A KNOWLEDGE-POOR PRONOUN RESOLUTION SYSTEM FOR TURKISH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

DİLEK KÜÇÜK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

SEPTEMBER 2005

Approval of the Graduate School of Natural and Applied Sciences

____
Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

____
Prof. Dr. Ayşe KİPER
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____
Dr. Meltem TURHAN YÖNDEM
Supervisor

Examining Committee Members:

Assoc. Prof. Dr. Göktürk ÜÇOLUK (METU, CENG)__________________________

Dr. Meltem TURHAN YÖNDEM (METU, CENG)__________________________

Assist. Prof. Dr. Bilge SAY (METU, COGS&SM)__________________________

Dr. Ayşenur BİRTÜRK (METU, CENG)__________________________

Dr. Onur Tolga ŞEHİTOĞLU (METU, CENG)__________________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Dilek KÜÇÜK

Signature :

ABSTRACT

A KNOWLEDGE-POOR PRONOUN RESOLUTION SYSTEM FOR TURKISH

Küçük, Dilek

M.S., Department of Computer Engineering

Supervisor: Dr. Meltem TURHAN YÖNDEM

September 2005, 68 pages

This thesis presents a knowledge-poor pronoun resolution system for Turkish which resolves third person personal pronouns and possessive pronouns. The system is knowledge-poor in the sense that it makes use of limited linguistic and semantic knowledge to resolve the pronouns. As pronoun resolution proposals for languages like English, French and Spanish, the core of the system is the constraints and preferences which are determined empirically.

The system has four modules: sentence splitting, pronoun extraction, forming the list of candidate antecedents and determination of the antecedent. It takes a Turkish text as input and rewrites this text with the considered pronouns replaced with their proposed antecedents. In order to compare the success rate of the system, two different baseline algorithms are implemented. The original system is tested against these baseline algorithms on two sample Turkish texts from different sources. Some suggestions to improve the success rate of the system and to extend the domain of the system are also presented.

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

ÖZ

TÜRKÇE İÇİN AZ BİLGİLİ BİR ADIL ÇÖZÜMLEME SİSTEMİ
Küçük, Dilek
Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi: Dr. Meltem TURHAN YÖNDEM

Eylül 2005, 68 sayfa

Bu tez, üçüncü kişi kişi adılları ile dönüşlülük adıllarını çözümleyen, Türkçe için az bilgili bir adıl çözümleme sistemi sunar. Sistem, adılları çözümlemek için sınırlı dilbilimsel ve anlamsal bilgi kullanması açısından az bilgilidir. İngilizce, Fransızca ve İspanyolca gibi diller için verilen adıl çözümleme önerileri gibi; sistemin merkezinde, deneysel olarak elde edilen sınırlamalar ve tercihler bulunur.

Sistemin dört modülü bulunmaktadır: cümlelere ayırma, adıl çıkarma, aday gönderge listesinin oluşturulması ve göndergenin tespit edilmesi. Sistem, bir Türkçe metni girdi olarak alır ve bu metni, göz önüne alınan adıllar önerilen göndergeleri ile değiştirilmiş olarak yeniden yazar. Sistemin başarı oranını karşılaştırmak için iki farklı temel algoritma gerçekleştirilmiştir. Asıl sistem, bu temel algoritmalara karşı, farklı kaynaklardan iki örnek metin üzerinde denenmiştir. Sistemin başarı oranını iyileştirmek ve alanını genişletmek için bir takım öneriler de sunulmuştur.

Anahtar Kelimeler : Artgönderim, Artgönderim Çözümlemesi, Adıl, Türkçe için Adıl Çözümlemesi.

ACKNOWLEDGMENTS

I am grateful to my thesis supervisor Dr. Meltem TURHAN YÖNDEM for her guidance, motivation and support throughout this study.

I want to thank my mother, my father, my sisters Emine and Nihal and my brothers Erkut and Doğan for their motivating support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

ABL	:	Ablative
ACC	:	Accusative
ANN	:	Artificial Neural Network
AOR	:	Aorist
AUX	:	Auxiliary
DAT	:	Dative
BFP	:	Brennan Friedman Pollard algorithm
FUT	:	Future
GEN	:	Genitive
GUI	:	Graphical User Interface
LRC	:	Left-Right Centering
LOC	:	Locative
MARS	:	Mitkov's Anaphora Resolution System
NLP	:	Natural Language Processing
NP	:	Noun Phrase
PAST	:	Past
PERS	:	Person
POSS	:	Possessive
POS	:	Part-Of-Speech
PROG	:	Progressive
RAP	:	Resolution of Anaphora Procedure

CHAPTER 1

INTRODUCTION

A word or phrase that refers to an entity that is mentioned previously is called an anaphor and this word or phrase that it refers to is called its antecedent. Anaphora is the situation in which an anaphor exists. The most widespread type of anaphora is the pronominal anaphora which is realized by anaphoric pronouns (Mitkov 2002). The process of identifying the antecedents of anaphors is called anaphora resolution.

Anaphora resolution is a commonly studied research area of Natural Language Processing (NLP). It is crucial for many application areas of NLP including information extraction, question answering, and text summarization.

Several algorithms have been proposed for anaphora resolution and also there is ongoing research on this topic. Most of these algorithms can be classified as a member of one of the two main categories. The first category comprises those algorithms that use extensive domain and linguistic knowledge, which can be termed as “discourse-based algorithms” (Brennan et al. 1987; Strube 1998; Tetreault 1999). The algorithms in the second category are usually called “knowledge-poor algorithms”, since they use a salience-based strategy based on syntax with limited usage of linguistic and domain knowledge (Kennedy and Boguraev 1996; Baldwin 1996; Mitkov 1998; Palomar et al. 2001; Tanev and Mitkov 2002; Trouilleux 2002). There are also other proposals for anaphora resolution that use statistics, machine learning approaches or semantics.

In this thesis, we describe a knowledge-poor pronoun resolution system for Turkish which resolves third person personal pronouns and reflexive pronouns which refer to proper person names in text.

The sequence of steps carried out by knowledge-poor pronoun resolution algorithms are:

1. Identification of anaphors
2. Location of the candidates for antecedents
3. Resolution of the anaphor using constraints and preferences (Mitkov 2002)

In this system, a similar strategy for Turkish is used. In order for the system to identify the personal and reflexive pronouns that refer to proper person names, the input text is preprocessed to mark these pronouns before it is given as input to the system.

To locate the candidate antecedents and to determine the constraints and preferences for Turkish, two empirical methods are carried out:

1. Manual analysis of a sample Turkish text
2. Questionnaire on native Turkish speakers

At the end of the manual analysis of the sample text, the following information is acquired:

1. The sentences from which the candidate antecedents are extracted to carry out the second step of the resolution procedure
2. Constraints and preferences to be used in the third step of the resolution procedure

Constraints are used to filter the candidates that cannot be the antecedent of a pronoun. The constraints for English include number and gender agreement, c-command constraints and selectional restrictions (Mitkov 2002). Among these, gender agreement is definitely not applicable to Turkish since Turkish pronouns do not denote gender. Number agreement is an applicable constraint to Turkish and is considered in our system. At the end of the manual analysis of the sample Turkish text, it is found that c-command constraints can not be applied directly to Turkish due to the knowledge-poor nature of the system and the properties of the reflexive pronouns in Turkish. However, these c-command constraints are modified and adapted to Turkish as two different constraints, namely, reflexive pronoun constraint and personal pronoun constraint. Selectional restrictions is an applicable constraint to Turkish but since it requires considerable semantic knowledge, it is not used in this system. Therefore, the constraints that are used in the system are number agreement, reflexive pronoun constraint, and personal pronoun constraint.

The preferences for Turkish which are used to sort the remaining candidates are quoted/unquoted text, recency, subject preference, first noun phrase (NP), nominal predicate, repetition, punctuation, and antecedent of zero pronoun preferences. First NP preference has already been used for Bulgarian (Tanev and Mitkov 2002). Nominal predicate preference is similar to the 'existential emphasis' preference used for English (Lappin and Leass 1994; Kennedy and Boguraev 1996). Recency, repetition, syntactic parallelism and subject preferences are extensively used for different languages (Lappin and Leass 1994; Mitkov 1998; Trouilleux 2002). During the empirical analysis, the applicability of these preferences to Turkish is verified. In addition to these, three preferences, namely, quoted/unquoted text, punctuation, and antecedent of zero pronoun preferences are determined empirically for Turkish. Among all these preferences, syntactic parallelism preference and subject preference are not used in the system since they require knowledge; however, a special case of subject preference, namely, nominative case preference is described and used in the system. A questionnaire on native Turkish speakers is carried out to verify that the constraints and preferences for Turkish are valid. Statistical methods are used to determine the statistical significance of the results of the questionnaire and the number of agreements necessary for a preference to be considered as valid.

The implementation steps of the system are sentence splitting, extraction of the third person pronouns and reflexive pronouns from the text, forming the list of candidates for antecedents of each of the extracted pronouns, and determining the antecedent of each extracted pronoun from left to right by applying constraints and preferences.

In this study, our claim is that in Turkish texts, we can get successful results in resolving personal and reflexive pronominal anaphors that refer to proper person names by employing a knowledge-poor anaphora resolution approach.

We examine existing anaphora resolution algorithms and their applicability to Turkish. We develop an anaphora resolution system for pronominal and reflexive anaphora referring to proper nouns that employs a knowledge-poor approach and compare the evaluation results of the system with a system that uses a baseline algorithm for Turkish as well as with the results of existing anaphora resolution systems implemented for languages including English, Spanish and French.

In Chapter 2, some background information about anaphora resolution and pronouns in Turkish is presented. Chapter 3 provides some background to the studies on anaphora resolution. In Chapter 4, details of the empirical studies on Turkish are provided. In Chapter 5, design and implementation of the pronoun resolution system for Turkish is described. Finally, Chapter 6 describes the testing results of the system and in Chapter 7 these results are discussed.

CHAPTER 2

BACKGROUND

2.1. Anaphora and Anaphora Resolution

A word or phrase that refers to an entity that has already been introduced into the discourse is called an anaphor. The entity that an anaphor refers to is called its antecedent. The process of identifying the antecedents of anaphors is called anaphora resolution. Anaphora is described as the cohesion which points back to some previous entity in a discourse, where the pointing back word or phrase is the anaphor. If an anaphor and its antecedent both refer to the same entity in the real world, they are called coreferential (Mitkov 2002).

According to the form of the anaphor, anaphora can be divided into six main groups, namely, pronominal anaphora, lexical noun phrase anaphora, one-anaphora, verb anaphora, adverb anaphora, and zero anaphora.

Pronominal anaphora occurs when the anaphoric word or phrase is a personal pronoun, possessive pronoun, reflexive pronoun, demonstrative pronoun or a relative pronoun.

> Doctors examined the child. *They* decided to make an operation.

When pronoun is mentioned before the antecedent, this situation is called *cataphora*.

> Before *she* left the house, Mary turned off all lights.

Lexical noun phrase anaphora occurs when the anaphor is a definite noun phrase or a proper name (Mitkov 2002).

Roy Keane has warned Manchester United he may snub their pay deal. *United's skipper* is even hinting that unless the future Old Trafford Package meets his demands, he could quit the club in June 2000. (Mitkov 2002)

One-anaphora is the case when the anaphoric expression is realized by a "one" noun phrase.

If you like those books, you can take *one* with you.

Verb anaphora occurs when then the antecedent of the anaphor is a verb.

Alice woke up early yesterday. So *did* her brother.

Adverb anaphora occurs when then the antecedent of the anaphor is an adverb.

John walked to the garden and stayed *there*.

Zero anaphora is the case when the anaphor is omitted but it is understood.

Mary left the house and ∅ began walking.

Some linguists also categorize anaphors according to their being in the same sentence with their antecedents or not. According to this categorization, an intrasentential anaphor refers to an antecedent which is in the same sentence as the anaphor, whereas an intersentential anaphor refers to an antecedent which is in a different sentence from that of the anaphor (Mitkov 2002).

2.2. Pronouns in Turkish

There are six types of pronouns in Turkish:

1. Personal Pronouns
2. Demonstrative Pronouns
3. Reflexive Pronouns
4. Possessive Pronouns
5. Interrogative Pronouns

6. Indefinite Pronouns

These pronouns are described in detail in the following sections.

2.2.1. Personal Pronouns

Personal pronouns in Turkish are *ben* (I), *sen* (you), *o* (he/she/it), *biz* (we), *siz* (you) and *onlar* (they). Unlike languages like English, German and French, in Turkish, third person singular pronoun, *o*, does not denote gender. The following example demonstrates the use of personal pronouns in Turkish:

O, futbol maçına gitti.

He football match-POSS-DAT go-PAST

'He went to the football match.'

However, third person singular pronoun has the same morphology with the demonstrative, *o (*that*)*, which is used as an adjective and should not be confused with it. The example below shows demonstrative, *o*, which is used as an adjective, in a sentence:

Çocuk, o okula gitmedi.

Child, that school go-NEG-PAST

'The child did not go to that school.'

Also, as will be explained below, third person singular and plural pronouns have the same morphology with the demonstrative pronouns *o* (that) and *onlar* (those).

2.2.2. Demonstrative Pronouns

In Turkish, demonstrative pronouns are *bu* (this), *şu* (this or that), *o* (that), *bunlar* (these), *şunlar* (these or those) and *onlar* (those). The following example shows the use of demonstrative pronouns:

Şu, güzel bir ev.

That beautiful one house

'That is a beautiful house.'

However, as in the case of pronominal pronouns, *bu, şu* and *o* can also be used in adjectival positions. The following is an example of this case:

Adam çocuğa o kitabı verdi.

Man child-DAT that book-ACC give-PAST

'The man gave that book to the child.'

2.2.3. Reflexive Pronouns

Reflexive pronouns in Turkish have two forms, namely, *kendi* (oneself) and *kendisi* (oneself). In the following example, *kendi* is used as a reflexive pronoun:

Yaşlı adam aynada kendine baktı.

Old man mirror-LOC himself-ACC look-PAST

'The old man looked at himself in the mirror.'

2.2.4. Possessive Pronouns

When added to the genitive case of a noun or pronoun, the pronominal suffix *–ki* makes a possessive pronoun (Lewis 2000). The example below demonstrates a possessive pronoun formed using the pronominal suffix *–ki*.

Ayşe'nin ayakkabısı eski, Ayla'nınki yeni.

Ayşe-GEN shoe-POSS old, Ayla-GEN-ki new

'Ayşe's shoes are old, the ones belonging to Ayla are new'

2.2.5. Interrogative Pronouns

When question words *ne* (what), *kim* (who), *nerede* (where), and *hangi* (which) are used as pronouns, they are called interrogative pronouns. An interrogative pronoun is exemplified in the following sentence:

Bu kitabı kim ister?

This book-ACC who read-AOR?

‘Who wants this book?’

2.2.6. Indefinite Pronouns

There are many different forms of indefinite pronouns in Turkish. Some of the words used as indefinite pronouns are *bazısı* (some people), *biri* (someone), *herkes* (everybody), and *kimi* (some people). An example of indefinite pronouns is given below:

Bazıları eve gitmiş.

Some people home go-PAST

‘Some people went home.’

CHAPTER 3

LITERATURE SURVEY

In this chapter, we describe existing research on anaphora resolution. In the first section, we outline studies based on Centering Theory which are usually called discourse-based anaphora resolution algorithms. In the second section, we explore the proposals which limit their use of linguistic and domain knowledge and these algorithms are usually termed knowledge-poor anaphora resolution algorithms. Finally, in the last section, we describe some of the important studies related to anaphora resolution in Turkish.

3.1. Discourse-Based Anaphora Resolution Algorithms

Centering Theory is one of the most influential theories on anaphora resolution. In this section, we explore the studies that describe the Centering Theory (Grosz et al. 1983; Grosz et al. 1995) and a study on discourse structure (Grosz and Sidner 1986) as well as some important discourse-based anaphora resolution algorithms that are based on Centering Theory.

3.1.1. Grosz, Joshi, and Weinstein [1983, 1995]

In their papers (Grosz et al. 1983; Grosz et al. 1995), Grosz, Joshi and Weinstein describe Centering Theory which provides an explanation for local coherence in a discourse. Centering Theory is used extensively by many researchers who study anaphora resolution. According to Centering Theory, each utterance, U, in a discourse has a backward-looking center, $C_b(U)$, and a set of forward-looking centers, $C_f(U)$ which integrate this utterance to the discourse. $C_b(U)$ serves to link U to the preceding discourse, while $C_f(U)$ provides a set of entities to which the succeeding discourse may be linked. There exists a language-specific ranking of forward looking centers, for instance, in English; C_f list is ranked according to grammatical roles. The most highly ranked element of the C_f list is the preferred center, C_p.

Three types of transitions exist across utterances, namely, center continuation, center retaining and center shifting. According to the relationships between backward-looking and preferred centers of consecutive sentences, one of these transitions holds among the sentences.

There exist two main rules of Centering Theory. The first rule states that if any element of $C_f(U_n)$ is realized by a pronoun in U_{n+1} then the $C_b(U_{n+1})$ must be realized by a pronoun also. The second rule of the theory states that sequences of continuation are preferred over sequences of retaining; and sequences of retaining are to be preferred over sequences of shifting. The two centering rules along with the partial ordering on the forward-looking centers constitute the basic framework of center management. These rules can explain a range of variations in local coherence.

3.1.2. Grosz and Sidner [1986]

Grosz and Sidner describe a theory of discourse structure in which, discourse structure is composed of three components, namely, linguistic structure, intentional structure, and attentional state (Grosz and Sidner 1986). Linguistic structure consists of discourse segments that the utterances in a discourse form. Intentional state is used to describe the discourse-relevant purposes of each linguistic segments and relationships between these purposes. Attentional state records the objects, properties, and relations that are salient at each point of discourse. This theory of discourse structure is important for anaphora resolution, because entities in attentional state are usually the best candidate antecedents for the anaphors in a discourse.

3.1.3. Brennan, Friedman, and Pollard [1987]

Brennan, Friedman and Pollard developed an algorithm for pronoun resolution in English by using the constraints and rules of Centering Theory (Brennan et al. 1987). They extended the Centering framework by distinguishing between smooth-shift and rough-shift. In the original Centering theory, a shift occurs when successive C_b's are not the same, whereas; in this work of Brennan and colleagues, a smooth-shift occurs when $C_b(U_N) \neq C_b(U_{N-1})$ and $C_b(U_N) = C_p(U_N)$ and a rough shift occurs when $C_b(U_N) \neq C_b(U_{N-1})$ and $C_b(U_N) \neq C_p(U_N)$. Transition orderings with this extension is as follows: continuing > retaining > smooth-shift > rough-shift.

The pronoun resolution algorithm using this extended Centering framework is as follows:

1. Generate possible C_b-C_f combinations which are called anchors.
2. Filter by constraints and rules of Centering Theory
3. Rank by transition orderings.

3.1.4. Strube [1998]

Strube proposes a model to describe the attentional state of the hearer in a discourse (Strube 1998). The model is similar to Centering Theory but it enables incremental processing of utterances which is a property lacking in the Centering theory. Strube's model consists of one construct called S-list, which is a list of discourse entities. It contains some discourse entities in the current and previous utterance. The elements of the S-list are ranked according to their utterance and their being hearer-old, mediated or hearer-new discourse entities. Hearer-old discourse entities are favored over mediated and hearer-new discourse entities, and mediated discourse entities are favored over hearer-new discourse entities.

Steps of anaphora resolution algorithm using S-list are:

1. If a referring expression is encountered
    a. If it is a pronoun, test the elements of the S-list in the given order until agreement constraints, binding and sortal constraints are satisfied.
    b. Update S-list; the position of the referring expression under consideration is determined by the S-list-ranking criteria which are used as an insertion algorithm.
2. If the analysis of utterance U is finished, remove all discourse entities from the S-list, which are not realized in U.

3.1.5. Tetreault [1999]

Tetreault presents a pronoun resolution algorithm based on Centering Theory, which performs better than the algorithm proposed by Brennan et al. (BFP) in 1978 (Tetreault 1999). This algorithm, namely, Left-Right Centering (LRC), is proposed since BFP lacks in incremental processing of pronouns and generating and filtering the elements of forward-looking centers causes a computational overhead in BFP. LRC algorithm works by searching for the antecedent in the current sentence, if it does not find an antecedent; it searches the previous Cf-lists left-to-right for an antecedent.

Tetreault compares his algorithm with the algorithms proposed by Hobbs (1977), Brennan et al. (1987) and Strube (1998), in addition to a baseline algorithm which favors the most recent noun phrase. The results of these comparisons reveal that LRC and Hobbs' algorithms perform better than the other algorithms examined. In the paper, it is stated that, these results are obtained since both algorithms search for referents intrasententially and then intersententially, and they search for their respective data structures in a salience-first manner.

3.2. Knowledge-Poor Anaphora Resolution Algorithms

Knowledge-poor anaphora resolution algorithms are the algorithms which do not use too much linguistic and domain knowledge to resolve the anaphors. As reported in (Mitkov 2002), the pressing need for the development of robust and inexpensive solutions to meet the demands of practical NLP systems encouraged many researchers to move away from extensive domain and linguistic knowledge and to embark instead upon knowledge-poor anaphora resolution strategies. In this section, we address two important anaphora resolution algorithms which influence knowledge-poor algorithms (Hobbs 1977; Lappin and Leass 1994), and then we describe some of the most well-known knowledge-poor proposals.

3.2.1. Hobbs [1977]

In his paper (Hobbs 1977); Hobbs describes two approaches to pronoun resolution in English. The first one is a simple, efficient, but naive algorithm working on the surface parse trees of the sentences in the text. His second approach is a complex semantic one which uses semantic analysis.

In the naive algorithm, surface parse tree of each sentence in the input text is used. This tree exhibits the grammatical structure of the sentence without permuting or omitting any of the words in the original sentence. The algorithm traverses the surface parse tree in a left-to-right depth-first manner looking for a noun phrase of the correct gender and number. One hundred consecutive examples of pronouns from each of three different texts were examined to test the performance of the naive algorithm. Overall, the algorithm worked in 88.3% of the cases. The algorithm together with selectional constraints worked 91.7% of the time. Jerry Hobbs' naïve approach remains one of the most influential works in the field and frequently serves as a 'classical' benchmark for evaluating current proposals (Mitkov 2002).

The semantic approach is based on semantic operations including detecting intersentence connectives, predicate interpretation, knitting, and identifying entities.

3.2.2. Lappin and Leass [1994]

The algorithm proposed by Lappin and Leass is called Resolution of Anaphora Procedure (RAP) and is based on measures of salience derived from syntactic structure and a simple dynamic model of attentional state (Lappin and Leass 1994). Both intrasentential and intersentential pronouns can be resolved using RAP.

When a pronoun *it* does not refer to anything specific, it is termed *pleonastic* (non-anaphoric) (Mitkov 2002) as in the example:

> *It* must be acknowledged that the truth was concealed.

Lappin and Leass's algorithm, RAP, identifies pleonastic pronouns and does not attempt to resolve these pronouns.

During the resolution procedure, by using morphological and syntactic filters, some of the candidate noun phrases (NPs) are filtered out, and remaining candidates are assigned salience measures according to predefined salience factor types. The candidate with the highest salience value is selected as the antecedent. The salience factor types used in this algorithm are:

1. *Sentence recency*: This factor gives preference to the candidates in recent sentences.

2. *Subject emphasis*: This preference is given to the NPs at subject positions.

3. *Existential emphasis*: This preference is given to the predicate nominals in existential constructions.

4. *Accusative emphasis*: This preference is given to the direct objects.

5. *Indirect object and oblique complement emphasis*: This preference if given to indirect objects and oblique complements.

6. *Head noun emphasis*: This preference is given to the NPs which are not contained in other NPs.

7. *Non-adverbial emphasis*: This preference is given to the NPs which are not contained in adverbial prepositional phrases (Lappin and Leass 1994).

RAP is tested on 360 pronoun occurrences which were randomly selected from a corpus of computer manuals containing 1.25 million words. The success rate of the algorithm is 86%, with 72% success rate for intersentential pronouns and 89% for intrasentential pronouns.

3.2.3. Kennedy and Boguraev [1996]

Kennedy and Boguraev's anaphora resolution algorithm is extended and modified version of Lappin and Leass's (1994) algorithm (Kennedy and Boguraev 1996). In this algorithm, a set of discourse referents is generated where each discourse referent contains information about itself and the context in which it appears. The absence of explicit information about configurational relations is the crucial difference between this algorithm and Lappin and Leass's algorithm. Configurational information is used in Lappin and Leass's algorithm both in the determination of the salience of a discourse referent (as in the case of head noun emphasis or non-adverbial emphasis) and in the disjoint reference filters (as in syntactic filter on pronoun-NP coreference). In Kennedy and Boguraev's algorithm, each discourse referent contains information about itself and the context in which it appears, the only information about its relation to other discourse referents being in the form of precedence relations (as indicated by the text position) (Mitkov 2002).

In this algorithm, coreference is represented in terms of equivalence classes of anophorically related discourse referents which are called "coref" classes. Coreference is determined by first filtering out those discourse referents that does not pass the agreement and disjoint reference filters and then selecting the most salient discourse referent after applying salience measures. The salience factor types used in the algorithm are sentence recency, context emphasis, subject emphasis, existential emphasis, possessive emphasis, accusative emphasis, indirect object emphasis, oblique complement emphasis, head noun emphasis and non-adverbial emphasis. The algorithm introduces two new salience factors in addition to the salience factor types of the algorithm proposed by Lappin and Leass's (1994):

1. *Context emphasis*: This preference is given to the NPs that are in the same discourse segment as the anaphor.

2. *Possessive emphasis*: This preference is given to the NPs whose grammatical function is possessive (Kennedy and Boguraev 1996).

After the resolution of a pronoun, this pronoun is added to the coref class of the discourse referent which is the antecedent of the pronoun and the salience of this coref class is recalculated.

The algorithm is tested on 27 texts taken from different genres. 231 of 306 third person pronouns were correctly resolved; therefore, the success rate of the algorithm is 75%.

3.2.4. Baldwin [1996]

In his work (Baldwin 1996), Breck Baldwin describes a high precision pronoun resolution engine called CogNIAC. This resolution engine resolves a subset of anaphors that do not require general world knowledge for successful resolution and it resolves pronouns only if their antecedents are not ambiguous.

The rules that CogNIAC uses are:

1. *Unique in discourse*: If there is a single possible antecedent *i* in the read-in portion of the entire discourse, then pick *i* as the antecedent.

2. *Reflexive*: Pick nearest possible antecedent in read-in portion of current sentence if the anaphor is a reflexive pronoun.

3. *Unique in current + prior*: If there is a single possible antecedent *i* in the prior sentence and the read-in portion of the current sentence, then pick *i* as the antecedent.

4. *Possessive pro*: If the anaphor is a possessive pronoun and there is a single exact string match *i* of the possessive in the prior sentence, then pick *i* as the antecedent.

5. *Unique current sentence*: If there is a single possible antecedent *i* in the read-in portion of the current sentence, then pick *i* as the antecedent.

6. *Unique subject/subject pronoun*: If the subject of the prior sentence contains a single possible antecedent *i*, and the anaphor is the subject of its sentence, then pick *i* as the antecedent (Baldwin 1996).

In order to compare the algorithm with Hobbs' naive algorithm, two lower precision rules are added to the original system for comparison reasons, these rules are Cb-picking and pick most recent:

1. *Cb-picking*: If there is a Cb *i* in the current finite clause that is also a candidate antecedent, then pick *i* as the antecedent.

2. *Pick most recent*: Pick the most recent potential antecedent in the text (Baldwin 1996).

In CogNIAC, above rules are tried beginning from the first to the last and if an antecedent is found, the other rules are not tried.

The algorithm correctly resolves 232 of the 298 of the personal pronouns in a sample narrative text, so, the success rate of the algorithm is 77.9%.

3.2.5. Mitkov [1998]

Mitkov presents a robust, knowledge-poor approach to resolving pronouns in technical manuals (Mitkov 1998). The algorithm takes as input the text preprocessed by a part-of-speech tagger. In this approach, when a pronoun is encountered in the input text, the noun phrases from the current and the two preceding sentences are extracted as candidates. Some of these candidates are eliminated by gender and number agreement filters. Mitkov uses the term "antecedent indicators" to denote the preference rules. After the application of the agreement filters, genre-specific antecedent indicators are applied to the remaining candidates and the candidate noun phrase with the highest aggregate score is declared as the antecedent.

Mitkov does not consider cataphora in this approach and non-anaphoric (pleonastic) occurrences of "it" are eliminated by a referential filter before the resolution process begins.

The antecedent indicators of Mitkov's approach are:

1. *Definiteness*: Definite noun phrases in previous sentences are more likely antecedents of pronominal anaphors than indefinite ones.

2. *Givenness*: Noun phrases in previous sentences representing the "given information" (theme) are deemed good candidates for antecedents.

3. *Indicating verbs*: If a verb is a member of the Verb_set = {discuss, present, illustrate, identify, summarise, examine, describe, define, show, check, develop, review, report, outline, consider, investigate, explore, assess, analyse, synthesise, study, survey, deal, cover}, the first NP following it is considered as the preferred antecedent.

4. *Lexical reiteration*: Lexically reiterated items are likely candidates for antecedent.

5. *Section heading preference*: If a noun phrase occurs in the heading of the section, part of which is the current sentence, then it is considered as the preferred candidate.

6. *"Non-prepositional" noun phrases*: A "pure", "non-prepositional" noun phrase is given a higher preference than a noun phrase which is part of a prepositional phrase.

7. *Collocation pattern preference*: This preference is given to candidates which have an identical collocation pattern with a pronoun. The collocation preference here is restricted to the patterns "noun phrase (pronoun), verb" and "verb, noun phrase (pronoun)".

8. *Immediate reference:* In technical manuals the "immediate reference" clue can often be useful in identifying the antecedent. The heuristics used is that in constructions of the form "...(You) V_1 NP ... con (you) V_2 it (con (you) V_3 it)", where con $\in$ {and/or/before/after...}, the noun phrase immediately after V_1 is a very likely candidate for antecedent of the pronoun "it" immediately following V_2 and is therefore given preference.

9. *Referential distance*: In complex sentences, noun phrases in the previous clause are the best candidate for the antecedent of an anaphor in the subsequent clause,

followed by noun phrases in the previous sentence, then by nouns situated 2 sentences further back and finally nouns 3 sentences further back.

10. *Referential distance*: In complex sentences, noun phrases in the previous clause are the best candidate for the antecedent of an anaphor in the subsequent clause, followed by noun phrases in the previous sentence, then by nouns situated 2 sentences further back and finally nouns 3 sentences further back (Mitkov 1998).

The success rate of the algorithm on technical manuals is 89.7%. Fully automated version of Mitkov's anaphora resolution system is called MARS (Mitkov's Anaphora Resolution System) (Mitkov et al 2002).

3.2.6. Palomar et al. [2001]

In this work, an algorithm for identifying noun phrase antecedents of personal pronouns, demonstrative pronouns, reflexive pronouns, and omitted pronouns (zero pronouns) in Spanish is described (Palomar et al. 2001). Algorithm contains following main components: identification of the type of the pronoun, constraints (morphological agreement (person, gender, and number), syntactic conditions on NP-pronoun noncoreference, and preferences.

Syntactic constraints are based on c-command and minimal-governing-category constraints. To obtain the different sets of preferences, a training corpus is used to identify the importance of each kind of knowledge that is used by humans when tracking down the NP antecedent of a pronoun. The antecedents of each pronoun in the text were identified, along with their configurational characteristics with reference to the pronoun and how often each characteristic is valid for the solution of a particular pronoun is determined. The order of importance was determined by first sorting the preferences according to the percentage of each configurational characteristic; that is, preferences with higher percentages were applied before those with lower percentages.

In evaluation phase, algorithm is tested on both technical manuals and literary texts. Over literary text corpora, the algorithm attained a success rate for anaphora resolution of 76.8%.

3.2.7. Tanev and Mitkov [2002]

The study presents the development and implementation of an architecture for language processing in Bulgarian called LINGUA, which includes modules for POS tagging, sentence splitting, clause segmentation, parsing, and anaphora resolution (Tanev and Mitkov 2002). LINGUA uses knowledge-poor, heuristically based algorithms for language analysis, in this way getting round the lack of resources for Bulgarian.

Anaphora resolution module of LINGUA resolves third-person personal pronouns, and is an adaptation of Mitkov's robust, knowledge-poor multilingual approach whose latest implementation is referred to as MARS.

The preferences for Bulgarian which have positive scores used in the pronoun resolution module of LINGUA are: first noun phrases, indicating verbs, lexical reiteration, section heading preference, collocation match, immediate reference, sequential instructions, term preference, selectional restriction pattern, adjectival NP, and name preference. The preference which has a negative score is prepositional noun phrases, that is, NPs appearing in prepositional phrases are assigned a score of -1. Two preferences, referential distance and indefiniteness may increase or decrease a candidate's score.

3.2.8. Trouilleux [2002]

In this paper, Trouilleux presents a robust system to resolve a subset of pronominal expressions in French (Trouillux 2002). This system implements a strategy similar to the one used in Lappin and Leass's RAP algorithm.

General strategy that is followed in this system is: syntactic analysis of the input text with identification of non-anaphoric pronouns, building a set of possible antecedents, discarding some of them based on a set of constraints, reducing antecedents to one based on a set of ordered preferences. Anaphoric NPs, NPs denoting a person or organization, NPs at subject positions, NPs occupying the same function as the object pronoun and NPs which are closer to the pronoun in intrasentential anaphora, are given preference by the system.

The results of this system shows that the preference for subject antecedent is stronger in intersentential anaphora than in intrasentential anaphora, confirming two observations already made by Baldwin (1995) and Tetreault (2001).

3.3. Studies Related to Anaphora Resolution in Turkish

There exist studies concerning different aspects of anaphora and anaphora resolution in Turkish. These studies include the research on overt and zero representations of anaphora in Turkish (Enç 1986; Erguvanlı-Taylan 1986), situation semantics approach to pronominal anaphora in Turkish (Tın and Akman 1994), discourse anaphora in Turkish from the perspective of Centering Theory (Turan 1995), resolution of dropped pronouns in Turkish (Turhan-Yöndem and Şehitoğlu 1997), and anaphora generation in Turkish (Yüksel and Bozşahin 2002).

3.3.1. Enç [1986]

In her study (Enç 1986); Enç states that there are two kinds of discourse anaphora in Turkish, namely, pronominal and zero anaphora. According to her study, the choice of the type of anaphor depends on whether the topic of discourse is maintained or a new topic is introduced. If a new topic is introduced into the discourse, an overt pronoun is used, whereas if the topic is maintained zero anaphor is used. Enç also argues that another use of overt pronouns is to contrast the referent of the pronoun with the referent of another NP.

3.3.2. Erguvanlı-Taylan [1986]

In her work on pronominal versus zero representation of anaphora in Turkish (Erguvanlı-Taylan 1986), Erguvanlı-Taylan emphasizes that Turkish employs pronominal and zero representations of anaphora to convey coreference. Coreferentiality with another NP can be expressed by three situations: zero anaphora, pronominal anaphora or, zero or pronominal anaphora.

When the presence of a pronominal anaphora signals distinct reference, zero anaphora is essential. In cases where zero anaphora makes the sentence ungrammatical, pronominal anaphora is obligatory. In conjoined structures with the anaphoric expression being a non-subject of the second sentence in the conjoined structure, either the pronominal or zero anaphora may be used.

In the study, it is also emphasized that discourse context determines the interpretation of anaphoric expressions when the anaphoric relations extend beyond boundaries.

3.3.3. Tın and Akman [1994]

Tın and Akman describe an approach to anaphora resolution in Turkish in the framework of situation semantics (Tın and Akman 1994). The computational environment they use is called BABY-SIT which is a framework employing situation-theoretic constructs.

According to situation theory (Devlin 1991), individuals, properties, relations, spatio-temporal locations, and situations are the basic ingredients. A discourse situation involves the expression uttered, its speaker, the spatio-temporal location of the utterance and the addressee(s). The utterance of an expression constrains the world in a certain way, depending on how the roles for discourse situations, connections, and the described situation are occupied. In interpreting the utterance of the expression in a context, there is a flow of information, partly from the linguistic form encoded in the expression and partly from the contextual factors provided by the utterance situation. The meaning of the expression and hence its interpretation are influenced by other factors such as stress, modality, and intonation. However, the situation in which the expression is uttered and the situation described by this utterance seem to play the most influential roles (Tın and Akman 1994).

3.3.4. Turan [1995]

In her study, Turan studies three important issues related to discourse anaphora in Turkish from the perspective of Centering theory (Turan 1995). The first issue is the determination of the noun phrases that contribute to the Cf-list in Turkish. It is shown that nonreferential expressions as well as referential expressions serve as antecedents to pronouns. The second issue is the factors that affect the ranking of the Cf-list in Turkish. Unlike English in which Cf-list is ranked according to grammatical roles, it is stated that Cf-list is ranked according to thematic roles in Turkish. The third problem is to determine the discourse functions of overt pronouns, zero pronouns and full NPs in subject position in Turkish. The discourse functions of null vs. overt pronouns vs. full NPs and the way in which they pattern in Centering transitions are discussed in the study.

3.3.5. Turhan-Yöndem and Şehitoğlu [1997]

Turhan-Yöndem and Şehitoğlu describe intrasentential resolution of Turkish dropped pronouns in a phrase-structure grammar (Turhan-Yöndem and Şehitoğlu 1997). They state that the resolution scheme for dropped pronouns depends on the constituent order.

Resolution rules for different surface orders and an implementation for an HPSG based parser are introduced.

The following rules are used in the system developed:

- Nominal objects preceding the dropped pronoun are candidate antecedents.
- If the dropped pronoun belongs to the first NP of the sentence, following nouns are candidate antecedents.
- Nominal objects in the post-verbal position are candidate antecedents of any dropped pronon before the verb.
- Relativized nouns cannot be antecedents of the dropped pronouns.

It is also stated in the study that semantic clues are as important as the surface order since native speakers eliminate most of the ambiguities semantically by their knowledge about the world and the objects and therefore it is essentially important to take semantics into account for producing only valid ambiguities.

3.3.6. Yüksel and Bozşahin [2002]

In this study on contextually appropriate anaphora generation in Turkish, Yüksel and Bozşahin describe a reference planning system, the goal of which is to retain, drop or replace the full NPs in the generation so that the resulting discourse is quite natural (Yüksel and Bozşahin 2002). The system uses a set of rules for binding relations and Centering Theory to model local and nonlocal reference. Local reference is planned by binding rules. According to the binding rules, the antecedents of reflexive pronouns must be in the so-called local domain, whereas the antecedents of personal pronouns must be outside the local domain (Mitkov 2002). In the system, nonlocal reference is planned by an interaction of binding and centering rules.

Pro-drop is the situation where pronominal subjects, objects and specifiers of possessive NPs are dropped (Yüksel and Bozşahin 2002). In the anaphora generation system, pro-drop is handled at the final stage of reference planning.

The system is tested as part of a machine translation system and also as a stand-alone system. The success rate of the system is found as 70% which is comparable to existing anaphora generation systems proposed for languages such as Chinese.

CHAPTER 4

ANALYSIS: EMPIRICAL STUDIES ON TURKISH

To our knowledge, there exists no knowledge-poor pronoun resolution system proposed and implemented for Turkish. In order to implement such a system, the strategy employed by most of the knowledge-poor proposals for languages other than Turkish, will be used.. As described in Mitkov (1998), these systems use the following strategy:

1. Identification of anaphors

2. Location of the candidates for antecedents

3. Resolution of the anaphor using constraints and preferences

The pronoun resolution system for Turkish will attempt to resolve only the third person pronouns and possessive pronouns which refer to proper person names. Therefore, in the first step, the system will identify only these pronouns and will attempt to resolve later. In order to carry out the second step, the number of sentences from which the proper person names will be extracted should be determined, and similarly, to perform the last step, the constraints and preferences for Turkish should be determined.

Empirical studies should be carried out to determine the number of sentences to consider when extracting the candidates and the constraints and preferences for Turkish. For analysis, two methods are used:

1. Manual analysis of a sample Turkish text

2. Questionnaire on native Turkish speakers to verify the findings of the manual analysis.

4.1. Empirical Analysis on a Sample Turkish Text

An empirical study on a Turkish child narrative (Ilgaz 2003a) is carried out in order to determine necessary information related to pronoun resolution in Turkish. This text consists of 8647 words and 455 third person personal and possessive pronouns. This study helped us determine the distance of sentences from which the candidates will be extracted as well as the constraints and preferences for Turkish.

First question is "what will be the distance to search for the antecedent?". The study shows that 7% of the pronouns in the text have their antecedents in the same sentence as the pronoun, 61% of the pronouns have their antecedents in the previous sentence, 9% of the pronouns have their antecedents in two sentences back, and 4% of the pronouns have their antecedents in three sentences back of the sentence containing the pronoun. That is, totally antecedents of 81% of pronouns in the text have their antecedents in the current sentence or in previous three sentences. Going further back does not increase this percentage too much; therefore we consider only those proper person names in the current sentence and three previous sentences as the candidates for the antecedent of an extracted pronoun.

The constraints and preferences that are determined throughout the study are described below.

4.1.1. Constraints

A constraint defines a property that must be satisfied in order for any candidate to be considered as a possible solution for the anaphor (Palomar 2001). Therefore, constraints are used to discard those candidates that cannot be the antecedent of a considered pronoun.

Constraints are applied before preference rules during pronoun resolution to decrease the number of candidates. After the application of the constraints, if the number of candidates is decreased down to one, preference rules will not be needed to be used. Otherwise, the preference rules are applied to find the antecedent.

Constraints that are proposed for English include number and gender agreement, selectional restrictions, and c-command constraints (Mitkov 2002).

Number and gender agreement requires that the pronoun and the antecedent must agree in number and gender. Number agreement is applicable to Turkish however; gender agreement is not applicable since the third person pronoun does not denote gender.

Selectional restrictions require that semantic restrictions that are applicable to the anaphor should apply to the antecedent as well (Mitkov 2002). This constraint is applicable to Turkish but since it requires considerable semantic knowledge, it is not implemented in the system.

C-command constraints are syntactic constraints on how noun phrases may corefer, which are imposed by Binding Theory (Chomsky 1981; Chomsky 1995). C-command relation can be described as follows: A node *A* c-commands a node *B* if and only if (Mitkov 2002):

1. *A* does not dominate *B*
2. *B* does not dominate *A*
3. The first branching node dominating *A* also dominates *B*.

One of the c-command constraints is that a reflexive anaphor must be c-commanded by its antecedent and another c-command constraint is that a pronoun cannot refer to a c-commanding NP within the same local domain. C-command constraints are applicable to Turkish; however, due to the knowledge-poor nature of the system, they are adapted to Turkish as two constraints: reflexive pronoun constraint and personal pronoun constraint. Details of these constraints are provided in the following sections.

As a result of the empirical study on the sample Turkish text, the following four constraints for Turkish are found to be applicable to Turkish, and they are explained in detail below:

1. Number agreement
2. Reflexive pronoun constraint
3. Personal pronoun constraint
4. Selectional restrictions

4.1.1.1 Number Agreement

Number agreement requires that the anaphor and its antecedent must agree in number. The constraint is applicable to many languages including English, Spanish and French. It helps filtering out plural candidates when a singular pronoun is to be resolved and similarly, filtering out singular candidates when resolving plural pronouns.

Ayşe$_i$ *okula gitti. [Ahmet ve Fatma]*$_k$ *onu*$_i$ *gördü.* $\varnothing_k$ *Ona*$_i$ *el salladılar.*

Ayşe school-DAT go-PAST. Ahmet and Fatma she-DAT see-PAST. She-ACC hand wave-PAST-PERS

‘Ayşe went to school. Ahmet and Fatma saw her. (They) waved hand to her.’

4.1.1.2 Reflexive Pronoun Constraint

This constraint is an adaptation of c-command constraints used in many anaphora resolution algorithms. The constraint is adapted as follows to be utilized with minimum amount of knowledge: reflexive pronoun constraint requires that the antecedent of a reflexive pronoun is the closest candidate to the pronoun.

Ali$_i$ *kendine*$_i$ *güvenir.*

Ali himself-DAT trust-AOR

‘Ali trusts himself.’

In Turkish, there are two forms of reflexive pronouns both of which denote himself/herself/itself: *kendi* and *kendisi*. However, our empirical studies show that there is a slight difference between these two forms. When the reflexive pronoun is in the form of *kendi*, the antecedent is the closest candidate to the anaphor whereas if the reflexive pronoun is in the form of *kendisi*, if there is only one candidate in the considered sentence, that candidate is proposed as the antecedent as in the case of *kendi*, if there are multiple candidates in the considered sentence, among the candidates the most probable antecedent is the one which is closer to the beginning of the sentence.

Ayla$_k$, Fatma$_i$'nın kendine$_i$ güvendiğini bilir.

Ayla, Fatma-GEN herself-DAT trust-NOM-POSS-ACC know-AOR

'Ayla knows that Fatma trusts herself.'

Ayla$_i$ Fatma$_k$'nın kendisine$_i$ güvendiğini bilir.

Ayla, Fatma-GEN herself-DAT trust-NOM-POSS-ACC know-AOR

'Ayla knows that Fatma trusts herself (her).'

4.1.1.3 Personal Pronoun Constraint

Personal pronoun constraint requires that in a simplex sentence, sentence with only one clause, the antecedent of a personal pronoun cannot exist in the same sentence as the pronoun. It filters out the candidates in the simplex sentence containing the considered personal pronoun.

Ayşe$_i$ onu$_k$ gördü.

Ayşe she-ACC see-PAST

'Ayşe saw her.'

In our study on the Turkish narrative, we see that among the 455 pronouns, 417 of them are personal pronouns. Among these 417 personal pronouns, only 10 (2.4%) of them have their antecedents in the same sentence as the anaphor. Since this percentage is low, in our system, we decided to make use of this constraint without checking the sentences' being simplex or not. In other words, if the considered pronoun is a personal pronoun, the candidates in the same sentence with the pronoun are filtered out due to personal pronoun constraint.

4.1.1.4 Selectional Restrictions

Selectional restrictions constraint requires that pronoun and its antecedent must satisfy the same semantic properties. If this constraint is applied, those candidates which do not possess the semantic property of the pronoun are filtered out.

> *Zehra$_k$ sessizce bekliyor, Ayla$_i$ şarkı söylüyordu. Herkes onun$_i$ sesini duydu.*
>
> Zehra silently wait-PROG-PAST, Ayla song sing-PROG-PAST. Everybody she-GEN voice-POSS-ACC hear-PAST.
>
> ‘Zehra was waiting silently, Ayla was singing. Everybody heard her voice.’

In this example, since somebody who is silent cannot be heard, selectional restriction filters out the candidate *Zehra* which cannot be the antecedent of the pronoun.

Since this constraint requires considerable semantic knowledge, it is not used in the knowledge-poor system for Turkish.

4.1.2. Preferences

In knowledge-poor systems, after the application of the constraints, preferences rules are used to sort the remaining candidates. Each preference rule usually has a score associated with it which is added to the scores of the candidates satisfying the preference to determine the overall score of a candidate. At the end of this procedure, the candidate with the highest total score is proposed as the antecedent.

At the end of the empirical study on the representative Turkish text, the following preference rules are determined for Turkish:

1. Quoted/Unquoted Text Preference
2. Recency Preference
3. Subject Preference
    a. Nominative Case Preference

4. First Noun Phrase Preference
5. Nominal Predicate Preference
6. Repetition Preference
7. Punctuation Preference
8. Antecedent of Zero Pronoun Preference
9. Syntactic Parallelism Preference

Some of these preference rules had already been used in knowledge-poor anaphora resolution systems for other languages. For instance, first NP preference has already been used for Bulgarian (Tanev and Mitkov 2002), nominal predicate preference is similar to the 'existential emphasis' preference used for English (Lappin and Leass 1994; Kennedy and Boguraev 1996), recency, repetition, syntactic parallelism and subject preferences are extensively used for different languages (Lappin and Leass 1994; Mitkov 1998; Trouilleux 2002). Some of the preference rules are determined in this empirical study which had not been used in the anaphora resolution systems for other languages. These preferences are namely, punctuation, antecedent of zero pronoun and quoted/unquoted text preferences. The details of these preferences are provided in the following sections.

Among these preference rules, syntactic parallelism preference is not implemented in our system. In addition to this, instead of subject preference, nominative case preference is implemented.

In most knowledge-poor pronoun resolution algorithms, preferences have corresponding scores denoting the importance of the preference. Application of preferences to the candidate antecedents means assigning the score of the preference to the candidate if it satisfies the preference. The candidate with the highest aggregate score is proposed as the antecedent. In this system, the preferences are used in a similar way. The score assigned to each preference is determined by a learning system which uses a neural network to learn the optimal scores assigned to the preference rules. Details of this learning procedure are presented in subsequent sections.

Details of these preferences are described in the following sections. The preferences are presented in increasing score order.

4.1.2.1 Quoted/Unquoted Text Preference

If the pronoun considered is in quoted text, it is more likely that its antecedent is in quoted text also. Same way, if the pronoun considered is in unquoted text, it is more likely that its antecedent is in unquoted text also.

> *"Bugün Ayşe$_i$'yi gördüm" dedi Zerrin. "Ben de onu$_i$ dün görmüştüm" dedi Murat.*
>
> "Today Ayşe-ACC see-PAST-PERS" say-PAST Zerrin. "I too she-ACC yesterday see-PAST-PAST-PERS" say-PAST Murat.
>
> ' "(I) saw Ayşe today." said Zerrin. "I has seen her yesterday too" said Murat.'

In this example, the most salient entity to be the antecedent of the pronoun in the second sentence is *Ayşe*, since both the pronoun and *Ayşe* are in quoted text.

4.1.2.2 Recency Preference

Recency preference is given to the NPs that are in closer sentences to the sentence that contains the pronoun.

> *Ali oyun oynuyordu. Murat$_i$ da geldi. $\varnothing_i$ Oyunu sevdi.*
>
> Ali game play-PROG-PAST. Murat too come-PAST. Game-ACC
>
> like- PAST.
>
> 'Ali was playing a game. Murat came too. (He) liked the game.'

In this example, the candidate *Murat* is more salient than *Ali* since it resides in a more recent sentence that the other candidate *Ali*.

4.1.2.3 Subject Preference

This preference is given to the NPs at subject positions in sentences.

"Günaydın" dedi Murat$_i$. Ali ona$_i$ baktı.

"Good Morning" say-PAST Murat. Ali he-DAT look-PAST.

'"Good Morning" said Murat. Ali looked at him.'

In our knowledge-poor system, we do not use syntactic knowledge; therefore our system does not possess the ability to detect subjects, objects, verbs etc. in a sentence. However, we know that subjects are usually in nominative case and objects may or may not be in nominative case. Therefore, a 'nominative case preference' rule, in which preference is given to proper nouns in nominative case, may improve the success rate of our system and this preference is employed in our system.

4.1.2.4 First Noun Phrase Preference

This preference is given to an NP if it the first phrase in the sentence containing it.

Ahmet$_i$ Ali'yi gördü. $\varnothing_i$ Koştu.

Ahmet Ali-ACC see-PAST. Run-PAST.

'Ahmet saw Ali. (He) ran.'

In this example, *Ahmet* is more salient than *Ali*, since it is the first phrase in its sentence.

4.1.2.5 Nominal Predicate Preference

This preference is given to the NPs in the nominal predicates of this type of sentences.

Bu çocuk Ali$_i$'ydi. $\varnothing_i$ Sinirli görünüyordu.

This child Ali-PAST. Angry seem-PROG-PAST.

'This child was Ali. (He) seemed angry.'

In this example, *Ali* has increased salience to be the antecedent of the pronoun in the following sentence, since it is in the nominal predicate in the first sentence.

4.1.2.6 Repetition Preference

This preference is given to the NPs that are repeated in the text more than the other candidates.

Ayşe$_i$ parka gitti. ∅$_i$ Zeynep'le oyun oynadı. ∅$_i$ Şarkı söyledi.

Ayşe park-DAT go-PAST. With Zeynep game play-PAST. Song sing-PAST.

'Ayşe went to the park. (She) played game with Zeynep. (She) sang a song.'

In this example, the candidate *Ayşe* is repeated in the first and second sentences whereas *Zeynep* exists only in the second sentence, therefore, *Ayşe* is more salient than *Zeynep*.

4.1.2.7 Punctuation Preference

This preference is given to an NP if it has a comma following it. This comma increases the salience of NP that precedes it.

Yolda Tekin$_i$, Ali'ye seslendi. ∅$_i$ Çok yorgundu.

Way-LOC Tekin Ali-DAT call-PAST. Very tired-PAST

'On the way Tekin called Ali. (He) was very tired.'

In the above example, the comma following *Tekin* increases the salience of this proper person name.

4.1.2.8 Antecedent of Zero Pronoun Preference

If there exist consecutive sentences containing zero pronouns, then these zero pronouns usually refer to the same entity, which is the antecedent of the zero pronoun of the first sentence in this sequence. If the considered pronoun is a zero pronoun, *antecedent of zero pronoun preference* is given to the candidates that are the antecedents of zero pronouns in previous sentences.

∅ᵢ Eve yürüdü. ∅ᵢ Kapıda durdu. ∅ᵢ Kapıyı çaldı.

Home-DAT walk-PAST. Door-LOC stop-PAST. Door-ACC knock-PAST.

'(He) walked home. (He) stopped at the door. (He) knocked the door.'

In this example, the antecedent of the dropped (zero) pronoun in the very first sentence has increased salience to be the antecedent of the pronouns in the second and third sentences.

4.1.2.9 Syntactic Parallelism Preference

Syntactic parallelism preference is given to the NPs with the same syntactic function as the anaphor.

Ali Tekin$_i$'i yolda görmüş. Murat onu$_i$ okulda görmüş.

Ali Tekin-ACC way-LOC gör-AUX. Murat he-ACC school-LOC see-PAST.

'Ali saw Tekin on the way. Murat saw him at school.'

In this example, *Tekin* is a more salient candidate than *Ali* and *Murat* to be the antecedent of the pronoun in the second sentence, since it has a similar syntactic structure as the pronoun.

4.2. Questionnaire on Native Turkish Speakers

The results of the text analysis on the sample Turkish text are further verified by native Turkish speakers. This verification is completed by a questionnaire of 17 questions on 48 native Turkish speakers from different age, gender and job groups. This questionnaire is provided in Appendix A. Each question tests the validity of at least one of the 9 preference rules or 4 constraints stated above.

In (Keller 2000), it is pointed out that, to minimize biases, Schütze (1996) suggests a number of recommendations that should be applied when designing, applying and evaluating an acceptability judgment experiment. In our questionnaire, acceptability of examples sentences were our second concern, our main concern was to determine how native Turkish speakers resolve third person personal and reflexive pronouns referring to proper names. However, some of the recommendations provided in (Keller 2000) were perceived as applicable to our questionnaire and these are explained below.

In the procedure of gathering judgments, when selecting the subjects, linguists should be excluded as informants as their judgments may be confounded by theoretical bias (Keller 2000). Following this recommendation, we applied our questionnaire to subjects who are not linguists but native speakers of Turkish.

In (Keller 2000), it also is stated that the number of subjects used has to be large enough so that statistical test can be carried out on the data. This recommendation is also utilized in our study. The questionnaire is applied on 48 native speakers so that Cochran's Q test can be carried out on the data. The application of this test is explained in the following paragraphs.

Instructions that are given to the subjects may have an important influence on the judgment results, and in Schütze (1996), Schütze argues that the instructions should be as specific as possible in defining these terms, preferably making reference to relevant examples (Keller 2000). In our study, instructions are provided to the subjects however, these instructions are not based on any study in the literature.

As the last recommendation that is applicable to our questionnaire, it is pointed out that statistical methods should be used when evaluating the data gathered after the experiment (Keller 2000). Following this recommendation, statistical analysis of the questionnaire is carried out using Cochran's Q statistics.

Cochran's Q test can be used to evaluate the relation between two variables which are measured on a nominal scale. If the data from the research can be arranged in a two-way table consisting of N (subjects) rows and k (categories) columns, it is possible to test the null hypothesis that the proportion (or frequency) of responses of a particular kind is the same in each column, except for chance differences (Siegel and Castellan 1988). According to Cochran's Q test, if the null hypothesis is true, the categories are randomly distributed, and if the number of rows is not too small

$$Q = \frac{k(k-1) \sum_{j=1}^{k} (G_j - G^*)^2}{k \sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2}$$

is distributed approximately as chi square with df = k-1

where G_j = total number of “successes” in *j*th column,

G^* = mean of the G_j

L_i = total number of “successes” in the *i*th row. (Siegel and Castellan 1988)

Table 5.1 Usage of preference rules by subjects [Usage:1 Non-usage: 0]

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**L_i**
1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	4
2	1	1	0	1	1	0	1	1	1	0	0	0	0	0	1	1	1	10
3	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	5
4	1	1	0	1	1	0	0	1	0	0	0	0	0	0	1	0	1	7
5	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	4
6	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	4
7	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	15
8	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	14
9	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	13
10	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	11
11	1	1	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0	7
12	1	1	1	1	0	1	1	1	1	0	0	0	1	0	0	1	1	11
13	1	1	0	1	1	1	1	1	0	0	0	1	1	1	0	1	1	12
14	1	1	0	1	1	1	1	1	0	1	0	0	0	0	0	1	0	9
15	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	13
16	1	1	0	1	1	0	1	1	1	0	0	1	0	0	1	1	1	11
17	0	0	0	1	1	0	1	1	0	0	0	1	0	1	0	1	1	8
18	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	1	6
19	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	14
20	1	1	0	1	0	1	1	1	1	0	1	1	0	0	0	1	0	10
21	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	1	13
22	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	13
23	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	15
24	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	15
25	1	1	0	1	0	1	0	1	0	1	1	1	1	0	0	1	1	11
26	1	0	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0	6
27	1	1	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	6
28	1	1	0	1	1	1	0	1	0	0	0	0	1	0	0	1	0	8
29	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	1	1	7
30	1	0	0	1	1	0	1	1	1	0	1	1	1	0	0	0	1	10
31	1	1	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	13
32	1	0	1	1	1	1	0	1	1	1	0	0	0	1	0	0	1	10
33	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	15
34	1	1	0	1	0	1	0	0	0	0	0	0	1	0	1	0	1	7
35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17

36	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	12
37	1	0	1	1	0	0	0	1	0	1	1	0	1	1	0	0	0	8
38	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	0	1	13
39	1	0	0	1	1	1	0	1	0	0	0	0	1	1	0	0	1	8
40	0	0	1	1	0	0	1	1	0	1	1	1	1	0	0	1	1	10
41	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	15
42	0	0	0	1	0	0	1	1	0	0	0	0	0	0	1	1	1	6
43	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	15
44	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	0	1	10
45	0	0	1	1	1	0	0	1	1	1	0	0	1	1	0	1	1	10
46	1	1	0	1	0	0	0	1	0	1	1	1	1	0	0	1	0	9
47	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	14
48	0	0	1	1	0	0	0	1	0	0	1	1	0	0	0	1	0	6
Gj	40	32	22	48	31	28	31	43	21	19	20	22	26	22	15	34	36	

Table 5.1 shows the results of the questionnaire of 17 questions given as columns on 48 native Turkish speakers given as rows. In this table, columns denote each of the 17 (k) questions of the questionnaire and the rows denote the human subjects (N). A value 1 in a cell means that the i^{th} subject made use of the preference or the constraint which is applicable in the j^{th} question when answering this question, whereas a 0 means that the subject did not make use of the corresponding constraint or preference.

The sampling distribution of Q is approximated by the chi-square distribution with df = k – 1. Q is equal to the value of chi-square if there is only chance difference between variables, if Q is greater than the value of chi-square then it means that there is a statistical association between the variables.

Table 5.2 Chi-square for df = 16

df	p = 0.30	p = 0.20	p = 0.10	p = 0.05	p = 0.02	p = 0.01	p = 0.001
16	18.42	20.46	23.54	26.30	29.63	32.00	39.29

Using the data presented in Table 5.1, Q is calculated as 135. The degree of freedom is k – 1 = 16, if we check the value of the Chi square table for this value in Table 5.2, we see that the result will be significant if Q is greater than 39.29 for p = 0.001. Since Q is 135, this result is highly significant.

The number of cases necessary for a preference rule to be considered as applicable can also be found using a partial Cochran's Q method. Partial Cochran's Q method requires the calculation of different Q values for different G_i (0 to 48) values. In our study, for $G_i >= 19$, Q_i is found significant at the $p = 0.001$ level. Therefore, we conclude that 19 among 48 subjects should make use of the corresponding preference or constraint in a particular question for the constraint or preference to be considered as applicable for Turkish.

Using the findings of the partial Cochran's Q method, our 4 constraints and 9 preference rules described before, are verified by the human subjects since these constraints and preferences are used by more than 19 subjects. However, the preference rule in the 15th question is not verified by human subjects. 15th question was testing the validity of a preference which can be named as 'accusative case preference'. This preference is not described previously, since it is not verified by the native Turkish speakers. 'Accusative case preference' can be applied by giving preference to NPs in accusative case over NPs in dative case. There are no other preference rules which are not verified by the results of the questionnaire and all the constraints and preferences that are verified are described in the previous section.

4.3. Determination of the Preference Scores

The application of preference rules in our knowledge-poor pronoun resolution system is the assignment of scores of the preference rules to the candidates which satisfy these rules. Then the aggregate score of a candidate is calculated by summing the scores assigned to it. The candidate with the highest aggregate score is proposed as the antecedent of the pronoun.

The determination of scores of the preferences is a very important task. If the optimal values are not used, this may lead the system to propose incorrect candidates as antecedents. The optimal values for the scores can be best approximated by employing a machine learning approach so that the scores can be learned.

Most of the existing knowledge-poor anaphora resolution algorithms use empirical observations to determine the preference scores (Lappin and Leass 1994; Kennedy and Boguraev 1996; Mitkov 1998). The optimization of preference scores is applied to the

Mitkov's original approach which was described in Mitkov (1998) using a genetic algorithm (Orasan and Evans 2000).

In this system, the score assigned to each preference rule is determined by training a neural network (a perceptron). The inputs of the perceptron are the preference rules for Turkish determined by empirical studies and weights of these inputs are the scores of the preference rules.

The training of the perceptron is performed using delta rule. According to the delta rule algorithm, at the beginning, each weight is initialized to some random value, in our case this value is +1. The learning rate is taken as 0.05 and the threshold is taken as +5.

Since in our system reflexive pronouns are resolved using the reflexive pronoun constraint, preference rules are utilized only for personal pronouns. Therefore we trained our perceptron using 20 examples, taken from the questionnaire that is applied to native Turkish speakers and the sample Turkish text that is analyzed, each including a personal pronoun.

After this training phase, the weights of the perceptron are taken as the scores of the preference rules. These final scores of the preference rules are presented in Table 5.3.

Table 5.3 Preference rules and corresponding scores.

Preference Rule	**Score**
Quoted/Unquoted Text Preference	+2.20
Recency Preference	+2.15
Nominative Case Preference	+1.85
First NP Preference	+1.40
Nominal Predicate Preference	+1.20
Repetition Preference	+1.20 if repetition is more than once
Punctuation Preference	+1.15
Antecedent of Zero Pronoun Preference	+1.05

CHAPTER 5

DESIGN AND IMPLEMENTATION OF THE PRONOUN RESOLUTION SYSTEM FOR TURKISH

Pronoun resolution system for Turkish takes a Turkish text which is preprocessed manually to mark the considered pronouns and the system outputs a rewritten version of the input text in which the considered pronouns are replaced with their antecedents.

Design and implementation of the system will be described in two sections. In the first section, the preprocessing step is explained. The input to the pronoun resolution system for Turkish should be preprocessed to mark the considered third person personal and reflexive pronouns that refer to proper person names. In the second section, the architecture of the system will be presented. The system is composed of four modules: sentence splitting, pronoun extraction, forming the list of candidate antecedents and determination of the antecedent.

5.1. Preprocessing to Mark Overt and Zero Pronouns

Turkish is a pro-drop language, that is, pronouns in subject and object positions may be omitted but they are nevertheless understood. Our pronoun resolution system will resolve third person personal pronouns and reflexive pronouns that refer to proper person names including overt pronouns as well as zero pronouns at subject positions.

By employing a preprocessing procedure, the pronoun resolution system is freed from the burden of resolving the forms of '*o*' which are non-anaphoric. There are two cases in Turkish where a form of '*o*' is used non-anaphorically:

1. The third person personal pronoun, *o*, can be used as a demonstrative in adjectival position.

O *kitabı* *görmedim.*

That book-DAT see-NEG-PAST-PERS

(I) didn't see that book.

2. The dative case of third person personal pronoun *ona* (he/she-DAT) can be used non-anaphorically in the idiomatic phrases like *sözüm ona* (seemingly).

   Sözüm ***ona****, Ayşe, Ayla'yı kandıracak.*

   Seemingly, Ayşe Ayla-ACC deceive-FUT

   'Seemingly, Ayşe will deceive Ayla.'

In order for the system to detect the considered pronouns, the input text is processed by the author to mark considered overt pronouns and to bring zero pronouns to surface and mark them. In the preprocessing step, *[o]* (will be referred to as 'overt pronoun sign') is appended at the end of considered overt pronouns, the considered zero pronouns are brought to the surface and *[z]* (will be referred to as 'zero pronoun sign') is appended at the end of them to mark them accordingly. The following text is taken from Metu Turkish Corpus (Say et al. 2002) and preprocessed as described above.

Original text:

Haris, annesiyle Ayşe'yi hayretle kendisine bakar durumda bıraktıktan sonra Muhsinler'in evine koştu. Cebinde parası da vardı. Haris, okulda kimseciklerin bilmediği bir işte çalışıyor; hafta sonlarında evlere sabah gazetelerini dağıtıyordu. Kazandığı parayı harcamıyor, biriktiriyordu. Muhsin, Haris'i kapıda beklemiyordu. Zili çaldığında, onun yerine annesi çıktı. Muhsin'in, bu akşam şehre gitmek üzere, az önce evden ayrıldığını söyledi. Haris, pencereden atlamak zorunda kalmamanın sevinciyle deliye dönmüş olduğundan, buluşacakları yerin Muhsinler'in evi olmadığını unutmuştu.

Oysa Muhsin, onu okul kütüphanesinin önünde beklemekteydi. Haris ona, evden nasıl ayrıldığını anlattı.

‘Haris ran to Muhsin’s house after leaving (his) mother and Ayşe looking at him in surprise. There was money in (his) pocket. Haris was doing a job that nobody at school knows about; carrying morning papers to houses at weekends. He was not spending but saving the money (he) earned. Muhsin was not waiting for Haris at the door. When he rang the bell, (his) mother appeared instead of him. She said that Muhsin had just left the house to go to the city that night. Haris, having been very happy for not having to jump from the window, had forgotten that the place that (they) would meet was not Muhsin’s house. But Muhsin was waiting for him in front of the city library. Haris told him how (he) left the house.’

Preprocessed version of the original text:

Haris, annesiyle Ayşe'yi hayretle ***kendisine[o]*** *bakar durumda bıraktıktan sonra Muhsinler'in evine koştu. Cebinde parası da vardı. Haris, okulda kimseciklerin bilmediği bir işte çalışıyor; hafta sonlarında evlere sabah gazetelerini dağıtıyordu.* ***O[z]*** *Kazandığı parayı harcamıyor, biriktiriyordu. Muhsin, Haris'i kapıda beklemiyordu.* ***O[z]*** *Zili çaldığında,* ***onun[o]*** *yerine annesi çıktı. Muhsin'in, bu akşam şehre gitmek üzere, az önce evden ayrıldığını söyledi. Haris, pencereden atlamak zorunda kalmamanın sevinciyle deliye dönmüş olduğundan, buluşacakları yerin Muhsinler'in evi olmadığını unutmuştu. Oysa Muhsin,* ***onu[o]*** *okul kütüphanesinin önünde beklemekteydi. Haris* ***ona[o]****, evden nasıl ayrıldığını anlattı.*

‘Haris ran to Muhsin’s house after leaving (his) mother and Ayşe looking at **him[o]** in surprise. There was money in (his) pocket. Haris was doing a job that nobody at school knows about; carrying morning papers to houses at weekends. **He[z]** was not spending but saving the money (he) earned. Muhsin was not waiting for Haris at the door. When **he[z]** rang the bell, (his) mother appeared instead of **him[o]**. She said that Muhsin had just left the house to go to the city that night. Haris, having been very happy for not having to jump from the window, had forgotten that the place that (they) would meet was not Muhsin’s house. But Muhsin was waiting for **him[o]** in front of the city library. Haris told **him[o]** how (he) left the house.’

The words in parenthesis in English translation are those pronouns that are dropped in original Turkish text above; however, since they are not at subject positions in the original text, they are not marked during preprocessing.

5.2. Architecture of Turkish Pronoun Resolution System

After the input text is preprocessed, the resulting text is given as input to the actual pronoun resolution system. The implementation steps of the system are:

- Splitting the input text into sentences
- Extraction of the third person pronouns and reflexive pronouns from the text
- Creating the candidate list of antecedents for each of the extracted pronouns
- Determining the antecedent of each extracted pronoun from left to right by applying constraints and preferences.

These steps are described in the following sections in detail.

5.2.1. Sentence Splitting

The first step of a pronoun resolution algorithm is to split the input text into sentences. For our pronoun resolution system, dot (.), three dots (…), exclamation mark (!) and question mark (?) are used as sentence separators. A sentence splitter that uses these separators is implemented. The output of the sentence splitter is a sequence of sentences each having a unique sequence number.

5.2.2. Pronoun Extraction

In the preprocessing step, third person personal pronouns and reflexive pronouns that refer to proper nouns are marked according to their being overt or zero. Pronouns that are preprocessed in this step are the pronouns that will be resolved by our system. Those words in the input text which are marked with overt and zero pronoun signs will be extracted to be resolved in pronoun extraction step.

5.2.3. Forming the List of Candidate Antecedents

In a knowledge-poor pronoun resolution system, after a pronoun is extracted from the input text, a list of candidates for the antecedent of corresponding to this pronoun has to be formed. The search scope of candidates for the antecedents of pronouns varies in existing proposals for different languages. For instance in Mitkov's robust pronoun resolution system for English, candidates are taken from the current and two preceding sentences. For pronominal anaphors, the search scope is usually limited to the current and two or three preceding sentences (Mitkov 2002).

In the analysis phase of our pronoun resolution system, we empirically analyze a sample Turkish text and find out that candidate antecedents for a pronoun should be taken from the current sentence and three preceding sentences. Therefore, in this system, when a pronoun is extracted from input text, the proper nouns in the current sentence and in three preceding sentences are used to form the candidate list for this anaphor. In the current sentence, proper nouns to the left of the considered pronoun are extracted as candidates, that is, our system does not try to resolve cataphora, in which the pronoun may precede its antecedent.

Proper nouns are easily identifiable in Turkish texts since they are the only words capitalized in a sentence except for the sentence initial words. However, we have to extract only those proper nouns which are person names since we consider only personal and reflexive pronouns. Also, since sentence initial words in each sentence are capitalized, it is not easy to determine whether a word at the sentence initial position is a proper person name or an ordinary name. In order to accomplish the task of extracting person names, we use a Turkish person names dictionary consisting of 9060 names. This dictionary is very useful to our study; however there are still problems because of those person names which may well be used as ordinary words. For instance, ‘Cesur’ may be used as a male name, also as an adjective since it means ‘brave’. Similar proper names problems in English are also explained in Mitkov (2002).

One important point to note here is the case of plural pronouns. The system expects the candidates of plural pronouns to include the plural suffix or consist of multiple proper nouns joined with ‘ve’ (and) or ‘*ile*’ (with). The following are extracted as candidates of plural pronouns:

(1) *Mehmetler* (usually denotes ‘*Mehmet and his family*’)

‘Mehmet-PLU’

(2) *Ali ve Mehmet*

‘Ali and Mehmet’

(3) *Ali ile Mehmet*

‘Mehmet with Ali’

(4) *Ali’yle Mehmet* (short for ‘*Ali ile Mehmet*’)

‘Mehmet with Ali’

When the system cannot find candidates in the above forms, set generation is used. Set generation creates set-level referents that can serve as antecedents for plural pronouns (Rich and Luperfoy 1988). The proper person names in the same sentence are joined with *ve* (‘*and*’) forming a set. This set is extracted as a candidate for plural pronouns for which candidates in the expected forms do no exist. For instance, the set ‘*Ali ve Mehmet*’ is generated and extracted as a candidate for the plural pronoun in the following sentences. In the second sentence, the considered plural personal pronoun ‘*onlar*’ is marked with the zero pronoun sign (*[z]*) in the preprocessing step, since it is a dropped (zero) pronoun at subject position.

`Ali, Mehmet'le karşılaştı.` **`Onlar[z]`** `Birlikte yürüdüler.`

`Ali, Mehmet-WITH meet-PAST. They[z] Together walk-PAST-PERS.`

`'Ali met Mehmet. They walked together.'`

5.2.4. Determination of the Antecedent

Determination of the antecedent of a pronoun is the final step in the pronoun resolution process. Similar to the existing knowledge-poor pronoun resolution proposals for other languages, at this step our system will apply the constraints and preferences, which were determined in the analysis step, to the candidate antecedents to resolve a pronoun.

The constraints for Turkish include number agreement, reflexive pronoun constraint, personal pronoun constraint and selectional restrictions. In this knowledge-poor system, we implemented number agreement, reflexive pronoun constraint and personal pronoun constraint. As stated before, we assume the personal pronoun constraint to hold true for all kinds of sentences instead of only simplex sentences. The system will discard those candidates which do not agree in number with the pronoun. If the pronoun is reflexive, it will keep only those candidates in the current sentence whereas if the pronoun is personal, it will filter out the candidates that are in the same sentence with the pronoun. After the constraints are applied, if there is no candidate in the search scope, then the system cannot resolve this pronoun and reports the pronoun as "ambiguous". If only one possible candidate is available, this candidate is the antecedent. If more candidates exist, then preferences are applied to the remaining candidates.

Among the preference rules for Turkish determined during analysis, syntactic parallelism preference and subject preferences are not implemented in our system since these preferences require considerable linguistic and semantic knowledge. However, in order to improve the success rate of the system, a special case of subject preference, namely nominative case preference is implemented in our system. By doing this, our system makes partial use of the subject preference at the same time keeping its knowledge-poor nature. Therefore, the preference rules considered by our system are: quoted/unquoted text, recency, nominative case, first noun phrase, nominal predicate, repetition, punctuation, and antecedent of zero pronoun preferences.

Each preference rule has an associated score determined during analysis. Each rule is applied to each candidate. If a preference rule is satisfied by the candidate, the corresponding score is added to the aggregate score of the candidate, if not, then the aggregate score is not changed. After the application of all the considered preferences, the candidate with the highest aggregate score is selected as the antecedent. In case of a tie, the candidate which is more recent is taken as the antecedent.

The system takes part of Turkish text as input, which is preprocessed manually to mark the considered overt and zero pronouns as described above. In the text, these pronouns are replaced with their antecedents and this modified version of the text is given as the output.

When the pronouns are replaced with their antecedents, the system formats the antecedent so that the pronoun and its antecedent have the same case. To illustrate, if '*onu*' (he-ACC) is the considered pronoun and its antecedent is '*Ali*', then the system replaces '*onu*' with '*Ali'yi*' (Ali-ACC). By doing this, the coherence of the output text is maintained.

For instance, if the following text, which is preprocessed manually, is given as input to this system:

> Haris, pencereden atlamak zorunda kalmamanın sevinciyle deliye dönmüş olduğundan, buluşacakları yerin Muhsinler'in evi olmadığını unutmuştu. Oysa Muhsin, **onu[o]** okul kütüphanesinin önünde beklemekteydi. Haris **ona[o]**, evden nasıl ayrıldığını anlattı.
>
> *'Haris, having been very happy for not having to jump from the window, had forgotten that the place that (they) would meet was not Muhsin's house. But Muhsin was waiting for **him[o]** in front of the city library. Haris told **him[o]** how (he) left the house.'*

The output of the system will be as follows:

Haris, pencereden atlamak zorunda kalmamanın sevinciyle deliye dönmüş olduğundan, buluşacakları yerin Muhsinler'in evi olmadığını unutmuştu. Oysa Muhsin, **Haris'i** okul kütüphanesinin önünde beklemekteydi. Haris **Muhsin'e** evden nasıl ayrıldığını anlattı.

*'Haris, having been very happy for not having to jump from the window, had forgotten that the place that (they) would meet was not Muhsin's house. But Muhsin was waiting for **Haris** in front of the city library. Haris told **Muhsin** how (he) left the house.'*

For the first pronoun '*onu*', candidates are *Haris* and *Muhsinler*. With the application of the number agreement constraint, *Muhsinler* (a plural form denoting the family of Muhsin) is filtered out. Therefore, the only candidate *Haris* is the antecedent. For the second pronoun '*ona*', candidates are *Haris*, *Muhsinler* and *Muhsin*. As in the case of the first pronoun considered, *Muhsinler* is filtered out due to the number agreement constraint. *Haris* is filtered out due to the personal pronoun constraint that we assume to hold true most of the time. The only candidate, *Muhsin*, is proposed as the antecedent after the application of the constraints.

If we consider the following two sentences:

Ali yolda Mehmet'i gördü. Ona selam verdi.

'Ali saw Mehmet on the way. (He) greeted him.'

After they are manually preprocessed to mark the considered overt and zero pronouns with overt pronoun sign (*[o]*) and zero pronoun sign (*[z]*) respectively, they are given as input to the system in the following form:

Ali yolda Mehmet'i gördü. **O[z] Ona[o]** selam verdi.

*'Ali saw Mehmet on the way. **He[z]** greeted **him[o]**.'*

The output of the system is:

Ali yolda Mehmet'i gördü. **Ali, Mehmet'e** selam verdi .

‘Ali saw Mehmet on the way. ***Ali*** *greeted* ***Mehmet****.’*

For the first zero pronoun, the candidates are *Ali* and *Mehmet*, and none of them is filtered out during the application of the constraints. Therefore, preference rules are applied to these candidates, and their scores are computed as follows:

For *Ali*:

Quoted/Unquoted Text Preference (Both the pronoun and *Ali* is unquoted) : +2.20

Recency Preference (*Ali* is in the previous sentence) : +2.20

Nominative Case Preference (*Ali* is in nominative case) : +1.85

First NP Preference (*Ali* is the first NP in its sentence) : +1.40

The total score for *Ali* is +7.65.

For *Mehmet*:

Quoted/Unquoted Text Preference (Both the pronoun and *Mehmet* are unquoted) : +2.20

Recency Preference (*Mehmet* is in the previous sentence) : +2.20

The total score for *Mehmet* is +4.40.

Since the candidate *Ali* has a higher aggregate score, it is proposed as the antecedent of the first pronoun.

When resolving the second pronoun, the candidates are again *Ali* and *Mehmet*, but this time, due to the simplex sentence constraint, since *Ali* is replaced with the first pronoun considered, *Ali* is filtered out and the only candidate *Mehmet* is selected as the antecedent for the second pronoun.

Knowledge-poor pronoun resolution system for Turkish which executes as described above is implemented using Java programming language. The development environment is the Eclipse Workbench which is an open-source platform. A graphical user interface (GUI) is developed for the system using the Swing package of Java. This GUI enables the user to input a text file in the local machine containing the preprocessed text to be resolved, the output of the system is written to a text file in the same path as the input file. GUI also enables the user to enter the text to be resolved to an input text area and the output text is written to an output text area.

CHAPTER 6

TESTING

In order to test the performance of the presented knowledge-poor pronoun resolution system, two sample texts from different sources are used. As described in the first section of Chapter 5, input texts are preprocessed manually to mark the overt and zero pronouns that the system will consider before it is given as input to the system. In order to compare the performance of the system, two different baseline algorithms are implemented. The first baseline algorithm applies only the constraints to the candidates and resolves a pronoun if it has a single candidate remaining after the application of the constraints. The second baseline algorithm selects the most recent candidate as the antecedent after the application of constraints. Similar baseline algorithms had already been used in most proposals of knowledge-poor pronoun resolution algorithms for languages including English, Spanish and French, so they are implemented and used in the presented knowledge-poor system for Turkish. These baseline implementations take their input text after it is manually preprocessed like the original knowledge-poor algorithm.

In the first experiment, the system is evaluated on a sample text from Metu Turkish Corpus (Say et al. 2002). This sample text which was taken from the corpus was a narrative by Aysel Kumru Korkut called ‘Kimse Beni Anlamıyor’ (‘Nobody Understands Me’). The text contains 4140 words with 190 marked pronouns after the preprocessing step. 20 of these pronouns were reflexive and 170 of them were personal pronouns, also, 67 of the pronouns were overt and 123 of them were zero pronouns. After the application of the constraints, it is observed that 90 of 190 pronouns had single candidate in their candidate lists, that is, 90 pronouns were resolvable without the application of the preference rules.

Both the knowledge-poor algorithm for Turkish and baseline algorithms are tested on this sample text. The baseline algorithm using only the constraints resolved 90 of 190 pronouns (47.3%) correctly and the baseline algorithm favoring the most recent candidate resolved 130 of the 190 pronouns (68.4%) correctly. Finally, knowledge-poor algorithm resolved 162 of the 190 pronouns (85.2%) correctly. These results are shown in Table 6.1. The metrics used in this testing phase are success rate (recall) and precision which are calculated with the following formulas:

Success rate (Recall) = Number of pronouns correctly resolved /
Number of pronouns identified

Precision = Number of pronouns correctly resolved / Number of pronouns attempted

Table 6.1 Success Rate (Recall) and Precision of the Implemented Algorithms in the First Experiment

	Baseline algorithm using only the constraints	Baseline algorithm favoring the most recent candidate	Knowledge-poor algorithm
Success rate (Recall)	47.3%	68.4%	85.2%
Precision	100%	70.6%	88%

The second sample that is used to test the system is a Turkish child narrative (Ilgaz 2003b) of 11315 words which include 190 personal pronouns and 15 reflexive pronouns that refer to proper person names, therefore, totally 205 pronouns that the system will attempt to resolve. 156 of these 205 pronouns are zero pronouns and 49 of them are overt pronouns. The baseline algorithm using only the constraints correctly resolved 100 pronouns (48.7%), the baseline algorithm proposing the most recent candidate correctly resolved 135 pronouns (65.8%) and finally the knowledge-poor algorithm correctly resolved 151 of 205 (73.6%) considered pronouns. The results of testing the knowledge-poor and baseline algorithms on this sample text are presented in Table 6.2.

Table 6.2 Success Rate (Recall) and Precision of the Implemented Algorithms in the Second Experiment

	Baseline algorithm using only the constraints	Baseline algorithm favoring the most recent candidate	Knowledge-poor algorithm
Success rate (Recall)	48.7%	65.8%	73.6%
Precision	100%	81.3%	90.9%

The knowledge-poor system correctly resolved 162 of the 190 pronouns in the first experiment. When the cases where the algorithm fails are analyzed, it is seen that 15 of the 28 incorrect resolutions are due to the personal pronoun constraint that is employed by the algorithm. In 6 of the remaining 13 cases, a candidate is not found in the current and three preceding sentences. In the analysis step, the probability that the antecedent is in the current and three preceding sentences was found as 81%. The remaining 7 incorrect resolutions can be attributed to multiple reasons including the extraction of non-proper names as candidates and semantic reasons.

In the second experiment on the second sample, the original system correctly resolved 151 of 205 pronouns. 39 of the 54 failures of the system are due to the non-existence of antecedent in the current and three preceding sentences. The incorrect resolution of the remaining 15 cases is due to the reasons stated for the first experiment.

These two text samples used in the experiments are chosen due to their having sufficient number of pronouns that our knowledge-poor pronoun resolution system considers. It could well be tested on other samples including newspaper articles, but since these samples might not contain many pronoun occurrences that refer to proper person names, testing results could be misleading. Metu-Sabancı Turkish Treebank (Oflazer et al. 2003) is also examined to test the system on samples from this treebank. However, there were not sufficient example samples in the treebank containing reflexive and third person personal pronouns that refer to proper person names. Therefore, it can be stated that our system for Turkish is not genre-specific, since it can be used for any sample Turkish texts of any genre.

CHAPTER 7

DISCUSSION

Knowledge-poor pronoun resolution algorithms developed for languages other than Turkish have been presented in Chapter 3. The success rates of these algorithms, as reported in their respective studies, cannot be compared, since they are developed for different languages and those which are developed for the same language like English are not evaluated on the same test corpus. However, to summarize the information provided in Chapter 3, success rates of the systems developed for other languages including English, French, Spanish and Bulgarian are presented with the success rate of the knowledge-poor pronoun resolution system for Turkish in Table 7.1.

Table 7.1 Success rates of different knowledge-poor pronoun resolution systems as reported in respective studies.

Pronoun Resolution System	**Success Rate**
Baldwin's System for English (Baldwin 1996)	77.9%
Kennedy and Boguraev's System for English (Kennedy and Boguraev 1996)	75%
Lappin and Leass's System for English (Lappin and Leass 1994)	86%
Mitkov's System for English (Mitkov 1998)	89.7%
Palomar et al.'s System for Spanish (Palomar et al. 2001)	76.8%
Trouilleux's System for French (Trouilleux 2002)	74.8%
Tanev and Mitkov's System for Bulgarian (Tanev and Mitkov 2002)	75%
Knowledge-poor System for Turkish (on sample text 1)	85.2%
Knowledge-poor System for Turkish (on sample text 2)	73.6%

The presented knowledge-poor resolution system for Turkish is different from the earlier proposals for other languages in the following aspects:

- The considered reflexive and third person personal pronouns refer to proper person names.

- The pronouns satisfying the above condition are marked manually in the preprocessing step, so, there were no difficulty in pronoun extraction stage of the algorithm. Therefore, there are no cases where this system attempts to resolve a pronoun not satisfying the above constraint.

In addition to the above points, the constraints and preferences for Turkish are not exactly the same with that of other knowledge-poor anaphora resolution systems in other languages.

As for constraints, gender agreement is not applicable for Turkish whereas this constraint is commonly used for other anaphora resolution systems. Number agreement is an applicable constraint for Turkish and is used in our system. The remaining two constraints, namely, reflexive pronoun constraint and personal pronoun constraint, are determined by modification and adaptation of the c-command constraints for Turkish as explained in the first section of Chapter 4.

Concerning the preference rules, some of the preference rules employed in this system had already been used in the systems for other languages. As described in Chapter 4, first NP preference has already been used for Bulgarian (Tanev and Mitkov 2002), nominal predicate preference is similar to the 'existential emphasis' preference used for English (Lappin and Leass 1994; Kennedy and Boguraev 1996), recency and repetition are extensively used for different languages (Lappin and Leass 1994; Mitkov 1998; Trouilleux 2002). Apart from these common preferences, three remaining preferences, namely, quoted/unquoted text, antecedent of zero pronoun, and, punctuation preferences are not common and they are used in our system. Usually, in other knowledge-poor anaphora resolution systems, quoted text is not dealt with; therefore these systems do not employ a preference rule to deal with quoted text. In this system for Turkish, a quoted/unquoted text preference rule is employed just to make a distinction between antecedents in quoted and unquoted text when a pronoun to be resolved is in quoted or unquoted text. Punctuation preference is not very common for languages other than Turkish, in Turkish comma can be used to emphasize the noun phrase (NP) preceding it in a sentence, which increases the salience of this NP. In addition, antecedent of zero pronoun preference is not used for other languages since it is applicable to languages in only pro-drop languages. Since pro-drop is commonly used in Turkish texts, the pronoun resolution system for Turkish makes use of the preference.

In future studies, the domain of the system may be extended and the success rate of the system may be improved. We restricted our system as a pronoun resolution system that resolves third person personal pronouns and reflexive pronouns that refer to proper person names, as a further study, the system may be extended and improved in the following directions:

- The system may be extended to resolve pronouns with noun phrase (NP) antecedents by making use of an NP extractor for Turkish. The NP extractor can be used to extract the NPs preceding the pronoun which will be used as candidates. This extension cannot be made at the moment since there is no available NP extractor for Turkish.

- The pronoun extraction module can be improved by employing a part-of-speech (POS) tagger to discard non-anaphoric occurrences of overt pronouns, so that only anaphoric occurrences of the overt pronouns will be taken into account. Also by using a parser together with the POS tagger, dropped third person pronouns can detected without human intervention. With this and the previous extension, there will be no need for the preprocessing of the input text and the system will be a fully-automated system. Since there is no available POS tagger and parsers for Turkish that can be integrated to the system to make it fully automated, this extension cannot be made at the moment.

- Subject preference and syntactic parallelism preference rules can be implemented to increase the success rate of the system by integrating NLP tools such POS taggers and shallow parsers to the system. Similar to the previous extensions, this extension is not possible at the moment due to the unavailability of the stated NLP tools.

- Personal pronoun constraint can be improved by making use of a parser to determine whether a sentence is a simplex sentence or not. In the system presented, personal pronoun constraint requires that the antecedents of the personal pronouns are not in the same sentence with that of the pronouns.

However, in a sentence which is not simplex, the antecedent of a personal pronoun may well be in the same sentence with that of the pronoun. If a Turkish parser to detect whether a sentence is simplex or not, were available, the personal pronoun constraint could be improved so that the candidates for antecedents of personal pronouns can be extracted from the sentence containing the pronoun.

CHAPTER 8

CONCLUSION

Anaphora resolution algorithms can usually be classified as a member of two different types of algorithms, namely, discourse-based algorithms and knowledge-poor algorithms. Discourse-based algorithms make extensive use of linguistic and domain knowledge to resolve anaphors whereas knowledge-poor algorithms deliberately limit their use of linguistic and domain knowledge. Due to the need for robust and efficient tools for anaphora resolution, knowledge-poor anaphora resolution systems have emerged in 1990s.

In this study, a knowledge-poor pronoun resolution system which attempts to resolve third person personal and reflexive pronouns in Turkish is presented. The steps in a knowledge-poor anaphora resolution system are identification of anaphors, location of the candidates for the antecedents and resolution of the anaphor using constraints and preferences. In this system, the anaphors considered are third person personal and reflexive pronouns that refer to proper person names, therefore, in order to carry out the first step of the algorithm, the input text is preprocessed manually to mark the overt and zero pronouns that refer to proper person names.

In order to determine the search scope to locate the candidates for the antecedents and constraints and preferences for Turkish, empirical studies are carried out. Constraints are used to discard the candidates that cannot be the antecedents of the pronouns. Preferences are used to sort the remaining candidates after the application of the constraints. After the empirical analysis on a sample Turkish text, the search scope of the candidate antecedents is determined as the current and three preceding sentences. The proper person names in the current and three preceding sentences constitute the candidates for the antecedents of the pronouns. The constraints determined for Turkish are number agreement, reflexive pronoun constraint, personal pronoun constraint, and selectional restrictions. The preferences determined for Turkish are quoted/unquoted text preference, recency preference, subject preference, first noun phrase preference, nominal predicate preference, repetition preference, punctuation preference, antecedent of zero pronoun preference, and syntactic parallelism preference. Subject and syntactic parallelism preferences are not implemented in the system since they require considerable linguistic knowledge; however, a special case of subject preference, nominative case preference, is implemented in the system. Among all the preferences, quoted/unquoted text, punctuation and antecedent of zero pronoun preferences are used in a knowledge-poor pronoun system for the first time with our system for Turkish. The remaining preferences have been extensively used in knowledge-poor pronoun resolution systems for other languages including English, Spanish and French. The constraints and preferences that are determined empirically are verified by a questionnaire of 17 questions on native Turkish speakers. Each of the preferences has corresponding scores which are used to resolve the pronouns correctly. The scores of the preferences are determined by making the system learn these scores by training an artificial neural network (ANN).

The implementation steps of the knowledge-poor pronoun resolution system for Turkish are sentence splitting, pronoun extraction, forming the list of candidate antecedents and determination of the antecedents. After the pronouns are extracted and the list of candidate antecedents is formed, the antecedent is determined by the application of constraints and preferences for Turkish. The candidates satisfying certain preferences are assigned the corresponding scores of these preferences. Once the constraints and preferences are applied, the candidate with the highest aggregate score is selected as the antecedent.

After the implementation, the system is tested on two different test samples. The first sample is taken from the Metu Turkish Corpus and the success rate of the system on this sample is 85.2%. The failures of the system are due to the personal pronoun constraint, non-existence of the antecedent in the current and three preceding sentences, extraction of ordinary names as proper names, and semantic reasons. In the second experiment, a Turkish child narrative is used. The success rate of the system on this second sample text is 73.6%. The success rate of the system on the second sample is lower than its success rate on the first sample since there are more failures in the second experiment due to the non-existence of the antecedent in the current and three preceding sentences. The algorithm is compared with two different baseline algorithms. The first baseline algorithm applies only the constraints and therefore resolves a pronoun if it has single candidate after the application of constraints. The second baseline algorithm applies the constraints and selects the most recent candidate among the remaining candidates. On both of the sample texts, the system performed considerably better than the baseline algorithms.

Further studies to extend the domain of the system and to improve the performance of the system are also presented. These studies include extending the system to resolve pronouns with NP antecedents, to automatically detect overt and zero pronouns, and improving the performance of the system by implementing subject and syntactic parallelism preferences and improving the personal pronoun constraint which is implemented in the system. These further studies can be accomplished by making use of NLP tools like NP extractors, POS taggers and parsers. Due to the unavailability of these tools, these extensions and improvements cannot be made at the moment.

REFERENCES

Baldwin, B. (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, Madrid, Spain.

Brennan, S. E., Friedman, M. W. And Pollard, C. J. (1987) "*A Centering Approach to Pronouns*" In Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics (ACL'87), pages 155-162, Stanford, CA.

Chomsky, N. (1981). *"Lectures on Government and Binding"*. Foris, Dordrecht.

Chomsky, N. (1995). *"Knowledge of Language: Its Nature, Origin and Use"*. MIT Press, Cambridge, MA.

Enç, M. (1986). "*Topic Switching and Pronominal Subjects in Turkish*", D. Slobin and K. Zimmer eds. Studies in Turkish Linguistics. Amsterdam: John Benjamins.

Erguvanlı-Taylan, E. (1986) "*Pronominal versus Zero Representation of Anaphora in Turkish*", D. Slobin and K. Zimmer eds. Studies in Turkish Linguistics. Amsterdam: John Benjamins.

Grosz, B., Joshi, A. K. and Weinstein, S. (1983) "*Providing a Unified Account of Definite Noun Phrases in Discourse*", Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics (ACL'83), 44-50. Cambridge, Massachusetts.

Grosz, B., Joshi, A. K. and Weinstein, S. (1995) "*Centering: A Framework for Modelling the Local Coherence of Discourse*", Computational Linguistics, 21(2), 44-55.

Grosz, B. and Sidner, C. (1986) "*Attention, Intentions, and the Structure of Discourse*", Computational Linguistics.

Hobbs, J. R. (1978) "*Resolving Pronoun References*". Lingua, 44, 339-352

Ilgaz, R. (2003a) "*Bacaksız Kamyon Sürücüsü*". Çınar Yayınları.

Ilgaz, R. (2003b) "*Bacaksız Tatil Köyünde*". Çınar Yayınları.

Keller, F. (2000). "*Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*". Ph.D. Thesis, University of Edinburgh.

Kennedy, C. and Boguraev, B. (1996) "*Anaphora for Everyone: Pronominal Anaphora Resolution without a Parser*" In Proceedings of the 16th International Conference on Computational Linguistics (COLING'96), 113-118. Copenhagen, Denmark.

Lappin, S. and Leass, H. (1994) "*An Algorithm for Pronominal Anaphora Resolution*", Computational Linguistics, 20(4), 535-561.

Lewis, G. L. (2000) *"Turkish Grammar"* Oxford University Press., Oxford, UK.

Mitkov, R. (1998) "*Robust Pronoun Resolution with Limited Knowledge*", In Proceedings of the 18th International Conference on Computational Linguistics (COLING'98)/ACL'98 Conference. Montreal, Canada.

Mitkov, R. (2002) "*Anaphora Resolution*" Longman.

Mitkov, R., Evans, R. and Orasan, C. (2002). "*A new, fully automated version of Mitkov's knowledge-poor pronoun resolution method*". In Al. Gelbukh (Ed.), Computational Linguistics and Text Processing, 168-186. Springer.

Oflazer, K., Say, B., Hakkani-Tür D. Z. and Tür, G. (2003). "*Building a Turkish Treebank*". In Anne Abeille, editor. Building and Exploiting Syntactically-annotated Corpora. Kluuwer Academic Publishers.

Orasan, C. and Evans, R. (2000). "Experiments in Optimising the Task of Anaphora Resolution". Proceedings of ICEIS 2000, 191-195. Stafford, UK.

Palomar, M., Ferrandez, A., Moreno, L., Martinez-Barco, P., Peral, J., Saiz-Noeda, M. and Muñoz, R. (2001). "*An Algorithm for Anaphora Resolution in Spanish Texts*", Computational Linguistics, 27(4), 545-567.

Rich, E. and LuperFoy, S. (1988). "*An Architecture for Anaphora Resolution*" Proceedings of the Second Conference on Applied Natural Language Processing (ANLP-2), 18-24. Texas, USA.

Say, B. Zeyrek, D., Oflazer, K. and Özge U. (2002) "*Development of a Corpus and a Treebank for Present-day Written Turkish*", in Proceedings of the Eleventh International Conference of Turkish Linguistics.

Schütze, C. T. (1996). "*The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*". Chicago: University of Chicago Press.

Siegel, S. and Castellan, N., J. (1988). "*Nonparametric Statistics for Behavioral Sciences*", McGraw-Hill.

Strube, M. (1998). "*Never Look Back: An Alternative to Centering*". In Proceedings of the 17th Int. Conference on Computational Linguistics and 36th Annual Meeting of ACL, Page 1251-1257.

Tanev H. and Mitkov R. (2002) "*Shallow Language Processing Architecture for Bulgarian*", In Proceedings of the 19th International Conference on Computational Linguistics (COLING'02).

Tetreault, J. R. (1999). "*Analysis of Syntax-based Pronoun Resolution Methods*". In Proceedings, 27th Annual Meeting of the Association for Computational Linguistics, pages 602–605.

Tın, E. and Akman, V. (1994) "*Situated Processing of Pronominal Anaphora*", In Trost, Harald, Eds. Proceedings KONVENS '94 - 2. Konferenz "Verarbeitung natuerlicher Sprache", pp. 369-378, Vienna, Austria.

Trouilleux, F. (2002) "*A Rule-Based Pronoun Resolution System for French*", In Proceedings of the 4th Discourse Anaphora and Anaphora Resolution Colloquium (DAARC'02).

Turan, Ü. D. (1995) "*Null vs. Overt Subjects in Turkish Discourse: A Centering Analysis*", Ph.D. Thesis, University of Pennsylvania.

Turhan-Yöndem, M. and Şehitoğlu, O. (1997) "*Resolution of Dropped Pronouns*", In the Proceedings of the ESSLLI'97 Student Session.

Yüksel, Ö. and Bozşahin, C. (2002) "Contextually Appropriate Reference Generation". Natural Language Engineering 8 (1): 69-89. Cambridge University Press.

APPENDIX A

QUESTIONNAIRE

1. Ali yolda Mehmet'i gördü. Ona selam verdi.
   I. Bu iki cümle gramer olarak doğru mudur?
      a. Evet
      b. Hayır
   II. İlk soruya cevabınız 'Evet' ise ikinci cümlede selam **veren** kimdir?
      a. Ali
      b. Mehmet
      c. Anlaşılır değil
   III. İlk soruya cevabınız 'Evet' ise ikinci cümlede selam **verilen** kimdir?
      a. Ali
      b. Mehmet
      c. Anlaşılır değil

2. Mehmet'i yolda gördü Ali. Ona selam verdi.
   I. Bu iki cümle gramer olarak doğru mudur?
      a. Evet
      b. Hayır
   II. İlk soruya cevabınız 'Evet' ise ikinci cümlede selam **veren** kimdir?
      a. Ali
      b. Mehmet
      c. Anlaşılır değil
   III. İlk soruya cevabınız 'Evet' ise ikinci cümleden selam **verilen** kimdir?
      a. Ali
      b. Mehmet
      c. Anlaşılır değil

3. Evden çıkınca Zeynep, Ayşe'yi gördü. "Merhaba" dedi.
   I. Bu iki cümle gramer olarak doğru mudur?
      a. Evet
      b. Hayır
   II. İlk soruya cevabınız 'Evet' ise ikinci cümlede "Merhaba" diyen kimdir?
      a. Zeynep
      b. Ayşe
      c. Anlaşılır değil

4. Bahri eve yürüdü. Kapıda durdu. Kapıyı çaldı.
   I. Bu üç cümle gramer olarak doğru mudur?
      a. Evet
      b. Hayır
   II. İlk soruya cevabınız 'Evet' ise ikinci cümlede kapıda duran kimdir?
      a. Bahri

b. Anlaşılır değil

III. İlk soruya cevabınız 'Evet' ise üçüncü cümlede kapıyı çalan kimdir?

a. Bahri

b. Anlaşılır değil

5. Ayşe oyun oynuyordu. Zeynep de ona katıldı. Oyunu sevdi.

I. Bu üç cümle gramer olarak doğru mudur?

a. Evet

b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede Zeynep'in katıldığı kişi kimdir?

a. Ayşe

b. Zeynep

c. Anlaşılır değil

III. İlk soruya cevabınız 'Evet' ise üçüncü cümlede oyunu seven kimdir?

a. Ayşe

b. Zeynep

c. Anlaşılır değil

6. Ayşe oyun oynuyordu. Zeynep geldi. Oyunu sevdi.

I. Bu üç cümle gramer olarak doğru mudur?

a. Evet

b. Hayır

II. İlk soruya cevabınız 'Evet' ise üçüncü cümlede oyunu seven kimdir?

a. Ayşe

b. Zeynep

c. Anlaşılır değil

7. Murat'la selamlaşan çocuk Ali'ydi. Sinirli görünüyordu.

I. Bu iki cümle gramer olarak doğru mudur?

a. Evet

b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede sinirli görünen kimdir?

a. Murat

b. Ali

c. Anlaşılır değil

8. "Dün Zerrin'i gördüm" dedi Ayşe. "Ne yapıyordu?" diye sordu Ahmet.

I. Bu iki cümle gramer olarak doğru mudur?

a. Evet

b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede ne yaptığı **sorulan** kimdir?

a. Zerrin

b. Ayşe

c. Ahmet

d. Anlaşılır değil

9. Belma oturuyordu. Aylin oyun oynuyordu. Sibel onu izliyordu.

I. Bu iki cümle gramer olarak doğru mudur?

a. Evet

b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede Sibel'in izlediği kimdir?

a. Belma

b. Aylin

c. Sibel

d. Anlaşılır değil

10. Murat eve koştu. Zile bastı. Muhsin onu kapıda beklemiyordu. Hayal kırıklığına uğramıştı.

I. Bu dört cümle gramer olarak doğru mudur?
  - **a.** Evet
  - **b.** Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede zile basan kimdir?
  - **a.** Murat
  - **b.** Muhsin
  - **c.** Anlaşılır değil

III. İlk soruya cevabınız 'Evet' ise üçüncü cümlede Muhsin'in kapıda beklemediği kimdir?
  - **a.** Murat
  - **b.** Muhsin
  - **c.** Anlaşılır değil

IV. İlk soruya cevabınız 'Evet' ise dördüncü cümlede hayal kırıklığına uğrayan kimdir?
  - **a.** Murat
  - **b.** Muhsin
  - **c.** Anlaşılır değil.

11. Serap Öğretmen Murat'ın kendine güvendiğini biliyordu.

I. Bu cümle gramer olarak doğru mudur?
  - **a.** Evet
  - **b.** Hayır

II. İlk soruya cevabınız 'Evet' ise bu cümlede Murat'ın güvendiği kimdir?
  - **a.** Serap Öğretmen
  - **b.** Murat
  - **c.** Anlaşılır değil

12. Serap Öğretmen Murat'ın kendisine güvendiğini biliyordu.

I. Bu cümle gramer olarak doğru mudur?
  - **a.** Evet
  - **b.** Hayır

II. İlk soruya cevabınız 'Evet' ise bu cümlede Murat'ın güvendiği kimdir?
  - **a.** Serap Öğretmen
  - **b.** Murat
  - **c.** Anlaşılır değil

13. Ali Serap Öğretmen'e Murat'ın kendine söylediklerini anlattı.

I. Bu cümle gramer olarak doğru mudur?
  - **a.** Evet
  - **b.** Hayır

II. İlk soruya cevabınız 'Evet' ise bu cümlede Murat'ın birşeyler söylediği kişi kimdir?
  - **a.** Ali
  - **b.** Serap Öğretmen
  - **c.** Murat
  - **d.** Anlaşılır değil.

14. Muhsin Aylin'i okulda görmüş. Murat onu sinemada görmüş.

I. Bu iki cümle gramer olarak doğru mudur?
  - **a.** Evet
  - **b.** Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede Murat sinemada kimi görmüştür?
  - **a.** Muhsin
  - **b.** Aylin

c. Anlaşılır değil

15. Murat, Ali'yi Serap Öğretmen'e şikayet etmişti. Onu yolda gördü.

I. Bu iki cümle gramer olarak doğru mudur?
   a. Evet
   b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede **gören** kimdir?
   a. Murat
   b. Ali
   c. Serap Öğretmen
   d. Anlaşılır değil

III. İlk soruya cevabınız 'Evet' ise ikinci cümlede **görülen** kimdir?
   a. Murat
   b. Ali
   c. Serap Öğretmen
   d. Anlaşılır değil

16. "Dün Ayşe'yi gördüm" dedi Ayla. Oya ona inanmamıştı.

I. Bu iki cümle gramer olarak doğru mudur?
   a. Evet
   b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede Oya'nın inanmadığı kimdir?
   a. Ayşe
   b. Ayla
   c. Oya
   d. Anlaşılır değil

17. Ayşe Oya'yla konuşmayı seviyordu. Sabah erkenden kalktı. Koşarak evden çıktı.

I. Bu üç cümle gramer olarak doğru mudur?
   a. Evet
   b. Hayır

II. İlk soruya cevabınız 'Evet' ise ikinci cümlede sabah erkenden kalkan kimdir?
   a. Ayşe
   b. Oya
   c. Anlaşılır değil

III. İlk soruya cevabınız 'Evet' ise üçüncü cümlede koşarak evden çıkan kimdir?
   a. Ayşe
   b. Oya
   c. Anlaşılır değil

APPENDIX B

THE CODE OF THE PROGRAM

Source code can be obtained from the site:

http://www.ceng.metu.edu.tr/~120329/adilcozumlemesi.zip