Anaphora in Turkish

Nazik Dinçtopal-Deniz

The Graduate Center, The City University of New York

Abstract

Turkish reflexive *kendisi* (self), seems to deviate from the Principles of Binding Theory as proposed within the Government and Binding framework and from the Minimalist movement approaches to referential dependencies among noun phrases. In the traditional approach, binding of *kendisi* is predicted to be governed by Binding Principle A since it is considered to be a reflexive. However, data show that binding of *kendisi* conforms to both Principle A and B.

This paper suggests that Turkish reflexive *kendi-si* (self-3SGPoss) marked with the third person singular possessive, is preceded by a null possessor, which is influential in its binding. However, binding of *kendisi* differ when it was preceded by a null and overt pronominal as possessors. It is argued that this difference is due to the fact that overt and null pronominals differ in their distribution in possessive phrases and in subject positions generally. They show the same distribution only in (in)direct object positions. If they were the null and overt counterparts of the same type of pronominal as the traditional grammars suggested their distribution would not differ.

This paper follows the arguments proposed by Cardinaletti and Starke (1999) and Safir (2004) and proposes that the null pronoun *pro* and the overt pronoun are not the same type of pronominal. *Pro* is a weak pronoun which only co-refers with an antecedent mentioned in the context. Its overt counterpart, *o* (she/he/it), on the other hand, is a strong pronoun with capacity to make independent reference. As a result, the overt pronoun functions as a topic shifter (Cardinaletti & Starke, 1999; Safir, 2004). The differences in their distribution are discussed in relation to their being weak and strong pronouns.
1. Introduction

1.1. Binding Theory in Principles and Parameters:

Binding Theory (BT) developed by Chomsky (1980, 1981, 1982) within the ‘Principles and Parameters’ (PP) Theory of syntax, is a module of the grammar that regulates referential dependencies between noun phrases (NPs) such as follows:

(1) John\textsubscript{i} thinks that Bill\textsubscript{j} hit himself\textsubscript{ij}.
(2) John\textsubscript{i} thinks that Bill\textsubscript{j} saw him\textsubscript{ij}.
(3) He\textsubscript{i} thinks that John\textsubscript{i} is a student.

In sentence (1), the reflexive ‘himself’ is co-referential with the subject of the embedded clause ‘Bill’ but not with the subject of the matrix clause, ‘John’. In sentence (2), on the other hand, the pronoun ‘him’ cannot be co-referential with the subject of the embedded clause ‘Bill’ but it can co-refer with the subject of the matrix clause, ‘John’. And finally in sentence (3), the proper name ‘John’ does not co-refer with another NP within the complex sentence. Such distributional patterns resulted in classification of NPs according to the values of features [+anaphoric] and [-pronominal]:

(i) Anaphors [+anaphoric, -pronominal]: himself, herself, each other, one another, etc.
(ii) Pronominals [-anaphoric, +pronominal]: he, she, him, her, etc.
(iii) R-expressions [-anaphoric, -pronominal]: Bill, John, etc.

The distribution of NPs with the features above was hypothesized to be governed by the three universal Binding Principles (Chomsky, 1981):

(i) Principle A: an anaphor must be bound in its governing category
(ii) Principle B: a pronoun must be free in its governing category
(iii) Principle C: an R-expression must be free
Where governing category is:

(i) \( \alpha \) is the governing category for \( \beta \) if and only if \( \alpha \) is the minimal category containing \( \beta \) and a governor of \( \beta \), where \( \alpha = \text{NP or S} \).

(Chomsky, 1981: 188)

And binding is:

(i) \( \alpha \) is X-bound by \( \beta \) if and only if \( \alpha \) and \( \beta \) are coindexed, \( \beta \) c-commands \( \alpha \), and \( \beta \) is in an X-position.

(Chomsky, 1981: 184)

The ‘governing category’ in the definition of binding is later replaced with ‘local domain’ (e.g., Chomsky, 1986), which has also undergone several revisions. In the initial formulation of the local domain, it is defined as the minimal clause (IP) which contains the bindee. According to this definition, the local domain for the sentences above was the embedded that-clause and the binding properties of sentences such as the ones above could successfully be handled under this definition. The anaphor himself is bound within its local domain, i.e., that-clause in (1); the pronoun him is free in its local domain and bound by an antecedent outside that-clause in (2); and finally the R-expression John is free in (3). However, there were problematic sentences for this definition of the local domain as well as for some other later definitions such as follows:

(4) John read Bill’s description of himself *j/*j him *j/*j.

If the local domain was the minimal clause, then in sentence (4), the whole sentence would be the local domain since it is the only minimal clause. Yet, as the indices on the anaphor himself and the pronoun him show, the matrix clause does not function as the local domain. Such examples have led to the introduction of the term ‘governor’ and ‘subject’ into the definition of the local domain. Local domain is defined as the minimal maximal projection containing the
bindee, the governor of the bindee, and a SUBJECT accessible to the bindee. Accessible means that co-indexation of SUBJECT and bindee should not violate any principles of grammar. An accessible SUBJECT could be either a lexical subject or the agreement (AGR) (Chomsky, 1986). However, this definition of local domain could not handle *i-within-i filter* cases such as the following:

(5) John$_i$ thinks that [pictures of himself$_i$ are hideous].

(6) John$_i$ thinks that [pictures of him$_i$ are hideous].

If the local domain includes an accessible subject which can be AGR or a lexical subject, then for the sentences above the embedded clauses would be the local domains. However, this would wrongly predict that the anaphor *himself* in (5) would not be able to co-index with an antecedent outside its local domain, i.e., John. Yet, it does. Furthermore, the pronoun *him* in (6) would be able to pick up *John* as its antecedent which is outside its local domain, which is not the case.

After several other attempts, local domain is defined as the minimal Complete Functional Complex (CFC) in the sense that all grammatical functions compatible with its head are realized in it. Thus, the local domain is the minimal CFC which contains the bindee and a governor of the bindee in which the bindee’s Binding Principle could, in principle, be satisfied. (Chomsky, 1995).

1.2. Problems with the Binding Theory

Although the final definition of the local domain was less restrictive than the earlier versions, pronouns and anaphors in several other languages such as Turkish, *kendisi*, Japanese, *zibun* (Enç, 1989) Dutch, *zich*, Norwegian *seg*, and Italian *sè* (Reinhart & Reuland, 1993) and Icelandic *sig* (Wexler & Manzini, 1987) still posed a challenge to the Binding Principles. These
observations have led to different versions of Binding such as parametric variation of governing category across languages and lexical parameterization within a specific language (Wexler & Manzini, 1987), the ‘Reflexivity’ approach (Reinhart & Reuland, 1993) or revision of features for NPs (Enç, 1989) or a movement analysis for anaphors as in the Minimalist Program (Hornstein, 2006).

1.3. Aim/Overview of the paper

This paper examines the distribution of so-called reflexive kendisi (self) in Turkish with reference to the traditional BT as proposed within the Principles and Parameters framework. Turkish reflexive kendisi has been reported to deviate from BT-A since it could co-refer with an antecedent within and outside its local domain. Section 2 sets the problem and reviews the literature on previous investigations of kendisi. Section 3 investigates the structure of genitive possessive phrases in Turkish and proposes that kendisi marked with the third person singular possessive is preceded by a null pronominal which is influential in the binding of the whole phrase. However, the fact that kendisi does not show the same distribution when possessed by an overt and null possessor will lead to Section 4 which focuses on the distribution of null and overt pronouns in Turkish. Section 4 will provide literature on syntactic as well as pragmatic accounts for the differing distributions of null and overt pronouns. And finally Section 5 will conclude the paper by following upon syntactic accounts claiming that null and overt pronouns do not belong to the same NP type.
2. Turkish Reflexive, *kendisi*

2.1. Problems with the binding of *kendisi*

Turkish NP, *kendisi*, is a so-called reflexive that has received a lot of attention since its distribution deviated from the traditional BT-A. Literature on Turkish reflexives reports that the bare form of Turkish reflexive pronoun is *kendi* (e.g., Göksel & Kerslake, 2005, Gürel, 2002). *Kendi* can be translated as ‘self’ or ‘essence’ into English. As a result of nominal agreement in Turkish, *kendi* receives a possessive suffix to agree with the person and the number of its possessor. The third person singular form of the reflexive is reported to appear in two forms: *kendi* and *kendi-si* since it can optionally receive the third person singular agreement marking –sI (Göksel & Kerslake, 2005). Analyzed as anaphors so far, both *kendi* and *kendisi* have been predicted to co-refer with antecedents within their local domains due to their [+anaphoric] feature. However, it is evident that in Turkish, it is possible for *kendisi* to have an antecedent both within and outside its local domain if local domain is taken to be the minimal clause (Enç, 1989; Gürel, 2002, 2004; Kornfilt, 1997):

Ali Ayşe-Gen self-3SgPoss-Acc like-Nom-3SgPoss-Acc think-Prog

‘Ali thinks that Ayşe likes himself/herself.’

Ali Ayşe-Gen self-3SgPoss-Acc like-Nom-3SgPoss-Acc think-Prog

‘Ali thinks that Ayşe likes himself/herself.’

Sentence (7) is compatible with BT-A, when the local domain is taken to be the embedded sentence in which the anaphor’s BT principle could *in principle* be satisfied, i.e., it can be locally bound. As a result, the anaphor *kendi* is bound by the antecedent ‘Ayşe’ within its local domain.
and it cannot co-refer with the matrix clause subject, Ali. Sentence (8), on the other hand, is problematic if we take the same CFC (i.e., the embedded clause) as the local domain. As the indices on kendisi indicate, kendisi could co-refer with the subject of the embedded clause, ‘Ayşe’ or the matrix subject ‘Ali’ or it can pick up another antecedent outside the sentence within the discourse. If the embedded clause is taken as the local domain in Turkish as it was in (7), k and m readings are predicted to be ungrammatical. If the embedded clause is not a good candidate for local domain in the sense that the BT principle of the anaphor could not be met in principle and the matrix clause is analyzed as the local domain, then at least the m reading is predicted to be ungrammatical. Furthermore, the matrix clause as the local domain will not meet the BT principles of the anaphor in (7). Since both of the NPs, namely kendi and kendisi, are assumed to share the same nominal property, namely [+anaphoric], a local domain selected for one is predicted to satisfy the other’s BT principle, as well.

In addition to their different distributions, kendi and kendisi cannot always be interchangeably used:

(9) a. Cem Bey’le konuşabil-im mi-yim?
    Cem Mr.-with talk-Psb-Aor Cl-1Sg
    ‘May I talk to Mr. Cem?’

    b. Kendisi/*kendi henüz gel-me-di.
    Self-3SgPoss yet arrive-Neg-Past
    ‘Himself has not arrived yet.’

Use of a bare form of reflexive would result in ungrammaticality in (9) just like its English counterpart. In this context, since the bare form of the reflexive, kendi, is not bound by an antecedent within its local domain, which would be the matrix clause, the result is an ungrammatical sentence. However, the inflected from, kendi-si, behaves like a pronominal rather
than an anaphor and could co-refer with an antecedent mentioned within the context outside its local domain. However, considering the sentences (7) and (8), it is difficult to call the reflexive a pronoun since it has the capacity to co-refer with an antecedent within its local domain when there exists one.

Thus, *kendisi* has been analyzed as a problematic case for BT-A in the literature, so far.

### 2.2. Earlier explanations to the binding of *kendisi*

Enç (1989) is one of the researchers who first investigated this problem in Turkish. According to Enç (1989) the feature system involved in the BT needed to be revised. She claimed that the feature system in the BT is not sufficient to capture the distribution of NPs observed in natural languages and it does not entail semantic binding. She proposed that some other features such as ‘binder’ and ‘licenser’ are necessary to account for the distribution of NPs in natural languages.

The binder feature is [+B] for NPs which need a sentence-internal antecedent that semantically binds it and [-B] for the ones which do not need a sentence-internal semantic binder. This feature is proposed to account for possible semantic as well as syntactic antecedents of a reflexive. The licenser feature is [+L] for the NPs that require a licenser and NPs with feature [+L] are subject to the BT-A, in other words, they are anaphors. These two features are claimed to be helpful to account for expressions such as Dogrib *ye* which share properties with both reflexives and pronominals. Such expressions would have [+L, -B] whereas a true reflexive would be [+L, +B]. However, these two features are not sufficient to account for the distribution of all pronominals such as Turkish *kendisi*. *Kendisi* does not need a sentence-internal antecedent. Enç mentions that it does not need a binder, at all. And if there is any possible binder in the context, there is no locality or non-locality condition for that. Thus, it would have the features [-B, -L, -pronominal].
However, it should be noted that *kendisi* is at least semantically dependent on an antecedent for its interpretation. In other words, it does not behave like a referential expression and pick up a unique referent. If a [+B] feature is assigned when there is a potential semantic binder within the context (not necessarily within the sentence), then Turkish *kendisi* must be [+B] because it needs to be bound semantically in the context regardless of having a local, non-local or disjoint reading. In that sense, it might be suggested that Turkish *kendisi* is [+B, -L, -pronominal].

However, Enç assigns [+B] only to sentence-internal binders. Therefore, this proposition does not make a clear distinction between a referential expression which does not need any semantic binder within the context and an NP such as *kendisi* which is dependent on another NP within the context to be referential.

Later on in her paper, Enç proposes that the additional two features (i.e., [B] and [L]) do not cover the case of reflexives where the binder and the licenser are the same. Thus, she adds a further [+ID] feature. A pronominal is [+ID] if the pronoun’s licenser and binder are co-indexed. Taking Turkish *kendisi* into consideration, then, we would ideally assign the features [+B, -L, +ID, -pronominal] since it is possible to co-index the binder and the licenser as in sentences (7) and (8) in Turkish. Furthermore, since it is possible to have the licenser within the sentence in Turkish it is even possible to say that it is [+L], although she assigns [-B, -L] features to it. As a result the local binding of *kendisi* cannot be accounted for. Therefore, Enç’s proposal does not seem to explain all the possible readings of Turkish *kendisi*, either.

Gürel (2002; 2004) investigated second language acquisition and first language attrition of overt and null pronouns’ distribution. In her analysis, the null pronominal, *pro*, shows a similar distribution to the reflexive, *kendisi*, rather than to its overt counterpart, *o* (he, she, it). This observation has led to the conclusion that *pro* is indeed the null counterpart of *kendisi* rather than
the overt pronoun, *o* since both *pro* and *kendisi* show anaphoric and pronominal distribution and are unconstrained in their binding. Following sentences compare the distributions of *kendisi*, overt pronoun *o*, and *pro* in possessor position of a genitive-possessive phrase (10), in subject position of a non-finite embedded sentence (11) and a finite embedded sentence (12):

(10) \(\text{Ali}_i [\text{o-nun}_i^*/k / \text{pro}_i/k / \text{kendi}-\text{si-nin}_i^*/k \text{kari-si-ni}]_m \text{ara-di}.\)
    \(\text{Ali she/he-Gen/} / \text{self-3SgPoss-Gen wife-3SgPoss-Acc} \text{ call-Past}\)
    ‘Ali called her/his/\(i//k\) himself/herself’s/\(i//k\) wife’

(11) \(\text{Ali}_i [\text{o-nun}_i^*/k/m/\text{pro}_i/k/m \text{kendi-si-nin}_i^*/k/m \text{yalancı ol-duş-u}]-nu \text{düşün-uyor}.\)
    \(\text{Ali she/he-Gen/} / \text{self-3SgPoss-Gen liar be-Nom-3SgPoss-Acc think-Prog}\)
    ‘Alii thinks that she/he/\(i//k/m\) pro\(i//k/m\) himself/herself\(i//k/m\) is a liar’

(12) \(\text{Ali}_i [\text{o}^*/k/m/\text{pro}_i/k/m \text{kendi-si}_i^*/k/m \text{kazan-di}]_m \text{san-iyor}.\)
    \(\text{Ali she/he/} / \text{self-3SgPoss win-Past think-Prog}\)
    ‘Ali thinks that she/he/\(i//k/m\) pro\(i//k/m\) himself/herself\(i//k/m\) won’

As the examples above show, in possessor positions of genitive possessive phrases, in subject position of non-finite embedded sentences and in finite embedded sentences, *pro* and *kendisi* show the same distribution. However, if *pro* and *kendisi* were the null and overt realizations of the same NP type, we would expect to see *pro* and *kendisi* to be interchangeably used without causing any syntactic ungrammaticality. However, use of a *pro* but not *kendisi* makes the following sentence ungrammatical.

    \(\text{Ali Kerem-Gen self-3SgPoss-Acc like-prog}\)
    ‘Ali likes Kerem’s self’.


Thus, the conclusion that *kendisi* is the overt counterpart of *pro* does not seem to be plausible. Furthermore, it does not explain the question why *kendisi* and *pro* are unconstrained in their binding, either.
As mentioned before, it has been reported that there are two third person singular forms for the reflexive, namely a bare form *kendi* and a morphologically inflected form *kendi-si*. As the examples (7), (8), and (9) showed, the bare form of the reflexive is compatible with BT-A. However, the form receiving overt morphology for third person singular deviates from BT-A. Thus, it is necessary to look at the nominal agreement morphology for the reflexive to see whether its distributional problem could be related to its overt morphology. The following section will examine deviation of *kendisi* from BT-A in light of the structure of genitive-possessive phrases in Turkish.

3. A Null Possessor, pro, Preceding *kendisi*

3.1. Structure of genitive-possessive phrases in Turkish

Turkish possessive phrases occur in the form of two NPs marked with genitive and possessive agreement as follows:

(NP+genitive) + (NP+possessive)

The first NP carrying the genitive case marker indicates the possessor and the second NP with the possessive marker indicates the possessed. It is possible to omit the possessor in genitive possessive constructions since the agreement marking on the possessed already indicates the person and number of the possessor:

(14) (Ben-im) ev-im
     I-Gen house-1SgPoss
     ‘My house’

(15) (Ali-nin) araba-si
     Ali-Gen car-3SgPoss
     ‘Ali’s car’

The following is the agreement paradigm for genitive-possessive phrases in Turkish:
Table 1. Turkish nominal agreement for genitive-possessive phrases.

<table>
<thead>
<tr>
<th>Person</th>
<th>Genitive</th>
<th>Possessive</th>
<th>Examples with ‘ev’ (house)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Sg</td>
<td>-(I)m</td>
<td>-(I)m</td>
<td>(ben-im) ev-im ‘my house’</td>
</tr>
<tr>
<td>2Sg</td>
<td>-(I)n</td>
<td>-(I)n</td>
<td>(sen-in) ev-in ‘your house’</td>
</tr>
<tr>
<td>3Sg</td>
<td>nIn</td>
<td>-(s)I</td>
<td>(o-nun) ev-i ‘her/his house’</td>
</tr>
<tr>
<td>1Pl</td>
<td>-(I)m</td>
<td>-(I)mlz</td>
<td>(biz-im) ev-imiz ‘our house’</td>
</tr>
<tr>
<td>2Pl</td>
<td>-(I)n</td>
<td>-(I)nIz</td>
<td>(siz-in) ev-iniz ‘your (pl) house’</td>
</tr>
<tr>
<td>3Pl</td>
<td>-(I)n</td>
<td>-(I)ArI</td>
<td>(onlar-in) ev-(ler)i ‘their house(s)’</td>
</tr>
</tbody>
</table>

When we look at the third person singular form of the agreement paradigm, we observe that the third person singular marking is morphologically realized with the suffix ‘(s)I’ with a deletable ‘s’. As Göksel and Kerslake (2005) mention, in Turkish, under certain circumstances, attachment of a suffix to a root or a stem results in vowel or consonant deletion, or epentheses since vowels do not occur next to each other. Therefore, if a suffix beginning with a vowel is attached to a root ending in a vowel, this results in either deletion of suffix-initial vowel or epentheses of the consonant ‘y’. First person possessive suffix is in the group of the suffixes that lose its initial vowel when attached to a stem ending in a vowel:

(16) silgi-m eraser-1SgPoss vs. ev-im house-1SgPoss
‘my eraser’ ‘my house’

Similarly, some suffixes have deletable initial consonants which normally occur to prevent vowel sequences. Examples to these suffixes are the genitive suffix, which has a deletable ‘n’ and third person possessive suffix which has a deletable ‘s’ (Göksel & Kerslake, 2005) as the examples below show:

(17) Ali’nin silgi-si eraser-3SgPoss vs. Kerem’In ev-i Kerem-Gen house-3SgPoss
‘Ali’s eraser’ ‘Kerem’s house’
As for the reflexive, the literature has reported that the third person singular form of Turkish reflexive appears in two forms: the bare form, *kendi* and an optionally marked third person singular form, *kendi-si* (e.g., Göksel & Kerslake, 2005). As a result, the following paradigm is proposed for the reflexive:

Table 2. Possessive agreement paradigm for *kendi*.

<table>
<thead>
<tr>
<th>Personal Pronoun</th>
<th>Genitive</th>
<th>Possessive</th>
<th>Kendi in Gen-Poss const.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Sg</td>
<td>-(I)m</td>
<td>-(I)m</td>
<td>(ben-im) kendi-m</td>
</tr>
<tr>
<td>2Sg</td>
<td>-(I)n</td>
<td>-(I)n</td>
<td>(sen-in) kendi-n</td>
</tr>
<tr>
<td>3Sg</td>
<td>-(s)I</td>
<td></td>
<td>(o-nun) kendi-(si)</td>
</tr>
<tr>
<td>1Pl</td>
<td>-(I)m</td>
<td>-(I)mIz</td>
<td>(biz-im) kendi-miz</td>
</tr>
<tr>
<td>2Pl</td>
<td>-(I)n</td>
<td>-(I)nIz</td>
<td>(siz-in) kendi-niz</td>
</tr>
<tr>
<td>3Pl</td>
<td>-(lAr)I</td>
<td>-(lAr)I</td>
<td>(onlar-in) kendi-leri</td>
</tr>
</tbody>
</table>

When Table 2, which shows the nominal agreement on *kendi* as a possessed NP, is compared to Table 1, which shows a common noun, ‘ev’ (house) as a possessed NP, it is observed that normally there is an optional ‘s’ rather than an optional morpheme, *sI*, for third person singular possessive marking, where the optionality of ‘s’ comes from vowel or consonant deletion or epenthesis conditions in Turkish. However, optional possessive marking of an NP for person is not common since it will fail to show person agreement or possessive relationship without overt possessive agreement morphology. Common nouns can only appear without overt possessive agreement when their possessor is overtly mentioned:

(18) Ben-im / o-nun ev
     I-Gen / she/he-Gen house
     ‘My / her/his house’
Without its possessor, ‘ben-im’ (my), ‘ev’ would only mean ‘house’ in the example above. It would not be able to indicate any person or possession relationship without its possessive agreement morphology or without an overt possessor preceding it. Furthermore, the optional marking of possessive suffix when its possessor is overtly expressed is not restricted to the third person singular as the example above indicates. The optionality of the possessive marking becomes available with the overt mention of the possessor with all persons.

Therefore, existence of two forms of reflexive, namely kendi and kendisi cannot be explained by the current nominal agreement system for Turkish possessive phrases. If the bare form of the reflexive is kendi as the literature suggests, it would not be able to indicate any agreement or possession relationship without its possessor overtly preceding it. Therefore, the bare form of the reflexive could be different than what has been assumed so far. The next section examines this in relation to the historical factors in reflexivization.

3.2. ‘kend-’ as the bare form of the reflexive and ‘kend-i-si’ as doubly-marked possessive NP

The reason why Turkish reflexives have turned out to exist in two forms could be related to the way the reflexive form evolved in the language. Schaldt (2000) investigated the origin and evolution of reflexive markers with 150 languages. According to his data, reflexive markers which are originally nominal forms, i.e., NPs, evolve into linguistic expressions of reflexivization. Reviewing 150 languages, Schaldt (2000) has shown common sources of reflexive markers. According to his analysis of frequency, the order of lexical elements that evolve into reflexive markers in time is: (i) body part names such as ‘head’, ‘arm’; (ii) nominal sources denoting something like ‘person’, ‘self’, ‘owner’; (iii) emphatic pronouns; (iv) object
Recall that Turkish reflexive *kendi* means ‘self’ or ‘essence’ (‘the person pronoun used to describe ‘essence’’ (in the official Turkish dictionary of the Institution of Turkish Language-Türk Dil Kurumu)). Therefore, it is a good candidate for a lexical element evolving into a reflexive marker according to Schaldt’s analysis. Furthermore, it still shows the nominal agreement paradigm in the language in the form of a possessive phrase. However, the question as to why only the third person singular form receives agreement marking optionally needs further attention having all other forms (i.e., 1Sg, 2Pl etc.) receiving nominal agreement all the time.

The assumption here is that the bare form of the reflexive is not ‘kendi’ as has been reported in the literature so far. Rather, ‘kend’ or ‘kent’ (due to the word-final devoicing in Turkish) is the root of the reflexive marker and it receives the third person singular form just like all other person markings and becomes ‘kend-i’ in the third person singular. After evolving successfully into a reflexive marker in the language, ‘kend’ does not appear in the form of a common noun meaning ‘essence’ any more. And the third person singular form, ‘kendi’ gains other uses such as an adjective meaning ‘own’ with the form marked for the third person, i.e., *kendi* (Lewis, 1967) such as follows:

```
(19) (Ben-im) kendi kitab-im
     1-Gen own book-1SgPoss
     ‘My own book’

(20) (O-nun) kendi kitab-i
     She/he-Gen own book-1SgPoss
     ‘Her/his own book’
```
Thus, after gaining other uses with its third person singular marking, the reflexive might have lost its form ‘kend-i’ ‘essence-3SgPoss’. As a result, the language might have marked the NP for 3SgPoss once again for third person singular as ‘kendi-si’ with an intended meaning of ‘essence-3SgPoss’.

What follows is the possessive agreement paradigm proposed for ‘kend’ and ‘kendi’ as the stem of the reflexive marker:

**Table 3. Kend- and kendi- as the root of Turkish reflexive marker:**

<table>
<thead>
<tr>
<th>Personal Pronoun</th>
<th>Possessive</th>
<th>‘kend’ as the bare form of the reflexive</th>
<th>‘kendi’ as the bare form of the reflexive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Sg -(I)m</td>
<td>kend-im</td>
<td>kendi-m</td>
<td></td>
</tr>
<tr>
<td>2Sg -(I)n</td>
<td>kend-in</td>
<td>kendi-n</td>
<td></td>
</tr>
<tr>
<td>3Sg -(s)I</td>
<td><strong>kend-i</strong></td>
<td><strong>kendi-si</strong></td>
<td></td>
</tr>
<tr>
<td>1Pl -(I)mIz</td>
<td>kend-imiz</td>
<td>kendi-miz</td>
<td></td>
</tr>
<tr>
<td>2Pl -(I)nIz</td>
<td>kend-iniz</td>
<td>kendi-niz</td>
<td></td>
</tr>
<tr>
<td>3Pl -(lAr)I</td>
<td>*kend-leri</td>
<td>kendi-leri</td>
<td></td>
</tr>
</tbody>
</table>

Taking the bare form of the noun as ‘kend’ is more compatible with the nominal agreement system that the language has for possessive phrases. Yet, the third person plural form of the noun ‘*kend-leri’ turns out to be ungrammatical (or non-existent in the language) if ‘kend’ is the root. This might pose a challenge to the assumption that ‘kend’ is the root of the noun. However, we have seen that a bare form of an NP cannot indicate any possession relationship. This observation provides strong evidence for ‘kend’ as the bare form and ‘kendi’ as the third person singular possessive form, which is able to show person and possession relationship. Furthermore,
the distribution of third person plural, ‘kendileri’ patterns with ‘kendisi’ rather than other forms of the anaphor, as follows:

(21) [Ali ve Ayşe], [on-lar-in_k kendi-leri-ni_i/k/m beğen-diğ-in]-i düşün-uyor.

Ali and Ayşe she/he-Pl-Gen self-3SgPoss.Acc like-Nom-3SgPoss.Acc think-Prog

‘[Ali and Ayşe], think that they_k like themselves i/k/m.’

Therefore, it provides evidence that both third person singular and third person plural forms are doubly possessive-marked after realization of the first possessive marking as a full anaphor.

The reason why the language would mark possession twice is not evident nor is it within the scope of this study. However, it is known that natural languages such as Urinara may optionally employ double or multiple marking of possession (Olawsky, 2006). Olawsky (2006) mentions that double marking of possession could be an optional feature of the grammar. Whether double marking of possession could indicate differences or not could be related to the diachronic changes of the system. He speculates that some kind of emphatic function might be implied in this usage.

For double-marking of possessives in Turkish and evolution of reflexives from nominal elements ‘can-i-st’ (soul-3SgPoss-3SgPoss) provides another example where ‘can’ takes two possessive morphemes and emphasizes the meaning of ‘my soul’ which is used as an address term indicating intimacy. According to Schaldt’s analysis, ‘can’ (soul) would also be a good candidate to evolve into a reflexive marker. Although it does not show any anaphoric relationship to another NP in the context yet, it indicates that double-marking of possessed is not totally unacceptable in Turkish.

Thus, the suggestion here is that all person forms of the reflexive in Turkish, including kendi, are morphologically marked and they are true anaphors. The doubly-marked third person singular
form *kendisi* and third person plural form *kendileri* deviate from BT-A. The true anaphor for third person singular is the bi-morphemic form ‘kend-i’, which has the feature [+anaphoric]. Due to historical changes in the language, this form is also realized as a mono-morphemic form, ‘kendi’ which provides input to ‘kendi-si’, which deviates from BT-A.

The deviation of *kendisi* from BT-A will be handled with reference to the genitive possessive structures in Turkish. Just like any common noun phrase marked with possessive agreement in a genitive possessive construction, such as ‘ø ev-i’ (house-3SgPoss), it is proposed that *kendi-si* is marked with third person singular possessive agreement and is preceded by a null pronominal in the form of a genitive possessive phrase. And this null form, namely, *pro*, influences the binding of the whole phrase as will be discussed in the next section.

### 3.3. Null pronominal, *pro*, preceding *kendisi* as a possessor:

Turkish is a pro-drop language allowing its subjects, objects and even possessors in genitive possessive phrases to be null unless the overt mention of them would make an emphatic or contrastive function or convey new information (Erguvanlı-Taylan, 1986). We have observed that due to the agreement marking on the possessed NP, it is possible to drop the possessor. If that is the case, there is a null element in the possessor position of the genitive possessive phrase. Thus, it is also possible that there is a *pro* before *kendisi* just like any other possessive-marked NPs such as the following:

(22) Ali [pro ev-i-ni] beğeniyor.  
Ali house-3SgPoss-Acc like-Prog  
‘Ali likes her/his house’

(23) Alii [pro kendi-si-ni,k] beğeniyor.  
Ali self-3SgPoss-Acc like-Prog  
‘Ali likes her/his self’
There are two possible claims as for the type of *kendisi* in this context. One claim is that *kendi-si* is a common noun possessed by a null pronominal. As mentioned in the previous section the third person possessive form of reflexive, ‘kend-i’ is also realized as a mono-morphemic form ‘kendi’ which means ‘own’ after evolving into a reflexive marker. This mono-morphemic form lacks the [+anaphoric] feature and acts as a common noun, maybe with an original meaning for ‘self’ or ‘essence’. The language marks the mono-morphemic form for third person singular just like a common noun. Thus, it does not show distribution of an anaphor. Instead, it has a possession relationship with its null possessor. And the null possessor, *pro*, being a pronominal is subject to BT principles, namely, BT-B. Consequently, binding of the whole phrase would be subject to BT-B and *kendisi* would receive non-local readings.

The second claim would be that *kendisi* still carries [+anaphoric] feature and it is bound by a null possessor. The null possessor could bind the anaphor within the genitive possessive phrase since the whole phrase bears person and number agreement (i.e., AGR). As a result, the genitive possessive phrase would qualify as a local domain. Thus the AgrP, which is the genitive possessive phrase in (23) would be the local domain for the anaphor *kendisi*. Accordingly, it will pick up the antecedent *pro* in the Spec position of the AgrP, as its antecedent and it will not violate the BT-A. The same AgrP would be the local domain for the *pro*, which is a null pronominal bearing the features [+pronominal, -anaphoric]. Thus, as a result of BT-B, it will be free in its local domain, i.e., AgrP, and will pick up an antecedent outside its local domain, which is ‘Ali’ or another entity mentioned in the context. Thus, *kendisi* will pick *pro* as its antecedent within its local domain and *pro* will pick its antecedent outside its local domain. Either way, the problem encountered for the reflexive seems to be solved.
Although both approaches seem to be plausible considering the binding of *kendisi*, I will keep the position that *kendisi* behaves like a common noun and the relationship between *pro* and *kendisi* is a possession relationship. The rationale for this approach is related to the fact that the language already employs a third person singular form of reflexive, i.e., *kendi*. Thus, it becomes redundant for the language to employ another form as a reflexive. Instead, there might be a need to use the noun ‘self’ as a common noun as well as a reflexive.

Although proposing that a null possessor precedes the reflexive in the form of a genitive possessive phrase eliminates the problem for the binding of *kendisi*, this approach brings further issues with regard to the binding of *pro* and the overt pronoun in Turkish because an equivalent sentence of (23) with an overt pronoun would not result in the same interpretation:

\[(24)\text{ Ali }\text{ [o-nun}_k\text{ kendisi-ni-si}_k\text{ ] begen-iyor.}\]

She/he-Gen self-3SgPoss-Acc like-Prog

‘Ali likes her/his self’

When a similar approach is carried to sentence (24) with an overt possessor rather than a null one, the reflexive could only co-refer with the overt pronominal, *o*. It is possible to say that *kendisi* is a common noun which belongs to its possessor, *o*. However, this time, it would be expected for the overt pronoun to be able to pick Ali as its antecedent just like the null pronoun since it is outside its local domain, i.e., AgrP (genitive possessive phrase). In other words, the chain of co-reference in (23) would also be expected to happen in (24) if both the null and overt pronominal are [-anaphoric, +pronominal] and if the AgrP qualifies as a local domain. So, the overt pronoun being [-anaphoric, +pronominal] NP type would be subject to BT-B and would be expected to be free in its local domain, which is AgrP (as a result of AGR), and would be able to pick up ‘Ali’ as its antecedent. Yet, it is not the case. The overt pronoun in (24) cannot co-refer
with ‘Ali’. It can only pick up an antecedent outside the sentence in which case it would not be able to give *kendisi*, the ‘i’ reading.

What we have observed so far is that the Turkish reflexive *kendisi* does not pose a challenge to the BT. It is a doubly-marked possessed NP which is possessed by a null pronominal, *pro*. Its deviance from the BT-A is a result of distribution of *pro*, which is a [-anaphoric, +pronominal] NP. However, *kendisi* shows different distributions with a null and an overt pronoun possessor, which seems to indicate that the null and overt pronouns are not subject to the same local domain when their binding conditions are considered. Or they are not the null and overt counterparts of the same NP type. Therefore, the next section will examine the distribution of overt and null pronominals in Turkish assuming that the problematic reflexive does not pose a challenge to the BT. If overt and null pronouns differ in their distribution (as also mentioned by Gürel, 2002) irrespective of being in the possessor position of *kendisi*, this would further support the hypothesis that *kendisi* is preceded by a null pronoun.

4. Null and Overt Pronouns

4.1. Distribution of null and overt pronouns in Turkish

Traditionally, it has been assumed that pro-drop languages have two forms of realizing subjects and objects: overtly and covertly. Although the use of overt and null pronouns depend on pragmatic and contextual factors such as avoiding ambiguities by using an overt pronoun, *pro* has been classified as a null element with the features [+pronominal, -anaphoric] just like its overt counterpart. Consequently, its distribution is predicted to be subject to BT-B. To see whether the overt and null pronouns both obey the BT-B, their distributions will be compared.
What follows provides examples with the overt pronoun \textit{o} and \textit{pro} in direct object positions in matrix sentences (25) and (26) and in an embedded clause (27):

(25) a. Ali\textsubscript{i} \textit{o-nu\textsubscript{i/k}} \textit{sev-iyor.} \\
     Ali \textit{she/he-Acc} \textit{like-Prog} \textit{‘Ali\textsubscript{i} likes her/him\textsubscript{i/k}’}

b. Ali\textsubscript{i} \textit{pro\textsubscript{i/k}} \textit{sev-iyor.} \\
     Ali \textit{like-Prog} \textit{‘Ali\textsubscript{i} likes pro\textsubscript{i/k}’}

(26) a. [Ali’n\textsubscript{i}n baba\textsubscript{-si}]\textsubscript{k} \textit{o-nu\textsubscript{i/k/m}} \textit{sev-iyor.} \\
     Ali-Gen father-3SgPoss \textit{she/he-Acc like-Prog} \textit{‘[Ali\textsubscript{i}’s father\textsubscript{k} likes her/him\textsubscript{i/k/m}’}

b. [Ali’n\textsubscript{i}n baba\textsubscript{-si}]\textsubscript{k} \textit{pro\textsubscript{i/k/m}} \textit{sev-iyor.} \\
     Ali-Gen father-3SgPoss \textit{like-Prog} \textit{‘[Ali\textsubscript{i}’s father\textsubscript{k} likes pro\textsubscript{i/k/m}’}

(27) a. Ali\textsubscript{i} [Ay\textsubscript{ş}e’n\textsubscript{x}n ]\textsubscript{k} \textit{o-nu\textsubscript{i/k/m}} \textit{sev-diğ-i]-ni} \textit{söyle-di.} \\
     Ali Ay\textsubscript{ş}e-Gen \textit{she/he-Acc like-Nom-3SgPoss-Acc} \textit{say-Past} \textit{‘Ali\textsubscript{i} said (that) Ay\textsubscript{ş}e\textsubscript{k} likes her/him\textsubscript{i/k/m}’}

b. Ali\textsubscript{i} [Ay\textsubscript{ş}e’n\textsubscript{x}n ]\textsubscript{k} \textit{pro\textsubscript{i/k/m}} \textit{sev-diğ-i]-ni} \textit{söyle-di.} \\
     Ali Ay\textsubscript{ş}e-Gen \textit{like-Nom-3SgPoss-Acc} \textit{say-Past} \textit{‘Ali\textsubscript{i} said (that) Ay\textsubscript{ş}e\textsubscript{k} likes pro\textsubscript{i/k/m}’}

In all (a) and (b) sentences above, both the overt and null pronoun behave consistently with BT-B. BT-B requires a pronominal to be free in its local domain. If we take the matrix clause in (25) and (26) and the embedded clause in (27) as local domains, both the overt and null pronouns conform to BT-B and they are free in their local domains.

The following example shows the distribution of \textit{pro} and overt pronoun in indirect object position of a non-finite embedded sentence:
(28) a. Ali$_i$ [Ayşe’nin$_k$ o-na/∗i/k/*m bak-ma-si]$_{-}$_{-}$_ni$ iste-m-yi$_{or}$.
   Ali$_i$ Ayşe-Gen she/he-Dat look-Nom-3sgPoss-Acc want-Neg-Prog
   ‘Ali$_i$ does not want Ayşe$_k$ to look at her/him$_{i/∗i/k/m}$’

b. Ali$_i$ [Ayşe’nin$_k$ pro$_i/∗i/k/m$ bak-ma-si]$_{-}$_{-}$_ni$ iste-m-yi$_{or}$.
   Ali$_i$ Ayşe-Gen look-Nom-3sgPoss-Acc want-Neg-Prog
   ‘Ali$_i$ does not want Ayşe$_k$ to look at pro$_i/∗i/k/m$’

Similar to the direct object positions, the distribution of $o$ and $pro$ does not differ as an indirect object.

However, when $o$ and $pro$ are used in the structure of genitive-possessive constructions, their distribution differs as the indices in the following examples show:

(29) a. Ali$_i$ [o-nun$_{i/∗i/k/m}$ kitab-1]$_{-}$_{-}$_ni$ sev-yi$_{or}$.
   Ali$_i$ she/he-Gen book-3SgPoss like-Prog
   ‘Ali$_i$ likes her/his$_{i/∗i/k/m}$ book’

b. Ali$_i$ [pro$_{i/∗i/k/m}$ kitab-1]$_{-}$_{-}$_ni$ sev-yi$_{or}$.
   Ali$_i$ book-3SgPoss like-Prog like-Prog
   ‘Ali$_i$ likes pro$_{i/∗i/k/m}$ (her/his) book’

(30) a. [Ali’nin$_{i}$_ kaba-si]$_k$_ [o-nun$_{i/∗i/k/m}$ kitab-1]$_{-}$_{-}$_ni$ sev-yi$_{or}$.
   Ali-Gen father-3SgPoss she/he-Gen book-3SgPoss like-Prog
   ‘Ali’$_{i}$’s$_{i}$ father$_{k}$ likes her/his$_{i/∗i/k/m}$ book’

b. [Ali’nin$_{i}$_ kaba-si]$_k$_ [pro$_{i/∗i/k/m}$ kitab-1]$_{-}$_{-}$_ni$ sev-yi$_{or}$.
   Ali-Gen father-3SgPoss book-3SgPoss like-Prog
   ‘Ali’$_{i}$’s$_{i}$ father$_{k}$ likes pro$_{i/∗i/k/m}$ (her/his) book’

In the examples above, if the genitive-possessive phrase is taken to be the local domain as a result of the nominal agreement, the indices on $pro$ are predicted. In other words, $pro$ picks up an antecedent (‘Ali’ in (29) and ‘Ali’s father’ in (30)) which is outside its local domain, i.e., the genitive possessive AgrP. However, if the AgrP is the local domain, then, we would expect the overt pronoun to be able to pick up any antecedent outside it (e.g., ‘Ali’ in (29) and ‘Ali’s father’ in (30)). Yet, it does not.
Similarly, when \( o \) and \( pro \) are used as subjects of embedded sentences, they show different distributions. This similarity might be related to the structure of embedded sentences in Turkish. In Turkish, most of the embedded sentences usually get nominalized and have the structure of genitive-possessive constructions. The subject of the embedded sentence receives genitive marker and the embedded verb carries the possessive suffix (Erguvanlı-Taylan, 1986). Thus, the distributions of \( pro \) and \( o \) differ just in the same way it was in simple genitive-possessive constructions:

\[
(31) \begin{align*}
\text{a. } & \text{ Ali } \quad [\text{o-nun}^{\text{ı/k}} \text{ akıllı} \quad \text{ol-duğ-u]-nu} \quad \text{söyle-di.} \\
& \text{Ali she/he-Gen intelligent be-Nom-3SgPoss-Acc say-Past} \\
& \text{‘Ali said (that) she/he}^{\text{ı/k}} \text{is intelligent’}
\end{align*}
\]

\[
\begin{align*}
\text{b. } & \text{ Ali } \quad [\text{pro}^{\text{ı/k}} \text{ akıllı} \quad \text{ol-duğ-u]-nu} \quad \text{söyle-di.} \\
& \text{Ali intelligent be-Nom-3SgPoss-Acc say-Past} \\
& \text{‘Ali said (that) pro}^{\text{ı/k}} \text{is intelligent’}
\end{align*}
\]

Again, if the local domain is taken to be the genitive possessive phrase due to the agreement, which is the non-finite embedded clause, in (31), \( pro \) behaves according to BT-B and picks an antecedent outside this local domain, e.g. ‘Ali’. In (31), on the other hand, the overt pronoun does not pick an antecedent outside the same local domain. In other words, it does not pick ‘Ali’ as its antecedent.

Whether this observation holds true only with genitive possessive phrases or not, it is also necessary to look at an embedded clause which is not in the form of a genitive possessive phrase. Following examples (from Gürel, 2002: 30) compare the overt and null pronoun in the subject position of a finite embedded clause:
And the following example compares the overt and null pronoun’s distribution in another type of non-finite embedded clause:

(34) Ali Ayşe’ye [o-nun-i/k / pro-i/k gel-me-sin]-i söyle-di.
    Ali Ayşe-Dat she/he-Gen come-Nom-3SgPoss say-Past
    ‘Ali told Ayşe that she/he-i/k/pro-i/k should come’

As the indices in the examples show, when used in subject positions in embedded clauses (whether in the form of a genitive possessive structure or not) and in the possessor positions of genitive possessive phrases, the distribution of pro and o differs. However, in object positions of the matrix clauses and embedded sentences, they show the same distribution.

The examples we have analyzed so far show that there are certain conditions such as genitive-possessive structures and subject positions of embedded sentences where pro and the overt pronoun do not show the same distribution. The following sections report on syntactic and pragmatic approaches to distributional differences for null and overt pronouns in Turkish.

4.2.1. The Avoid Pronoun Principle

Kornfilt (1991) considers pro as the representative of the syntactic class of pronominals and the overt pronoun as a marked form whose distribution is restricted by pragmatic as well as syntactic factors. When it is used as a subject, she claims that the overt pronoun is not able to co-refer with a potential antecedent. It is rather disjoint in its reference. Therefore, she prefers to call it an ‘obviative’ element. And she argues that the distributional properties of null and overt pronouns
in Turkish could be explained by the Avoid Pronoun Principle (APP) developed by Chomsky (1981) on the observation of PRO, later replaced with \textit{pro} for non-infinitive cases in pro-drop languages (Chomsky, 1982; Kornfilt 1991):


APP imposes a choice of a phonologically empty pronominal over an overt one. Kornfilt (1991) relates APP to agreement features in the language and provides a binding analysis for overt and null pronouns in Turkish. According to her analysis, a clause has weak agreement when its finiteness does not express all the relevant features such as number, person and Case. In such clauses an overt pronoun is used as a subject to satisfy the Case feature only. When the clause has all the relevant features (i.e., number, person, and possibly gender as well as Case), the clause has strong agreement, which is spelt out at the phonological level. The use of a null pronoun is preferred over an overt pronoun in such cases. When the embedded clause has strong AGR, it can function as a governing category (GC) since AGR can qualify as an accessible SUBJECT as a result of having all the relevant features. A weak AGR is a pseudo-AGR which cannot play a crucial role in Binding. It is only necessary to check Case feature. Kornfilt (1991: 68-69) provides two examples to show how APP influences Binding:

(36) \text{Askerler}_i \ [\text{pro}_{ij} \ \text{o\-ecek-lerin}]_e \ \text{inan-iyor-lar}.
Soldiers$_i \ \text{pro}_{ij} \ \text{die-Fut-3Pl-Dat} \ \text{believe-Prog-3Pl}
\text{‘The soldiers$_i$ believe that they$_{ij}$ will die.’}

(37) \text{Askerler}_i \ [\text{onlar-in-}i_{ij} \ \text{o\-ece\-g-in}]_e \ \text{inan-iyor-lar}.
Soldiers$_i \ \text{they-Gen-}i_{ij} \ \text{die-Fut-3Sg-Dat} \ \text{believe-Prog-3Pl}
\text{The soldiers$_i$ believe that they$_{ij}$ will die.}

In example (36), \textit{pro} has both a local and a non-local reading. However, its overt counterpart, \textit{onlar} (she/he-Pl), can only have a disjoint reading in (37). According to Kornfilt (1991) the
differences in their binding are related to the strength of AGR in the embedded clause. In (37), the third person plural agreement is not marked overtly. Thus, the embedded clause has a weak AGR feature. Since the AGR is weak, the embedded clause does not have a SUBJECT and it cannot function as a governing category. Therefore, the governing category is not the embedded clause but the matrix clause in (37). As the matrix clause is the GC, the overt pronoun, onlar, has a disjoint reading. It cannot pick ‘askerler’ (soldiers) as its antecedent since it is in the same GC with onlar (they). However, with the case of pro, the GC would differ. In the example (36), the embedded clause has a strong AGR since the embedded clause is marked overtly with a third person plural suffix. Since the strong AGR can function as SUBJECT, the embedded clause is predicted to be the GC. So, pro can pick ‘soldiers’ as its antecedent and this is compatible with BT-B because ‘soldiers’ is outside its GC.

Although Kornfilt’s predictions seem to work for cases similar to the ones above, there are other structures that need to be looked at. The following sentence is also from Kornfilt (1991: 67) and there is an overt pronoun despite the strong AGR, which would not be predicted by the APP:

(38) Askeleri [onlar-in*ij öl-ecek-lerin]-e inan-iyor-lar.
Soldiersi they-Gen*ij die-Fut-3Pl-Dat believe-Prof-3Pl
‘The soldiers believe that they *ij will die.’

In the sentence above, the AGR is strong enough to impose a choice of a null pronoun rather than an overt one. However, there is an overt pronoun. Furthermore, its binding is not influenced by the overtly marked, strong AGR. To put it differently, the verb in the embedded clause in (38) receives third person plural agreement marking overtly and therefore has a strong AGR which would provide a SUBJECT for the sentence. Thus, the embedded clause would be the GC and the local domain for the overt pronoun, onlar. If that was the case as Kornfilt’s APP version
suggests, the overt pronoun, *onlar*, would be able to pick ‘askerler’ (soldiers) as its antecedent since it is outside its local domain now. However, as the indices show, this is not the case. Neither the GC, nor the binding properties of the overt pronoun change with the existence of a strong AGR. Kornfilt’s application of the APP to Binding seems to fail to account for this observation.

Furthermore, in the sentences (31), (32), (33), and (11) which are repeated here as (39)-(40), if AGR is strong since the embedded clause is marked for third person singular, then the use of an overt pronoun is not predicted by the APP since a null pronominal would be preferred over the use of an overt one. In other words, the overt pronoun would be expected to be avoided. However, as the examples suggest both uses are available:

```
(39) Ali_i [o-nun/*i/*k/ /pro/*i/*k/ akıllı ol-duğ-u]-nu söyle-di.
    Ali she/he-Gen intelligent be-Nom-3SgPoss-Acc say-Past
    ‘Ali said (that) she/he/*i/*k/ /pro/*i/*k/ is intelligent’
```

```
(40) Ali Ayşe’ye [o-nun/*i/*k/*m/ /pro/*i/*k/*m/ yalancı ol-duğ-u]-nu söyle-di.
    Ali Ayşe-Dat she/he liar be-Nom-3SgPoss-Acc tell-Past
    ‘Ali told Ayşe that she/he/*i/*k/*m/ /pro/*i/*k/*m/ is a liar’
```

Furthermore, if *pro* is allowed, then the embedded clause marked with third person singular agreement marking would have strong AGR and would function as a GC so that *pro* can pick the subject of the matrix clause as its antecedent. Then, if that is the case, the same GC should be available to the overt pronoun as well and allow it to pick the subject of the matrix clause as its antecedent. However, this is not the case. It seems that the idea of strong AGR cannot handle the binding conditions of pronouns. Thus, we observe that the overt pronoun use seems to have a pragmatic function only which disambiguates the sentence between a local and non-local reading. However, binding relations of overt and null pronouns cannot be explained by the APP.
Furthermore, the APP is discussed to account for only the cases where *pro* and overt pronouns are used as subjects in embedded sentences (Kornfilt, 1991). In genitive-possessive constructions where *pro* or the overt pronoun is not used as subjects they still have different distributional properties. The examples are repeated below for convenience:

(41) a. \(\text{Ali} i\) o-nun\(_i^\text{k/m}\) kitab-ı-nı sev-iyor.  
    \(\text{Ali}\) she/he -Gen book-3SgPoss like-Prog
    ‘\(\text{Ali}\) likes her/his book’

    b. \(\text{Ali} i\) \(\text{pro}_i^\text{k/m}\) kitab-ı-nı sev-iyor.  
    \(\text{Ali} \) book-3SgPoss like-Prog like-Prog
    ‘\(\text{Ali}\) likes \(\text{pro}_i^\text{k/m}\) (her/his) book’

(42) a. \(\text{Ali'nin} i\) baba-sı\(_k\) o-nun\(_i^\text{k/m}\) kitab-ı-nı sev-iyor.  
    \(\text{Ali-Gen}\) father-3SgPoss she/he -Gen book-3SgPoss like-Prog
    ‘\(\text{Ali's}\) father\(_k\) likes her/his\(_i^\text{k/m}\) book’

    b. \(\text{Ali'nin} i\) baba-sı\(_k\) \(\text{pro}_i^\text{k/m}\) kitab-ı-nı sev-iyor.  
    \(\text{Ali-Gen}\) father-3SgPoss book-3SgPoss like-Prog
    ‘\(\text{Ali’s}\) father\(_k\) likes \(\text{pro}_i^\text{k/m}\) (her/his) book’

In addition, in cases where the embedded clause is finite as opposed to non-finite examples provided in Kornfilt’s analysis, *pro* patterns with the overt pronoun in its distribution in object positions and picks a non-local antecedent:

(43) \(\text{Ali} [\text{TP Ayşec} \text{pro}_i^*\text{k/m} / o-nun\(_i^*\text{k/m}\) sev-iyor] \) san-iyor.  
    \(\text{Ali}\) Ayşec she/he -Acc like-Prog think-Prog
    ‘\(\text{Ali}\) thinks that Ayşec\(_k\) likes \(\text{pro}_i^*\text{k/m}\) her/him\(_i^*\text{k/m}\)’

Thus, a better account which takes these cases into account as well as *pro* and overt pronoun as subjects is necessary to have a more comprehensive explanation for their distribution.
To summarize, so far, we have seen that pro and the overt pronoun show exactly the same distribution in:

(i) direct and indirect object positions where an overt pronoun is marked with accusative and dative cases respectively.

And pro and the overt pronoun differ in their distributions in positions such as:

(i) possessors in genitive-possessive constructions
(ii) subjects of finite and non-finite embedded clauses

As also proposed by Kornfilt (1991) pragmatic factors as well as syntactic factors play a role in the distribution of null and overt pronouns. The APP in its general approach is not rejected, here. Instead, applying the APP to the binding of pro in relation to strong and weak agreement in the clause seems not to be plausible when counterexamples are considered. As mentioned by Dimitriadis (1996) and Enç (1986), overt pronouns are indeed not redundant or optional in pro-drop languages. Nor do they seem to be used due to the weaknesses in the agreement paradigm in the sentence. Rather, they seem to play significant pragmatic functions within the discourse. The following section will provide brief information on the pragmatic functions of pro and the overt pronoun in Turkish.

4.2.2. Pragmatic factors for the distribution of the null and overt pronoun

Erguvanlı-Taylan (1986) and Enç (1986) mention discourse dependent uses of overt and null pronouns in Turkish. According to Erguvanlı-Taylan (1986), an overt pronoun functions as indicating emphasis or contrast and its occurrence is obligatory rather than optional under such discourse dependent conditions:
As examples (46) and (47) show, when there is no topic shift in the context, null pronoun use is preferred. Overt pronoun use results in ungrammaticality. Similarly, failing to use an overt pronoun when there is a topic shift also results in ungrammaticality as (48) and (49) indicate.

Thus, as the examples show, pro has the ability to refer to an existing antecedent within the context as well as within the sentence. However, the overt pronoun cannot refer to an already mentioned entity in the context and its use becomes even ungrammatical depending on contextual factors.
Applying the APP to the traditional BT and agreement paradigm of the language does not seem to account for the binding of pro and o in Turkish. As pragmatic factors have also shown, pro has the ability to refer to an existing antecedent within the context or within the sentence. The overt pronoun, on the other hand, shifts the topic and refers to an entity which is not mentioned within the sentence or the discourse. This difference in distribution might be related to their being different types of pronouns rather than being the null and overt counterparts of the same NP type. The following section will examine this and propose that null and overt pronouns do not belong to the same category.

5. Pro and Overt Pronoun as Different Pronominal Types

5.1. Pro as a weak pronoun and overt pronoun as a strong pronoun

Cardinaletti and Starke (1999) categorize pronouns into three different categories: strong, weak, and clitic. They assume that nominals can have CP and IP projections just like sentences:

Strong pronouns: \([CP [IP [NP pronoun]]]\)

Weak deficient pronouns: \([IP [NP pronoun]]\)

Clitic pronouns: \([N pronoun]\)

According to Cardinaletti and Starke (1999), the typology of pronouns is a set of forms in competition such that the form that is the most deficient is preferred over others since it has less structure. In other words, since clitics have less (syntactic) structure than weak pronouns (i.e., they are just N heads) and weak pronouns have less structure than strong pronouns, clitics are preferred over weak pronouns; and weak pronouns are preferred over strong pronouns when
available (cf. economy of representation, Chomsky, 1995). Thus, the order of deficiency is as follows:

critic < weak < strong

According to this categorization, an overt pronoun is a strong pronominal. Pro is a structurally deficient pronominal and belongs to the weak category. And the clitic is the most deficient with the least structure. If a deficient form is possible in a language, it will take precedence over a strong one. In other words, pro is always chosen over a strong pronoun, which is a similar approach to the APP (Chomsky, 1981).

Safir (2004) in his commentary on Cardinaletti and Strake’s (1999) work also suggests that there is a Weak Pronoun Competition (WPC) among pronominals. Accordingly, WPC selects the weakest pronoun as the optimal to represent a backgrounded, i.e., old information reading. To put it differently, WPC is a competition for topic reading, where only the winner can represent backgrounded topic reading. As a result, the weak form, i.e., pro, wins to represent the backgrounded reading.

According to Cardinaletti and Starke (1999) what makes a pronoun weak or strong is related to their range. Weak pronouns are incapable of bearing their own referential index, what is called as deictic potential in Safir (2004). This referential index is some sort of range restriction, which makes them either rangeless or they are associated with the range-restriction of an element prominent in the discourse. Only strong pronominals are full CPs (complement phrases) and only they contain [+human] specification under C, which gives them the referential index. Thus, they have the structure required for independent reference. However, deficient pronouns having no CP structure lack C and therefore they do not contain [+human] specification. Thus, they lack...
referential index and cannot make independent reference. Therefore, they are free to co-refer with any (prominent) antecedent.

Both Cardinaletti and Starke (1999) and Safir (2004) approaches are compatible with the pragmatic factors on null or overt subject selection in a consistent pro-drop language like Turkish. Unless the new sentence has a topic shift, a null pronominal is used which can only co-refer with an antecedent mentioned in the context. Use of an overt pronoun functions as a topic changer and results in an independent reference to another entity which is not previously mentioned in the context.

However, both the pragmatic and syntactic approaches reviewed so far focused on null subjects and overt subjects. As we have observed in section 4.1. a null and overt pronoun in object and indirect object positions of embedded and matrix clauses do not differ. If null and overt pronouns belong to different types, namely weak and strong respectively, they would also differ in their distribution in (in)direct object positions, as well. Thus, in the following section this difference is examined.

5.2. Pro as a weak pronoun in need to check its referentiality feature

According to Holmberg (2005), a generic interpretation via pro-drop is not available in consistent pro-drop languages. Instead, such languages resort to overt morphology (such as passive morphology in Turkish) or they use second person singular overt pronoun for generic interpretation. In partial pro-drop languages such as Finnish and Brazilian Portuguese, on the other hand, a null pronoun is either bound by (i.e., co-refers with) a higher DP or it has a generic interpretation. Brazilian Portuguese does not allow third person null subjects in a main clause as (50) shows. A null third person subject in an embedded clause would be bound by a higher
clause as in (51). When there is a third person null subject in the main clause, the meaning is
generic as in (52):

(50) Ele/*pro ganhou na loto.
    He won on-the lottery

(51) Pedro i disse que elei/proi*gandou na loto.
    Pedro said that he won on-the lottery

(52) Aqui não pode nadar.
    Here not can swim
    ‘One can’t swim here’

(Holmberg, 2005: 553)

Holmberg (2005) modifies Cardinaletti and Starke ‘s (1999) approach and claims that a null
subject is a φP. A φP is a deficient pronoun in the sense that it lacks referentiality and it has to
move to receive referential capacity. According to Holmberg (2005), Infl (I) has the feature [D]
that assigns referential capacity. In partial pro-drop languages if a φP does not move to Spec,IP
and remains in Spec,vP, it cannot co-refer with the subject of the matrix clause. This results in a
generic interpretation such as (52). In consistent pro-drop languages such as Turkish, Spanish
and Greek, however, a generic interpretation is not available with a null subject. According to
Holmberg (2005) a pro always checks its referentiality by moving to Spec, IP. It receives the
ability to refer to an entity or group from I containing D via movement. As a result, it will be
able to co-refer with a prominent antecedent in the discourse or in the matrix clause. According
to him, the relation between I and φP is an ‘Agree’ relation. φP has an unvalued D, ([uD])
feature which needs to get checked under I which has this feature. If pro does not move to
Spec,IP and does not check [uD] feature this will result in a generic reading with a null pronoun.
To explain the different and similar distributions of *pro* and *o* in Turkish, proposals of Cardinaletti and Starke (1999) and Holmberg (2005) will be followed. Accordingly, *pro* is a weak pronoun in the sense that it lacks referential index and it cannot show independent reference. It can only co-refer with an antecedent in the context. However, to do that it needs to check its referentiality, [D], feature with a head bearing this feature since *pro*’s [D] feature is uninterpretable or unvalued, i.e., [uD]. *Pro* needs to get [uD] feature checked so that it has the capacity to co-refer with an entity. As proposed by Holmberg (2005), IP bears the [D] feature, (cf. Chomsky, 1995) and it can check or value *pro*’s [uD]. Therefore, *pro*’s [uD] feature is checked under IP. As a result, *pro* could co-refer with an antecedent within the discourse rather than having a generic interpretation.

This approach can explain the referential dependency of *pro* in subjects of embedded and matrix clauses:

(53) [Ali, Ayşe’ye] [o-nun-*i/k* / *pro*/i/k gel-me-sin]-i söyle-di.  
Ali Ayşe-Dat she/he-Gen come-Nom-3SgPoss say-Past  
‘Ali told Ayşe that she/he-*i/k*/pro should come’

(54) [Ali] [o-nun*/*i/k* / *pro*/i/k nehir kenar-in-da koş-ma-sin]-1 sev-iyor.  
Ali she/he -Gen river side-3SgPoss-Loc run-Nom-3SgPoss-Acc like-Prog  
‘Ali likes her/his-*i/k*/pro running by the river side’

(55) [Ali] [o/*i/k* / *pro*/i/k iyileş-ti] san-iyor.  
Ali she/he recover-Past think-Prog  
‘Ali thinks that she/he-*i/k*/pro has recovered.’

(56) [Ali] [o/*i/k* / *pro*/i/k nehr-in kenarinda koş-uyor-du] diye bil-iyor.  
Ali she/he river-Gen side-3SgPoss-Loc run-Prog-Past Sub know-Prog  
‘Ali thinks that she/he-*i/k*/pro was running by the river side.’

In sentence (53) and (54), *pro* and the overt pronoun occupy the subject position of a non-finite embedded clause. In (55) and (56), they are in the subject position of a finite embedded clause.
In both cases, *pro* co-refers with the subject of the matrix clause, ‘Ali’ as well as another antecedent in the context (e.g., Ayşê in (53)). Applying Holmberg’s analysis to the differences observed in the sentences above, it could be argued that *pro* in need of checking its [uD] feature moves up to the first available IP (i.e, the embedded clause) which is the host for D feature, i.e., which can check this feature. As a result, it can co-refer with an antecedent above the IP, e.g., ‘Ali’. The overt pronoun, *o*, on the other hand, is strong and already bears the D feature. In other words, it has the capacity to make independent reference and to introduce a new topic within the discourse. Thus, it does not co-refer with the most recent topic, namely, ‘Ali’.

If we assume that a new clause, i.e., an IP, sustains the topic of the discourse mentioned before, it makes sense for an IP to co-refer with an antecedent mentioned previously in the discourse. If *pro* checks this feature under an IP it will consequently co-refer with the topic of a clause mentioned before the IP it checks its [uD] feature with.

In the following sentences, *pro* and the overt pronoun share the same distribution. However, this similarity does not make a *pro* a strong pronominal in such positions. Instead, applying the same rationale to direct object positions will give us the conclusion that *pro* is still a weak pronominal and overt pronoun is a strong one in object positions.

(57) \(\text{Ali}_{i} \ [\text{Ayşê}^{'n}i\text{n}_{k/m} \text{ o-nu}^{'i/k/m} / \text{pro}_{i}^{\ast/k/m} \text{ sevdiğin}]^{-i} \text{söyledi.}\)
\(\text{Ali} \ \text{Ayşê-Gen she/he-Acc} \ \text{like-Nom-3SgPoss-Acc} \ \text{say-Past}\)
‘Ali said that Ayşê likes her/him\(i^{i/k/m} / \text{pro}_{i}^{\ast/k/m} \).’

(58) \(\text{Ali}_{i} \ [\text{Ayşê}_{k} \text{ o-nu}_{i/k/m} / \text{pro}_{i}^{\ast/k/m} \text{ sev-iyor}] \text{ san-iyor.}\)
\(\text{Ali} \ \text{Ayşê} \ \text{she/he-Acc} \ \text{like-Prog think-Prog}\)
‘Ali\(i_{i} \text{thinks that Ayşê}_{k} \text{ likes her/him}\(i^{i/k/m} / \text{pro}_{i}^{\ast/k/m}\).’

In the sentences above, *pro* checks its referentiality under the IP, as a result, it can co-refer with an antecedent above the IP (e.g., ‘Ali’) rather than an antecedent within the IP (‘Ayşê’). In other
words, the IP is a continuation of the statement or the topic mentioned previously. Thus, checking its [uD] feature under IP, pro can co-refer with an antecedent that the IP does. Therefore, rather than co-referring with an antecedent within the IP such as ‘Ayşè’, pro co-refers with an antecedent that is mentioned before the IP under which its feature is checked, such as ‘Ali’. As for the strong pronoun, o, it has the referential index already, which gives it the capacity to shift the topic. Thus, it introduces a new topic which is not mentioned within the IP it is uttered. Thus, it is independent in its referentiality and does not pick up the most recent antecedent in the IP, i.e., ‘Ayşè’, but it can refer to another antecedent, which is not the most recent one, namely, ‘Ali’.

Similarly, as the indices in sentences (59) and (60) show, pro and the overt pronoun cannot co-refer with the matrix subject of the matrix IP when they are used in direct and indirect object positions of a matrix clause:

(59) Ali o-nu*i/k/m / pro*i/k/m sev-iyor.
    Ali she/he-Acc like-Prog
    ‘Ali likes her/him*i/k/m / pro*i/k/m.’

(60) Ali o-na*i/k/m / pro*i/k/m bak-iyor.
    Ali she/he-Dat look-Prog
    ‘Ali looks at her/him*i/k/m / pro*i/k/m.’

Similar to the direct and indirect objects in embedded sentences, pro can co-refer with an antecedent after checking its [uD] feature under matrix IP. Therefore, it has to co-refer with an antecedent mentioned before the IP it checks its features with. As for the strong pronoun, as mentioned before, it has its referential index and introduces a new topic. Thus, it does not co-refer with the most recent topic.
5.3. Summary and Conclusion

This paper has presented a problematic reflexive marker in Turkish, *kendisi*, which has kept researchers busy for a while as it deviated from the BT proposed in the GB Theory. It is proposed that *kendisi* does not pose a challenge the BT proposed in the GB Theory. The true anaphor for third person singular in Turkish is *kendi*, which conforms to the BT. *Kendisi*, marked with third person singular agreement is preceded by a null pronoun as a common noun possessed by *pro*. Therefore, *kendisi* is not a problematic reflexive. Instead, its distribution is influenced by *pro* as its possessor.

*Kendisi* shows different distributions when preceded by a null and overt pronoun. Further data show that *pro* and its overt counterpart, *o* do not show the same distribution generally although they have been considered to be the null and overt counterparts of the same NP type by the traditional grammars. In other words, both the null and overt pronouns have been reported to have [-anaphoric, +pronominal] features via which they would be subject to the BT-B.

The analysis in this paper has proposed that *pro* and *o* are not the same NP types. There are pragmatic as well as syntactic factors that guide their distribution. *Pro* is a weak pronoun which can only co-refer with an antecedent in the context. The overt pronoun, *o*, on the other hand, is a strong pronoun and has a referential index which makes it possible for the overt pronoun to show independent reference. Thus, rather than co-refering with an antecedent mentioned in the context, *o* has its own reference. What gives *pro* the capacity to co-refer with an antecedent mentioned in the context is the process of checking its uninterpretable D, i.e., [uD] feature under IP. As a result of its [uD] feature checking under IP, it has the ability to co-refer with an individual, which is mentioned before the IP it checks its [uD] feature with.
References:


