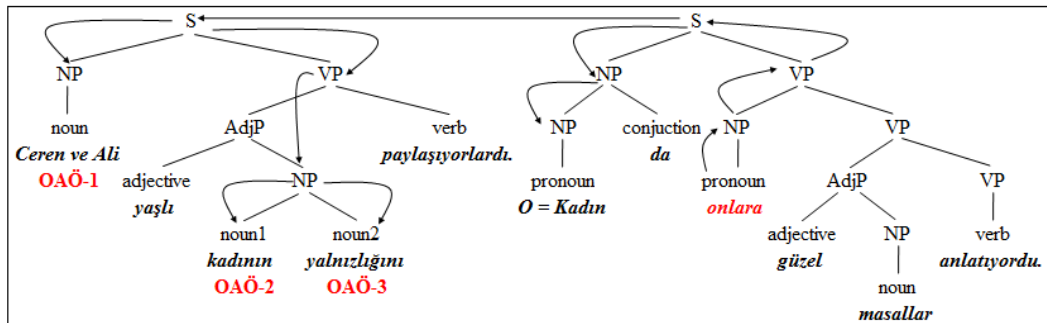
Bu OAÖ'lere ilk olarak program tarafından Filtre-1 uygulanır ve *Ceren ve Ali* sayı bakımından *o* adılı ile uyumlu olmadığı için elenir. Kalan OAÖ'lere Filtre-2 ve Filtre-3

sırasıyla uygulandığında elenen OAÖ olmaz. Filtre-4 yani sözdizimsel paralellik uygulandığında kalan OAÖ'lere, adılın özne konumunda olması ve OAÖ'lerin nesne konumunda olmasından dolayı her ikisi de elenir. Bu durumda, OAÖ'lerden en solda kalan, tümcenin öznesine en yakın konumda olan OAÖ, en iyi öncül olarak seçilir ve kadın OAÖ'sü *o* adılının en iyi öncülü olarak doğru bir şekilde bulunmuş olur. Şekil 4.48'de programın bu sonuçları ekran çıktısı olarak nasıl ürettiği gösterilmektedir.



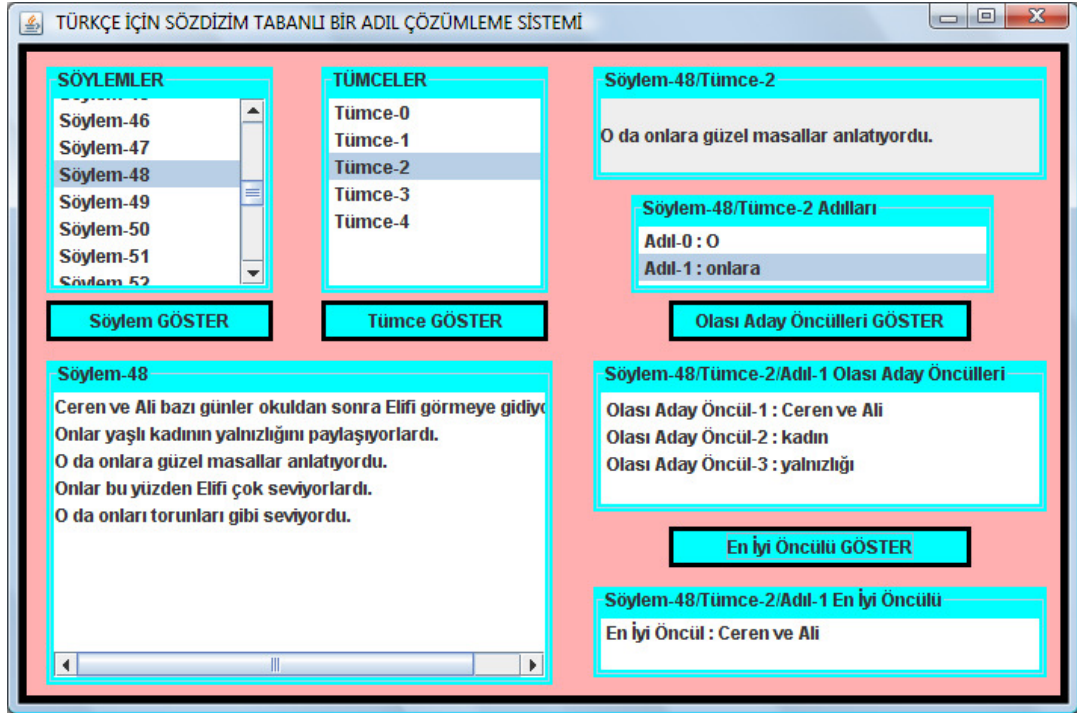
Şekil 4.48 Söylem 4.14'ün 'o' Adılı için Program Çıktısı

Çözümlenen *o* adılının yerine öncülü *kadın* konulup, şekil 4.49'daki ağaç üzerinde ikinci adıl araştırılır ve *onlara* adılı bulunup, bu adıl için çözümleme işlemi yeniden başlar. Yine aynı olası aday öncüller, *Ceren ve Ali*, *kadın* ve *yalnızlık* olarak bulunur.



Şekil 4.49 Söylem 4.14'ün Ağaç Yapısı ve 'onlara' Adılı için Algoritmanın İşleyişi

Daha sonra bu OAÖ'lere ilk olarak program tarafından Filtre-1 uygulanır. Filtre-1 sonunda kadın ve yalnızlık OAÖ'leri sayı bakımından onlara adılı ile uyumlu olmadığı için elenir. Geriye kalan *Ceren ve Ali* sayı bakımından *onlara* adılıyla uyumlu olduğu için ve kalan tek OAÖ olduğu için, onlara adılının en iyi öncülü olarak bulunmuş olur. Şekil 4.50'de de programın bu sonucunun ekran çıktısı olarak nasıl görüldüğü verilmiştir.

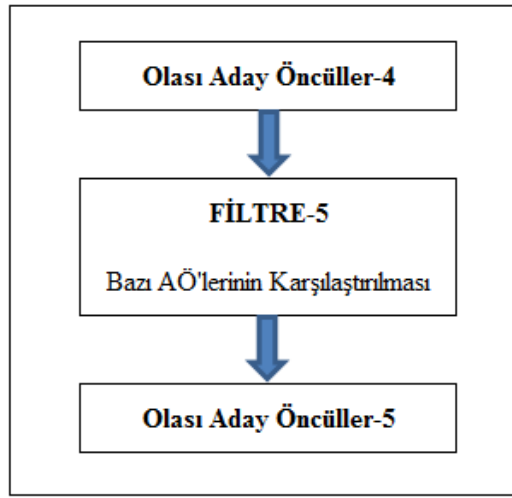


Şekil 4.50 Söylem 4.14'ün 'onlara' Adılı için Program Çıktısı

Eğer, adılın sözdizimsel rolü nesne ise, bu kriter uygulanmadan önceki OAÖ'lerden en sağdaki yani nesne konumuna en yakın olan öncül, en iyi öncül olarak seçilir.

3. Yakınlık Tercihi: Bu kriter gelinceye kadar, hâlâ olası aday öncül sayısı bire inememişse, sonuncu olan bu tercih uygulanarak en iyi öncül belirlenir. Bu tercih, olası aday öncüllerden adılı cümleye en yakın olan öncülün, en iyi öncül olarak seçilmesini sağlar. Literatürde adıl ile öncülü arasındaki mesafeye Hobbs mesafesi (Ge, Hale ve Charniak, 1998) denilir ve Hobbs mesafesi en kısa olan OAÖ'nün en iyi öncül olma olasılığı diğerlerine göre yüksektir.

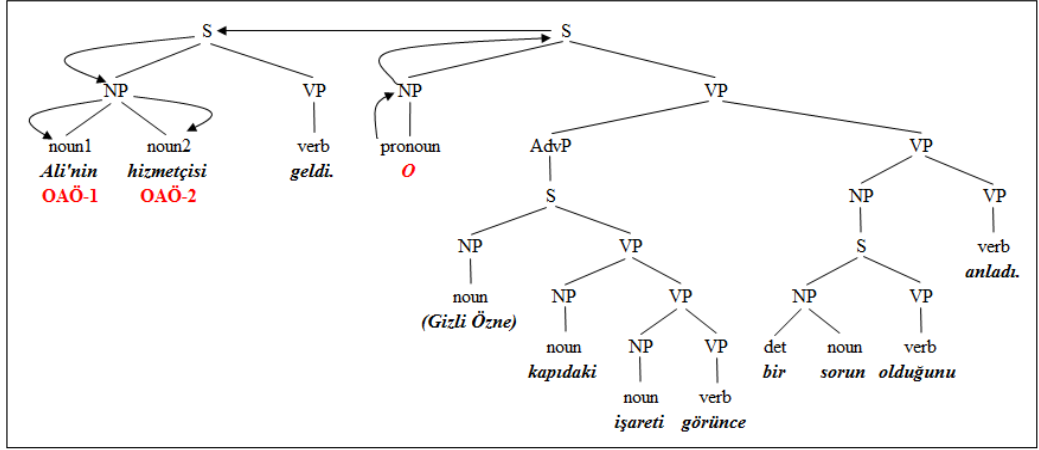
Programda, yakınlık tercihi uygulanmadan önce, sözdizimsel paralelliğe bakılan Filtre-4 çıkışından gelen OAÖ'lere Filtre-5 tarafından Şekil 4.51'de görüldüğü gibi bir ön işlem uygulanır. Filtre-5, önceden tanımlanmış bazı ad öbekleri ile Filtre-4 çıkışından gelen OAÖ'leri karşılaştırır. Önceden tanımlanmış ad öbekleri şunlardır: kez, defa, kere, sefer, her şey, hiç bir şey, hepsi, başkası. OAÖ'ler bu ad öbekleriyle karşılaştırılır ve eğer benzer ad öbekleri varsa OAÖ'ler içerisinde, bu OAÖ'ler elenir. Daha sonra kalan OAÖ'lere yakınlık tercihi uygulanarak adlı çözümlenmesi işleminin sonuna gelinmiş olur.



Şekil 4.51 Filtre-5

Söylem 4.15'in ikinci tümcesindeki *o* adlı için algoritma OAÖ'leri Şekil 4.52'de görüldüğü gibi *Ali* ve *hizmetçisi* olarak bulur.

Ali'nin hizmetçisi geldi. (4.15)
O kapıdaki işareti görünce bir sorun olduğunu anladı.



Şekil 4.52 Söylem 4.15'in Ağaç Yapısı ve 'o' Adılı için Algoritmanın İşleyişi

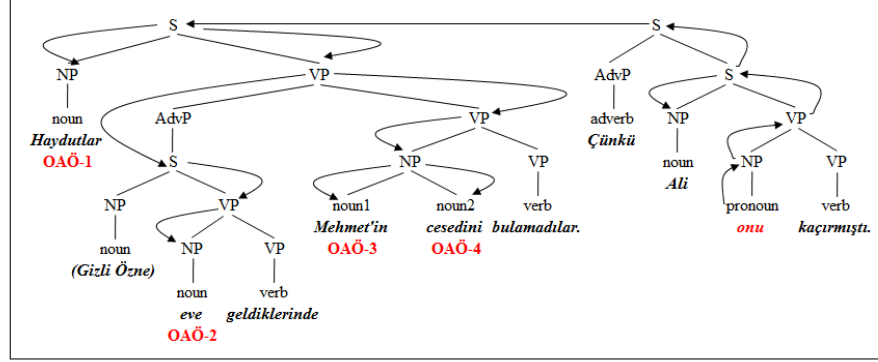
OAÖ'lere, program tarafından Filtre-1, Filtre-2 ve Filtre-3 uygulanır ve elenen herhangi bir OAÖ olmaz. Sözdizimsel paralelliğe bakan Filtre-4'e gelindiğinde, yine elenen OAÖ olmaz, her ikisi de *o* adılı gibi özne konumundadır. Bu durumda en son bulunan, ya da diğer bir deyişle adıla daha yakın olan OAÖ, yani *hizmetçisi*, *o* adılının en iyi öncülü olarak Şekil 4.53'de görüldüğü gibi seçilir ve doğru sonuç bulunmuş olur.

Şekil 4.53 Söylem 4.15'in 'o' Adılı için Program Çıktısı

Söylem 4.16'nın ikinci tûmcesindeki *onu* adlı için algoritma OAÖ'leri Şekil 4.54'de görüldüğü gibi *haydutlar*, *ev*, *Mehmet* ve *cesedi* olarak bulur.

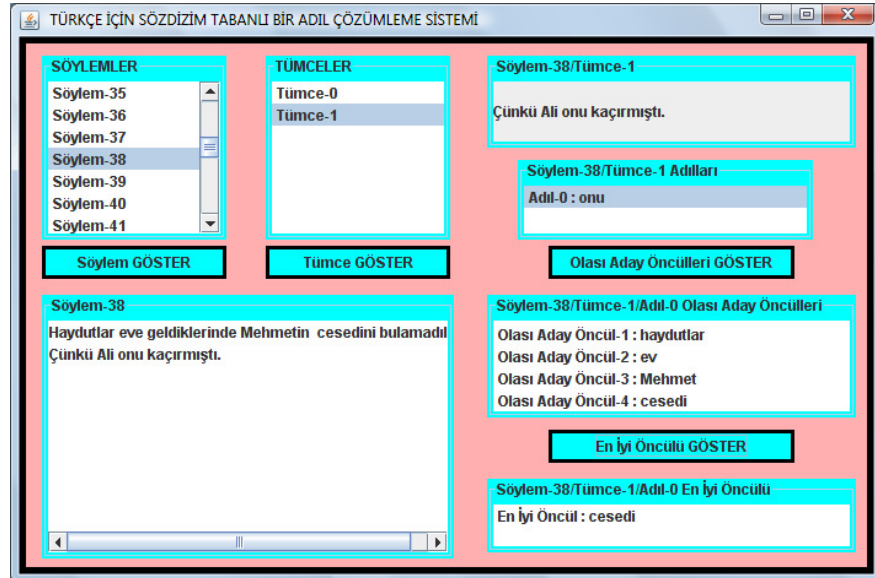
Haydutlar eve geldiklerinde Mehmet'in cesedini bulamadılar. (4.16)

Çünkü Ali onu kaçırmıştı.



Şekil 4.54 Söylem 4.16'nın Ağaç Yapısı ve 'onu' Adılı için Algoritmanın İşleyişi

Bulunan bu OAÖ'lere, önce program tarafından Filtre-1 uygulanır. OAÖ'lerden *eve* ve *haydutlar* elenir. Daha sonra Filtre-2 ve Filtre-3 uygulanır ve elenen herhangi bir OAÖ olmaz. Sözdizimsel paralelliğe bakan Filtre-4'e gelindiğinde, yine elenen OAÖ olmaz, her ikisi de *onu* adılı gibi nesne konumundadır. Bu durumda en son bulunan OAÖ, yani *cesedi*, *onu* adılıının en iyi öncülü olarak seçilir ve doğru bulunmuş olur. Programın ekran çıktısında da bu sonuç Şekil 4.55'deki gibi görülür.



Şekil 4.55 Söylem 4.16'nın 'onu' Adılı için Program Çıktısı

BÖLÜM 5

DEĞERLENDİRME

5.1 Performans Değerlendirme Ölçütleri

Doğal Dil Çözümleme çalışmalarının hepsinde olduğu gibi, sonuçların değerlendirilmesi artgönderim çözümlemesi alanı için de önemlidir. Fakat bu alanda geliştirilen sistemlerin performanslarının değerlendirilmesine yeterince önem verilmemiştir. MUC (Message Understanding Conference) toplantılarında bu soruna çözüm olarak, artgönderim çözümlemesi çalışmalarında anma ve kesinlik ölçütlerinin kullanılması tavsiye edilmiştir. Kullanımı uygun görülen bu ölçütler başarı oranını göstermede her ne kadar faydalı olsa da, bir artgönderim çözümleme sisteminin gerçek başarısı önceki çalışmaların sonuçlarıyla veya dayanaklarla karşılaştırılarak gösterilebilir. MUC tarafından tavsiye edilen bu ölçütler, ilk olarak istatistiksel başarımın hesaplanması için kullanılmışlardır. İki sınıfa ayrılmış bir veri kümesi içinde yapılan veri ayrıştırma işlemi sonucunda bir karışıklık matrisi (confusion matrix) ortaya çıkmaktadır (Kohavi ve Provost, 1998). Karışıklık matrisi içinde oluşan dört durum Tablo 5.1’de verilmektedir.

Tablo 5.1 Karışıklık Matrisi

		Tahmin Edilen Sınıf	
		Pozitif	Negatif
Gerçek Sınıf	Pozitif	DP	YN
	Negatif	YP	DN

Tasarlanan sistem, gerçekte doğru olanı doğru olarak tahmin etmesi durumunda DP (Doğru Pozitif), gerçekte doğru olanı yanlış olarak tahmin etmesi durumunda YN (Yanlış Negatif) ile ifade etmektedir. Ayrıca gerçekte yanlış olanın doğru olduğu ifade edilirse YP (Yanlış Pozitif) ve son olarak gerçekte yanlış olanın yanlış olduğu ifade edilirse DN (Doğru Negatif) ile tanımlanmaktadır.

Karışıklık matrisinden elde edilen bu değerlerle çeşitli ölçütler hesaplanmaktadır. Bunlardan *doğruluk* (accuracy) ölçütü tüm veri içinde doğru tahmin edilenlerin oranını ölçmeye imkân tanımaktadır. Bütün hata tiplerini dikkate alarak, pozitif ve negatif örnekleri aynı derecede önemsenmesini ve sınıflandırıcının toplam performansını değerlendirmeye yardımcı olmaktadır. Doğruluk ölçütü, veri kümesinde dengesiz dağılım var ise yeterli olamamaktadır. Bu durumda kullanılan *anma* (recall) ve *kesinlik* (precision) ölçütleri, sırasıyla, pozitif örneklerin negatif olarak sınıflandırılmasından oluşan hatalar ile negatif örneklerin pozitif olarak sınıflandırılmasından oluşan hataları belirtmektedirler. *F-ölçütü* (f-measure), anma ve kesinlik ölçütlerini, her ikisinin armonik ortalamasını olarak birleştirir. Doğruluk oranı, kesinlik, anma ve f-ölçütü aşağıdaki formüllerle hesaplanır (Kohavi ve Provost, 1998).

$$\text{Doğruluk Oranı (Accuracy)} = \frac{(DP + DN)}{(DP + DN + YN + YP)}$$

$$\text{Kesinlik (Precision)} = \frac{DP}{(DP + YP)}$$

$$\text{Anma (Recall)} = \frac{DP}{(DP + YN)}$$

$$\text{F-ölçütü} = \frac{(2 * \text{kesinlik} * \text{anma})}{(\text{kesinlik} + \text{anma})}$$

Karışıklık matrisi artgönderim çözümlemesi açısından yeniden düzenlenirse, DP değeri doğru tahmin edilen, YP değeri ise yanlış tahmin edilen artgönderimsel ilişkilerin sayısını ifade edecektir. Bu tezde verilen algoritmanın sonuçlarını değerlendirmekte kullanılan performans değerlendirme ölçütlerini bu değerlere uyarladığımızda aşağıdaki bağıntılar elde edilir.

$$\text{Doğruluk Oranı (Accuracy)} = \frac{\text{Sistem Tarafından Doğru Tahmin Edilmiş Tüm Durumlar}}{\text{Sistemdeki Tüm Durumlar}}$$

$$\text{Kesinlik (Precision)} = \frac{\text{Sistem Tarafından Doğru Şekilde Çözölmüş Artgönderim İlişkileri}}{\text{Sistem Tarafından İfade Edilmiş Toplam Artgönderim İlişkileri}}$$

$$\text{Anma (Recall)} = \frac{\text{Sistem Tarafından Doğru Şekilde Çözölmüş Artgönderim İlişkileri}}{\text{Derlemde Bulunan Toplam Artgönderimsel İlişkiler}}$$

5.2 Performans Sonuçları

Bu tez çalışmasının performansını değerlendirmek için sisteme iki dayanak (baseline) algoritması uygulanmıştır. Bu algoritmalar ile sistem için önerilen algoritma kıyaslanarak bir başarı oranı elde edilmektedir. Tablo 5.2'de sözdizim tabanlı bu sistem için oluşturulan 202 adıldan oluşan eğitim bütüncesine Dayanak-1 ve Dayanak-2 algoritmaları ile sistemin tamamının uygulanması sonucu alınan anma ve kesinlik oranları gösterilmektedir.

Tablo 5.2 Sözdizim Tabanlı Sistemin Eğitim Bütüncesi için Başarı Oranları

	DOĞRU	YANLIŞ	BELİRSİZ	Anma Oranı		Kesinlik Oranı	
Dayanak-1 (Algoritma + Kısıtlar)	110	3	89	110/202	54,46%	110/113	97,35%
Dayanak-2 (Algoritma + Kısıtlar + Yakınlık Tercihi)	140	62	0	140/202	69,31%	140/202	69,31%
Sözdizim Tabanlı Sistem (Algoritma + Kısıtlar + Tercihler)	171	31	0	171/202	84,65%	171/202	84,65%

Dayanak-1 algoritması, Hobbs'un Türkçe için yeniden uyarlanan algoritmasının bulunduğu olası aday öncüllere, sadece kısıtların uygulanmasıyla adılın en iyi öncülünü bulan algoritmadır. Olası aday öncüllere uygulanan bu kısıtlar, Bölüm 4.1.3'de açıklanmıştır. Bu kısıtlar olası aday öncüllerin yer ve zamanla ilgili semantik özelliklerini kontrol eden seçmeli kısıtlama, adıl ile olası aday öncüllerinin birbirleriyle sayı uyumluluğunu kontrol eden kısıtlama ve sözdizimsel Bağlama Kuramı kısıtlamalarını içerir. Programda bu kısıtların işleyişi Filtre-1 ve Filtre-2 sayesinde gerçekleştirilmektedir. Her filtre çıkışında olası aday öncül sayısı kontrol edilerek, eğer olası aday öncül sayısı bir ise ilgili adılın öncülü kalan bu olası aday öncül olarak kabul edilmektedir. Çoğunlukla Filtre-2 çıkışında bir tane olası aday öncül kalmaz. Dayanak-1 algoritmasını, olası aday öncüllere uyguladığımızda, son kısıtlamadan sonra olası aday öncüllerin sayısının her zaman bir olmadığı Tablo 5.2'de belirsiz sütunu altında görülmektedir.

Dayanak-1 algoritmasına uygulanan 202 adıldan, 113 tanesinin olası aday öncülleri sayısı, kısıtlar uygulandıktan sonra bire düşerken, 89 adıl için kısıtlar uygulandıktan sonra olası aday sayısı hala birden fazla olmaktadır. Dayanak-1 algoritması için anma ve kesinlik oranlarına bakıldığında 202 adıldan 110 tanesi doğru, 3 tanesi yanlış çözümlendiği ve 89 belirsiz durum olduğu görülmüştür. Sonuçta, anma oranı %54.46 ve kesinlik oranı %97.35 olarak bulunmuştur.

Dayanak-2 algoritması, olası aday öncüllere Dayanak-1 algoritması gibi ilk olarak kısıtları uygular ve olası aday öncül sayısının bire düşmediği durumlar için adıla en yakın olan olası aday öncülü seçen yakınlık tercihinin kullanarak adılı çözümler. Bölüm 4.1.4'de açıklanan yakınlık tercihinin Dayanak-2 algoritmasında kullanılması ile Dayanak-1'de oluşan belirsiz durumlar ortadan kalkar. Bu durum bulunan anma ve kesinlik oranlarının artmasına ve daha başarılı sonuçların alınmasına sebep olmaktadır. Dayanak-2 algoritmasına uygulanan 202 adıldan, 140 tanesinin doğru bir şekilde öncülü bulunurken 62 tanesinin öncülü yanlış bulunmuştur. Bu durumda alınan anma ve kesinlik oranları %69.31'dir. Sonuç olarak belirsiz durum olmadığı için her iki ölçütte sonuç olarak aynı değer bulunmuştur.

Sözdizim tabanlı sistemin eğitim bütüncesi, kısıtlar ve tercihlerden oluşan sisteme uygulandığında Tablo 5.2’de görüldüğü gibi 202 adıldan 171’i doğru bir şekilde çözümlenirken, 31 adıl yanlış çözümlenmiştir. Bu durumda anma ve kesinlik oranları için alınan performans sonuçları %84.65’dir. Görüldüğü gibi en iyi başarı oranları tüm kısıt ve tercihlerin olası aday öncüllere uygulanması ile alınmaktadır.

Sistemin performansını değerlendirmenin bir diğer yolu da modellenen sistemin başka bir sistemle karşılaştırılmasıdır. Bu tez kapsamında sunulan yaklaşımların sonuçları bu alanda yapılmış diğer çalışmalarla karşılaştırılabilmektedir. Ancak sağlıklı bir karşılaştırmanın önünde bir takım problemler vardır. Bu problemler şu şekilde aşağıda özetlenmiştir.

En çok görülen problem, karşılaştırılan iki sistemin farklı diller için yazılmış olmasıdır. Örneğin, İngilizce için yapılan bir çalışmayı Japonca için yapılmış başka bir çalışmayla birebir karşılaştırmak anlamsız olabilir. Bu çalışmada Hobbs’un İngilizce için yaptığı adıl çözümlenme algoritması, Türkçe için yeniden uyarlanmıştır. Bunun için sisteme Türkçenin dilbilgisel kuralları tanımlanarak, ağaç yapılarının belirlenen bu kurallara göre oluşturulması sağlanmıştır. Ağaç yapıları üzerinde adıl çözümlenmesinde kullanılacak olası aday öncüllerin bulunması yine bu kurallara göre yapılmıştır. Ancak yapılan bu değişiklikler sonucunda Hobbs’un İngilizce için yaptığı adıl çözümlenme sistemine benzer bir sistem modellenebilmiştir. Türkçe için modellenen bu sistemin, yine Türkçe için adıl çözümlenmesi yapabilen bir sistemle kıyaslanması daha anlamlı sonuçların alınmasını sağlamaktadır. Bu amaçla yapılan [Tüfekci, vd., 2007] çalışmasında, Türkçe için sözdizim tabanlı adıl çözümlenmesi yapabilen ve bu tezin kapsamını oluşturan sistem ile Türkçe için adıl çözümlenmesi yapabilen az bilgili sistemin (Küçük ve Turhan-Yöndem, 2007) karşılaştırılması yapılmıştır.

Karşılaştırılan iki sistem, aynı dil için oluşturulmuş olsa bile farklı bütünceler kullanıyor olabilmektedir. Her bütüncenin kendine göre bir karmaşıklığı olduğundan ortaya çıkan başarı oranlarının birbirleriyle karşılaştırılması gerekmemektedir. Bu yüzden bütüncelerin benzer hale getirilmesi, yapılacak karşılaştırmanın daha gerçekçi sonuçlar oluşturmasını sağlayacaktır. Tablo 5.3’deki 1 nolu bütünce, bu çalışma için

hazırlanmış olan Sözdizim Tabanlı Sistemin Eğitim Bütüncesidir. Diğer 2 nolu bütünce ise Tablo 5.3'deki az bilgili bir sistem için hazırlanmış olan Az Bilgili Sistemin Eğitim Bütüncesidir. Bütünce-1'de 202 adet adıl için çözümleme işlemi yapılmaktadır. Bu adillara özel isime gönderimde bulunan ya da bulunmayan diye herhangi bir ayırım uygulanmamıştır. Ancak az bilgili sistemde çözümlenen adillar için böyle bir ayırım söz konusudur. [Küçük ve Turhan-Yöndem, 2007] sadece özel isme gönderimde bulunan adillar için çözümleme yapabilen az bilgili bir sistem modellemiştir. İki farklı sistemin kıyaslanabilmesi için bütüncelerinin aynı ön işlemlerden geçirilip, benzer hale getirilmeleri gerekmektedir. Bu amaçla sözdizim tabanlı sistem ile az bilgili sisteme gerçekçi bir kıyaslama yapabilmek için, bütünce-1, bütünce-2'ye benzer hale getirilerek bütünce-1'deki adillar, 89 adet özel isme gönderimde bulunan ve 113 adet de özel isim dışındaki isimlere gönderimde bulunan adıl diye sınıflandırılmıştır. Bütünce-2 ise 190 adet özel isme gönderimde bulunan adıl ile 15 adet dönüşlü adıldan oluşmaktadır.

Tablo 5.3 Sistemde Kullanılan Bütünceler

Bütünce	Sistem Bütünceleri	Adillar (Özel İsme Gönderimde Bulunan)		Adillar (Özel İsim Dışındaki İsimlere Gönderimde Bulunan)		Toplam
		Adıl	Dönüşlü	Adıl	Dönüşlü	
1	Sözdizimsel Tabanlı Sistemin Eğitim Bütüncesi	89	0	113	0	202
2	Az Bilgili Sistemin Eğitim Bütüncesi	190	15	0	0	205

Az bilgili sistem için yeniden düzenlenen bütünce-1'in, Dayanak-1 ve Dayanak-2 algoritmaları ile sözdizim tabanlı sistem ve az bilgili sisteme yeniden uygulanmasıyla alınan sonuçlar, anma ve kesinlik ölçütleriyle Tablo 5.4'de ifade edilmiştir. Buna göre sözdizim tabanlı sistem 89 adıldan 72'sini doğru ve 17'sini yanlış olarak çözümler ve sonuçta %80.90 anma ve kesinlik oran değerlerini bulur. Az bilgili sistem ise 89 adıldan 85'ini doğru, 2'sini yanlış ve 2'sini de belirsiz olarak çözümler ve sonuç olarak da %95.51 anma ve %97.70 kesinlik oran değerlerini elde eder.

Tablo 5.4 Sözdizim Tabanlı Sistemin Eğitim Bütüncesinin Yeniden Uyarlanmış Hali İçin Başarı Oranları

	DOĞRU	YANLIŞ	BELİRSİZ	Anma Oranı		Kesinlik Oranı	
Dayanak-1 (Algoritma + Kısıtlar)	46	2	41	46/89	51,69%	46/48	95,83%
Dayanak-2 (Algoritma + Kısıtlar + Yakınlık Tercihi)	55	34	0	46/89	61,80%	46/89	61,80%
Sözdizim Tabanlı Sistem (Algoritma + Kısıtlar + Tercihler)	72	17	0	72/89	80,90%	72/89	80,90%
Az Bilgili Sistem	85	2	2	85/89	95,51%	85/87	97,70%

İki farklı adlı çözümleme sisteminin kıyaslanabilmesi için öncelikle bütüncelerin uygun hale getirilmesi gerekmektedir. Tablo 5.5’deki sonuçlara göre bütünce-1’in sözdizimsel sisteme uygulanmasıyla alınan anma ve kesinlik değerleri, %80.90 iken az bilgili sisteme uygulandığında alınan anma değeri %95.51 ve kesinlik değeri de %97.70 olmaktadır. Bütünce-2’nin, sözdizimsel sisteme uygulanmasıyla alınan anma oranı %69.47 ve kesinlik oranı %78.57 olarak bulunurken, az bilgili sistem için bu değerler anma %74.21 ve kesinlik %90.38 olmaktadır. Ortalama değerlere bakıldığında ise sözdizim tabanlı adlı çözümleme sistemi için iki farklı bütüncenin uygulanması sonucu alınan anma değeri %75.19 ve kesinlik değeri de %79.74 olarak bulunmuştur. Az bilgili sisteme iki farklı bütüncenin uygulanması sonucunda alınan ortalama anma değeri %84.86 ve kesinlik değeri de %94.04 olmaktadır. Bu sonuçlar bize göstermektedir ki, bu çalışmada sunduğumuz sözdizim tabanlı adlı çözümleme sistemi ile Türkçe için yapılmış diğer çalışmalarla boy ölçüşebilecek kadar başarılı sonuçlar elde edilmiştir.

Tablo 5.5 Sözdizim Tabanlı Adlı Çözümleme Sistemi ile Az Bilgili Adlı Çözümleme Sisteminin Kıyaslanması

Bütünce	Sözdizim Tabanlı Adlı Çözümleme Sistemi				Az Bilgili Adlı Çözümleme Sistemi			
	Anma		Kesinlik		Anma		Kesinlik	
1	72/89	80,90%	72/89	80,90%	85/89	95,51%	85/87	97,70%
2	132/190	69,47%	132/168	78,57%	141/190	74,21%	141/156	90,38%
Ortalama	75,19%		79,74%		84,86%		94,04%	

Kullanılan bütüncenin büyüklüğü de iki sistemin karşılaştırılmasını etkileyecek bir durumdur. Literatürdeki kimi çalışmalar 200 adıl içerirken kimi çalışmalar 5.000 adıl içerebilmektedir. Bu anlamda bütüncenin büyüklüğü de yaratılan sistemin güvenilirliğini belirler. Bu çalışmada sunulan sistem iki ayrı bütüncede toplam $202 + 190 = 392$ adet adıl çözümlemesi işlemi yapabilmektedir. Bu işlemler ne kadar çok olursa sistem güvenilirliği de buna paralel olarak artar. Bu çalışma sözdizimsel anlamda Hobbs'un Naif Yaklaşımı gibi literatürde çok önemli yere sahip bir yaklaşımın ilk kez Türkçe için oluşturulan bir modeli olup, bu kapsamda bu kapasitede bir bütünceyle ancak bir ilk olması açısından önemli bir çalışmadır. Bütünce kapasitesinin artırılması ve bu sayede bundan sonra gelecek çalışmalara bir temel olabilmesi açısından önemli bir çalışmadır.

İki farklı sistem karşılaştırılırken yaşanabilecek bir diğer problem de çözümlenecek artgönderimlerin kapsamıdır. Çünkü bazı çalışmalar sadece adılar üzerinde çalışırken bazı sistemler tüm gönderimsel ilişkiler üzerinde çalışabilmektedir. Türkçe için bunu açık ve gizli adılar olarak genişletmek de mümkündür. Bu tez kapsamında sadece açık adılar incelenmiştir. [Küçük, 2005] tarafından yapılan çalışmada da adılar ve dönüşlü artgönderimler çözümlenebilmekteyken, iki sistemin kıyaslanması söz konusu olduğunda dönüşlü artgönderimler başarı oranlarına dahil edilmemiştir.

Artgönderim çözümleme sistemlerindeki bir diğer önemli etken ise sistemin ne kadar otomatik bir şekilde çalıştığıdır. Çünkü gerçek ve güncel uygulamalar tam otomatik bir alanda çalışmak isterler. Artgönderim çalışmalarında bütünceyi elle işaretlemek sıklıkla başvurulmuş bir yöntemdir. Bu da sistemin tam otomatik olmamasına sebep olur. Bu çalışmada sistemde ağaç yapılarının oluşturulması için gerekli olan her bir sözcüğe ait bilgiler bütünceye elle girilmiştir. Karşılaştırılan sistemler açısından, elle girilen bilgilerin ne kadar olduğu da en az diğer unsurlar kadar önemlidir.

Az bilgili artgönderim çözümleme sistemleri son zamanlarda sıklıkla tercih edilmektedir. Sistemin az bilgiye ya da çok bilgiye dayanan bir sistem olması, onun performansını etkileyecektir. Bir sistem için ne kadar az bilgiye ihtiyaç duyulursa o

sistemin uygulanması çok daha kolay ve hızlı olur. Bu çalışmada da bu etken göz önünde bulundurularak sadece tümcelerın sözdizim özelliklerine dayanan, bunun dışında fazla bilgiye gereksinim duymadan adıl çözümlemesi yapabilen aslında bir az bilgili sistem tasarlanmıştır.

Bu tezde bahsedilen diğer artgönderim çalışmalarının sonuçları Tablo 5.6'da gösterilmiştir.

Tablo 5.6 Artgönderim Çalışmalarının Başarı Oranları

Çalışma	Dil	Yöntem	Sonuç
[Kennedy ve Boguraev, 1996]	İngilizce	Bilgi Tabanlı	% 75
[Baldwin, 1997]	İngilizce	Bilgi Tabanlı	% 92 kesinlik ve % 64 anma
[Mitkov, 2002]	İngilizce	Bilgi Tabanlı Yarı Otomatik	% 85 - 90
[Mitkov, 2002]	İngilizce	Bilgi Tabanlı Tam Otomatik	% 61.55
[Mitkov, 2002]	Japonca Adaptasyon	Bilgi Tabanlı Yarı Otomatik	% 75.8
[Mitkov, 2002]	Bulgarca Adaptasyon	Bilgi Tabanlı Yarı Otomatik	% 72.6
[Aone ve Bennet, 1995]	Japonca	Öğrenme Tabanlı	% 70.20, % 88.55, % 77.27 (anma, kesinlik, f-ölçütü)
[McCarthy ve Lehnert, 1995]	İngilizce	Öğrenme Tabanlı	% 85.4, % 87.6, % 86.5 (anma, kesinlik, f-ölçütü)
[Soon vd., 2001]	İngilizce	Öğrenme Tabanlı	% 52, % 68, % 60 (anma, kesinlik, f-ölçütü)
[Yüksel ve Bozşahin, 2002]	Türkçe	Bilgi Tabanlı	% 70
[Yıldırım, 2008]	Türkçe	Belirginlik Tabanlı	% 53
[Yıldırım, 2008]	Türkçe	Merkezeleme Yaklaşımı	% 71
[Yıldırım, 2008]	Türkçe	J48 - Karar Ağacı	% 69
[Yıldırım, 2008]	Türkçe	IBk (k=1)	% 75
[Küçük, 2005]	Türkçe	Bilgi Tabanlı	% 85, % 88, % 87 (anma, kesinlik, f-ölçütü)
Bizim Çalışma	Türkçe	Sözdizim Tabanlı Yaklaşım	% 84.65

Bu tabloda da görüldüğü gibi artgönderim problemi için bir çok dile yönelik çok sayıda çalışma yapılmıştır. Bu tablo yapılan çalışmaların bir kısmını temsil etmektedir. Ancak bunlar literatürdeki en önemli çalışmalar arasından seçilmiştir. Çalışmaların başarı oranları ve f-ölçütleri %60 ile %90 arasında değişmektedir. Bu sonuçlar mükemmel bir sistem için yeterli değildir. Dolayısıyla artgönderim çözümlemesi için daha çok çalışma yapılması gerekmektedir. Her ne kadar yukarıda sıralanan sebeplerden dolayı birebir karşılaştırma anlamsız olsa da çalışma sonuçlarının diğer çalışma sonuçlarına yakın olması hatta bu çalışmada aldığımız oranların diğer çalışmalarla boy ölçüşebilecek seviyede olması bu çalışmada tasarlanan Türkçe için sözdizim tabanlı adıl çözümleme sisteminin bu anlamda yeterli olduğu söylenebilir. Ama yine de Türkçe için daha bir çok çalışma ve deney yapılması gerekmektedir. Bu çalışma ve bahsedilen diğer çalışmalar bize gelecek için bir takım ipuçları vermektedir. Bu mana da değerlendirme yapmak faydalı olacaktır.

BÖLÜM 6

SONUÇ

Bu tez çalışması kapsamında Türkçe metinlerdeki artgönderimsel ilişkilerin bilgisayarlı çözümlenmesi araştırılmıştır. Artgönderim çözümlenmesi bilgisayarlı dilbilim araştırma alanının önemli bir alt başlığı olduğundan bir çok sistemin vazgeçilmez unsurudur. Türkçe için yürütülen bilgisayarlı dilbilim çalışmalarının azlığı ve bu alanlardaki çalışmalara duyulan ihtiyaç sebebiyle Türkçe için artgönderim çözümlenmesi bu tez kapsamında araştırılmıştır. Bu amaçla Türkçe'deki üçüncü kişi adılarını çözümlen sözdizim tabanlı bir adıl çözümlenme sistemi bilgisayarlı ortamda tasarlanmıştır. Kapsam olarak sadece adısal artgönderimler ele alınmıştır. Deneyler yaklaşık 12.940 sözcük ve 392 adıl içeren iki farklı bütünce üzerinde denenmiştir. Deney sonuçlarını değerlendirebilmek için bütüncedeki ilişkiler elle işaretlenmiştir.

Bu çalışmada izlenen yöntem, Hobbs'un Naif Yaklaşımına ve Chomsky'nin Bağlama Kuramına dayanan sözdizim tabanlı bir yaklaşımdır. Türkçede adıl çözümlenmesi yapmak için izlenen bu yöntem, şimdiye kadar Türkçe için yapılmış artgönderimsel çalışmalar içinde sözdizim tabanlı adıl çözümlenmesi yapabilen ilk çalışmadır. Sonuçlar göstermektedir ki, bu yöntem Türkçe için yapılmış diğer çalışmalarla boy ölçüşebilecek derecede başarılı bir yöntemdir.

Bu tez kapsamında tasarlanan sözdizim tabanlı adıl çözümleme sistemi, Dayanak-1 ve Dayanak-2 algoritmaları ile Türkçe için adıl çözümleme yapan bir diğer sistem olan az bilgili sistem (Küçük ve Turhan-Yöndem, 2007) ile karşılaştırılmıştır.

Sözdizim tabanlı sistem, Dayanak-1 ve Dayanak-2 ile kıyaslama sonucunda alınan başarı oranlarının %51.69 ve %61.80 olduğu görülmüştür. Sadece bu oranlar bile bu alanda yapılan diğer çalışmalarla boy ölçüşebilecek seviyededir. Sözdizim tabanlı ve az bilgili sisteme uygulanan bütüncelerin benzer hale getirilmesi sonucu toplamda 279 adıl her iki sisteme de öncüllerin çözümlenmesi üzere uygulanmıştır. Sözdizim tabanlı sistem, 279 adıldan 204'ünü doğru, 53'ünü yanlış ve 22'sini de belirsiz olarak çözümlenmektedir. Bu çözümleme sonucunda alınan ortalama performans değerleri sözdizim tabanlı sistem için, %75.19 anma ve %79.74 kesinlik oranları olarak bulunmuştur. Az bilgili sistem ise 279 adıldan 226'sını doğru, 19'unu yanlış ve 34'ünü de belirsiz olarak çözümlenmektedir. Bunun sonucunda alınan performans değerleri ise %84.86 anma ve %94.04 kesinlik oranları olarak bulunmuştur. Bu kıyaslama sonucunda bu çalışmada önerilen sistem açısından tatmin edici sonuçlara ulaşılmıştır.

Gelecekteki çalışmalara temel oluşturma açısından bu çalışma sözdizimsel özellikleri kullanan ve bunu yaparken de başka bir kaynağa ihtiyaç duymayan basit ve az bilgi kullanan bir sistemdir. İleride bu sistemi kullanacak çalışmalarda bütüncenin kapasitesinin genişletilmesi, sistemin güvenilirliğini arttıracaktır. Sistemin diğer yapılan çalışmalarla kıyaslanabilmesi için, tüm sistemler tarafından kullanılabilirliği ve erişimi kolay olan ortak bir veri tabanına ihtiyaç vardır. Ayrıca bu sistemde adılların öncülleri aranırken en fazla bir önceki tümceye kadar gidilebilmektedir. Az bilgili sistemde, ilgili adılın içinde bulunduğu tümceden en fazla 3 tümce öncesine kadar gidilip, olası aday öncül araştırması yapılabilinmektedir. Bu da belirsiz durumların ortadan kalkmasına ve bu yolla başarı oranlarının daha da yükselmesine sebep olmaktadır. Gelecekte bu çalışma da geliştirilip olası aday öncülleri tararken kullandığı araştırma alanı 3 tümceye kadar çıkarılarak daha yüksek performans sağlanabilir.

KAYNAKÇA

1. [Aone ve Bennet, 1996] Applying Machine Learning to Anaphora Resolution. Connectionist, statistical and symbolic approaches to learning for Natural Language Processing, 302-314. Berlin: Springer.
2. [Ariel, 1988] Referring and Accessibility. Journal of Linguistics, 24: 65-67.
3. [Ariel, 1990] Accessing Noun-Phrase Antecedents. London: Routledge.
4. [Baldwin, 1997] CogNIAC: High precision coreference with limited knowledge and linguistic resources. In Proceedings of the ACL'97/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution, 38-45.
5. [Balçı, 2006] Metindilbilim Açısından Bir Çözümleme. Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü Dergisi Sayı: 21, 191-204 s.
6. [Banfield, 1982]. Unspeakable Sentences: Narration and Representation in the Language of Fiction. Routledge & Kegan Paul, London.
7. [Brennan vd., 1987] A Centering Approach to Pronouns. In the 25th Annual Meeting of the Association of Computational Linguistics. Stanford.
8. [Carbonell ve Brown, 1988] Anaphora Resolution: A Multi-strategy Approach. Proceedings of the 12. International Conference on Computational Linguistics (COLING'88), Vol.I, 96-101, Budapest, Hungary.
9. [Chomsky, 1981] Lectures on Government and Binding. Dordrecht: Foris.
10. [Chomsky, 1982] Some Concepts and Consequences of the Theory of Government and Binding. Massachusetts: MIT Press.
11. [Chomsky, 1986] Knowledge of Language: Its Nature, Origin and Use. New York: Praeger.
12. [Chomsky, 1995] The Minimalist Program. Massachusetts: MIT Press.
13. [Cook ve Newson, 1996] Chomsky's Universal Grammar: An Introduction. Blackwell Publishing.
14. [Çeltek, 2003] Türkçe Sözlü Söylem Yapısı ve Artgönderim. Yüksek Lisans Tezi. Sosyal Bilimler Enstitüsü, Dokuz Eylül Üniversitesi.
15. [Çeltek ve Oktar, 2004] Türkçe Sözlü Söylemde Artgönderim Örüntüleri. Dilbilim Araştırmaları, 1-13.
16. [Enç, 1986] Topic switching and pronominal subjects in Turkish. D. Slobin and K.

- Zimmer eds. *Studies in Turkish Linguistics*. Amsterdam: John Benjamins.
17. [Erguvanlı-Taylan, 1986] Pronominal versus zero representation of anaphora in Turkish. D. Slobin and K. Zimmer eds. *Studies in Turkish Linguistics*. Amsterdam: John Benjamins.
 18. [Esmer, 2003] Şizofrenlerin Sözlü Anlatı Metinlerinin Çözümlemesi. Yüksek Lisans Tezi. Sosyal Bilimler Enstitüsü, Ankara Üniversitesi.
 19. [Fox, 1987] *Discourse Structure and Anaphora: Written and Conversational English*. Cambridge. Cambridge University Press.
 20. [Givón, 1983] *Topic Continuity in Discourse: A Quantative Cross-Language Study*. John Benjamins Publishing.
 21. [Givón, 1985] Iconicity, isomorphism, and non-arbitrary coding in syntax. J. Haiman, *Iconicity in Syntax*. Amsterdam: John Benjamins.
 22. [Givón, 1990] *Syntax: A functional-typological introduction*. Vol. II. Amsterdam: John Benjamins.
 23. [Grosz, 1995] Centering: A Framework for Modeling the Local Coherence of Discourse. In *Computational Linguistics*.
 24. [Grosz ve Sidner, 1986] Attention, intentions and the structure of discourse. In *Computational Linguistics*.
 25. [Günay, 2003] *Metin Bilgisi*. İstanbul: Multilingual Yayınları.
 26. [Güner, 2008] Türkçe için Derlem Tabanlı Bir Anafor Çözümleme Çalışması. Yüksek Lisans Tezi. Fen Bilimleri Enstitüsü, Trakya Üniversitesi.
 27. [Halliday ve Hasan, 1976] *Cohesion in English*. Longman English Language. Series 9. London: Longman.
 28. [Hinds, 1978] *Anaphora in Discourse*. Edmonton: Linguistic Research Inc.
 29. [Hinds, 1979] Organizational patterns in discourse. *Syntax and Semantics 12: Discourse and syntax* (s.135-157). New York: Academic Press.
 30. [Hobbs, 1976] Pronoun Resolution. Research Report 76-1, Department of Computer Sciences, City College, City University of New York. August 1976.
 31. [Hobbs, 1976] A Computational Approach to Discourse Analysis. Research Report 76-2, Department of Computer Sciences, City College, City University of New York. December 1976.
 32. [Hobbs, 1977] 38 Examples of Elusive Antecedents from Published Texts.

- Research Report 77-2, Department of Computer Sciences, City College, City University of New York, August 1977.
33. [Hobbs, 1978] Resolving Pronoun References. *Lingua*, Vol. 44, pp. 311-338. Also in *Readings in Natural Language Processing*, B. Grosz, K. Sparck-Jones, and B. Webber, editors, pp. 339-352, Morgan Kaufmann Publishers, Los Altos, California.
 34. [Huang, 2000] *Anaphora: A Cross-Linguistic Study*. Oxford: Oxford University Press.
 35. [Ingria ve Stallard, 1989] A Computational Mechanism for Pronominal Reference. *Proceedings of the 27th Annual Meeting of the ACL*, 262-271, Vancouver, British Columbia.
 36. [Kennedy ve Boguraev, 1996] Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser. In *Proceedings of the 16th International Conference on Computational Linguistics*, 113-118.
 37. [Kılıçaslan, 2004] Syntax of Information Structure in Turkish. *Linguistics*, 42-4, syf. 717-765.
 38. [Kohavi ve Provost, 1998] Glossary of Terms. In: Editorial for the Special Issue on Applications of Machine Learning and the Knowledge Discovery Process, *Machine Learning* 30(2-3).
 39. [Küçük, 2005] A knowledge-poor pronoun resolution system for Turkish. Master's thesis, Middle East Technical University.
 40. [Küçük ve Turhan-Yöndem, 2007] A knowledge-poor pronoun resolution system for Turkish. In *Proceedings of the 6th Discourse Anaphora and Anaphor Resolution Colloquium*.
 41. [Langacker, 1969] On Pronominalisation and the Chain of Command. *Modern Studies in English*, 160-186. Englewood Cliffs: Prentice Hall.
 42. [Lappin ve Leass, 1994] An Algorithm for Pronominal Anaphora Resolution. *Computational Linguistics*, 20(4), 535-561.
 43. [Longacre, 1979] The paragraph as a grammatical unit. *Syntax and Semantics 12: Discourse and Syntax* (s. 115-134). London: Academic Press.
 44. [Jurafsky ve Martin, 2000] *Speech and Language Processing*. Prentice-Hall.
 45. [Maingueneau, 1991] *L'Analyse do discours*. Paris, Hachette.
 46. [Maingueneau, 1996] *Les Termes clés de l'analyse du discours*. Paris, Seuil.

47. [Mann, 1988] Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text* 8(3): 243-281.
48. [McCarthy ve Lehnert, 1995] Using Decision Trees for Coreference Resolution. In *IJCAI*, 1050-1055.
49. [Mitkov, 1998a] Evaluating Anaphora Resolution Approaches. *Proceedings of the Discourse Anaphora and Anaphora Resolution Colloquium (DAARC'2)*, s. 164-172. Lancaster, UK.
50. [Mitkov, 1998b] Robust Pronoun Resolution with Limited Knowledge. In *Proceedings of the 17th International Conference on Computational Linguistics (COLING'98 / ACL'98)*, 869-875. Montreal, Canada.
51. [Mitkov, 2002] *Anaphora Resolution*. Pearson Education, ISBN 0 582 32505 6.
52. [Mitkov, Evans ve Orasan, 2002] A new, fully automatic version of Mitkov's knowledge-poor pronoun resolution method. In A1. Gelbukh (Ed.). *Computational Linguistics and Intelligent Text processing*, 168-186. Springer.
53. [Oktar ve Yağcıoğlu, 1997] Türkçe Söylem Yapısı ve Artgönderim. 8. Uluslararası Türk Dilbilimi Konferansı Bildirileri.
54. [Onursal, 2003] Türkçe Metinlerde Bağdaşıklık ve Tutarlılık. *Günümüz Dilbilim Çalışmaları*. Multilingual Yayınları, Dilbilim Dizisi, İstanbul, ss. 121-132.
55. [Özsoy, 1992] Türkçede Eşgönderge Örüntüsü. *Dilbilim Araştırmaları*, 77-85.
56. [Preiss, 2002] Anaphora Resolution with word sense disambiguation. In *Proc. of SENSEVAL-2*, s. 143-146.
57. [Quinlan, 1993] *Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA.
58. [Reinhart, 1981] *Pragmatics and Linguistics: An Analysis of Sentence Topics*. *Philosophica* 27, 53-94.
59. [Sezer, 1991] *Türkçe Sözdizimi*. Dilbilim ve Türkçe. Dil Derneği, Ankara.
60. [Soon, Ng ve Lim, 2001] A Machine Learning Approach to Coreference Resolution of Noun Phrases. *Computational Linguistics*, 27(4), 521-544.
61. [Tai, 1978] Anaphoric constraints in Mandarin Chinese narrative discourse. *Anaphora in discourse* (s. 279-338). Edmonton: Linguistic Research Inc.
62. [Tın ve Akman, 1994] Situated processing of pronominal anaphora. In Trost, Harald, Eds. In *Proceedings KONVENS '94 - 2. Konferenz "Verarbeitung*

- natuerlicher Sprache", pp. 369-378.
63. [Turan, 1995] Null vs. overt subjects in Turkish discourse: A centering analysis. Ph.D. thesis, University of Pennsylvania.
 64. [Turan, 1996] Türkçede Kişisiz Gizli Özne ve 'İnsan'. Dilbilim Araştırmaları. Ankara: Bizim Büro Basımevi, s. 245-249.
 65. [Turan, 1998] Zero Object Arguments and Referentiality in Turkish. The Mainz Meeting. Proceedings of the Seventh International Conference on Turkish Linguistics.
 66. [Turhan-Yöndem ve Şehitoğlu, 1997] Resolution of dropped pronouns. In Proceedings of the ESSLLI'97 Student Session.
 67. [Tüfekci ve Kılıçaslan, 2005] A Computational Model for Resolving Pronominal Anaphora in Turkish Using Hobbs' Naïve Algorithm. International Journal of Computational Intelligence, Volume 2, Number 1, 2005, ISSN: 1304-2386.
 68. [Tüfekci ve Kılıçaslan, 2007] A syntax-based pronoun resolution system for Turkish. In Proceedings of the 6th Discourse Anaphora and Anaphor Resolution Colloquium. Lagos (Algarve), Portugal.
 69. [Tüfekci, Küçük, Turhan Yöndem ve Kılıçaslan, 2007] Comparison of a Syntax-Based and a Knowledge-Poor Pronoun Resolution Systems for Turkish. 22nd International Symposium on Computer and Information Sciences, Middle East Technical University, Ankara, Türkiye.
 70. [Uzun, 2000] Ana Çizgileriyle Evrensel Dilbilgisi ve Türkçe. Multilingual, İstanbul.
 71. [Walker, 1989] Evaluating Discourse Processing Algorithms. In Proceedings of ACL-89.
 72. [Van Hoek, 1997] Anaphora and Conceptual Structure. Chicago: The University of Chicago Press.
 73. [Vardar, 1998] Açıklamalı Dilbilim Terimleri Sözlüğü. İstanbul: abc.
 74. [Yıldırım, 2008] Türkçe Derlemlerdeki Artgönderimlerin Tümdengelimli ve Tümevarımlı Yöntemlerle Çözümlemesi. Doktora Tezi. Fen Bilimleri Enstitüsü, Trakya Üniversitesi.
 75. [Yıldırım ve Kılıçaslan, 2007] A Machine Learning Approach to Personal Pronoun Resolution in Turkish. In: Proceedings of 20th International FLAIRS Conference,

FLAIRS-20, Key West, Florida.

76. [Yüksel ve Bozşahin, 2002] Contextually appropriate reference generation. *Natural Language Engineering*, 8 (1), 69-89.