Agreement as a Post-syntactic Phenomenon: Revisiting Turkish Nominal Compounds with -(s)İ(n)

Deniz Tat
University of Arizona

In this paper, I examine Turkish NN%N compounds bearing the -(s)İ(n) suffix and consider two commonly assumed grammar architectures for the derivation of compound words. In Lexicalism, a pre-syntactic grammar component is assumed to be responsible for deriving words, including compound nouns. Studies of Turkish NN%N compounds within the lexicalist framework generally take -(s)İ(n) to be a linker, historically derived from the 3rd person possessive agreement, specified for deriving NN%Ns. In a syntactic framework, non-idiosyncratic versions of such compounds are claimed to be derived entirely in the syntax and their structure is analogous to possessive phrases. In these studies, -(s)İ(n) either heads an AgrP or is derived by means of an AGREE relationship between the possessor and the possessee. I show that neither of these frameworks can fully account for Turkish NN%N facts, and propose that a post-syntactic morphology component must be responsible for the derivation of -(s)İ(n), extending the work of Bobaljik (2008) to Turkish.

1. Introduction
A first look at Turkish NN%N (henceforth NN) reveals that there is a suffix at the right-hand edge of the second constituent, with a set of exceptions. These compounds bear a phrasal stress. (See Kamali and Ikizoglu, this volume, for a detailed analysis). This suffix, which is also the 3rd person possessive marker, is only found in NN. Consider the following examples, where we have a semantically compositional endo-centric compound (1) and a semantically opaque exo-centric compound (2):

(1) masa-örtü-*(sü)
    table-cloth-SIN
    ‘table cloth’

(2) akşam-sefa-*(si)
    evening-joy-SIN
    ‘four-o’clock’ (a type of plant)

The status of this suffix has been a popular issue of debate in the previous literature. It has been claimed that it is an agreement marker just like its possessive counterpart (e.g. Kornfilt 1984; Yükseler, 1998; Dede 1978), a compound linker (e.g. Goksel and Haznedar 2007; Goksel 2009) or an ezafe (van Schaik 2002), a case-marker found in languages such as Persian, which marks the link between the head noun and its modifier as well as its possessor. This suffix has also been analyzed as a special nominal categorizer specified for compounds (Kharytonava 2011). Studies that consider -(s)İ(n) as a compound linker typically assume a pre-syntactic morphology component, which provides input to the syntax. (Kharytonava (2011) is an exception to this.) Studies that take this suffix as an agreement marker emphasize parallelisms between an NN and a possessive phrase, showing that the genitive-free possessor is analogous to differentially marked accusative case in object positions, such that the non-specific NPs are not morphologically accusative-marked (Kornfilt 1984). In this study, I agree with the position that -(s)İ(n) is an agreement
marker, but I claim that its derivation, takes place in a post-syntactic component, which takes syntactic forms as inputs and turns them into morphological forms before they receive a phonological form. In doing so, I assume the grammar architecture proposed in Distributed Morphology (henceforth DM) (Halle and Marantz 1993, Harley and Noyer 1999), and specifically, I assume that case and agreement are *dissociated morphemes* (Embick 1997, Bobaljik 2008).

In section 2, I present five observations regarding the behavior of NN. In section 3, I consider each of the premises given in section 2, assuming a pre-syntactic component for deriving compounds in section 3.1 and a syntactic component for deriving them in section 3.2. In section 3.3, I propose that compounds are derived in syntax but the -(s)I(n) suffix is realized post-syntactically. In section 4, I provide my concluding remarks.

### 2. Premises

#### 2.1 Premise 1: The plural

When NN compounds are pluralized, the plural marker does not follow, but instead precedes -(s)I(n). We see that this is again true of both compositional compounds, as in (5), and non-compositional ones, as in (6):

(3)  
- a. *mása-örtü-sú-ler  
  table-cloth-SIN-PL  
  ‘tablecloths’
- b. masa-örtü-ler-i  
  table-cloth-PL-SIN  
  ‘my table cloth’

(4)  
- a. *akşam-sefa-su-lar  
  evening-joy-SIN-PL  
  ‘four-o’clocks’
- b. akşam-sefa-lar-t  
  evening-joy-PL-SIN  
  ‘neighbor’s four-o’clock’

Regardless of whether we have a compositional compound like (3) or a non-compositional, idiosyncratic compound like (4), the plural marker precedes -(s)I(n), which must be accounted for by any theory of Turkish nominal compound formation.

#### 2.2 Premise 2: Possession

When an NN compound is in the possessee position in a possessive phrase, we have a puzzling case, where -(s)I(n) cannot co-occur with the possessive agreement marker (see Footnote 2). Consider the following examples:

(5)  
- benim masa-örtü-(sú)-m  
  my table-cloth-SIN-1SG.POSS  
  ‘my street cat’

(6)  
- komşu-nun akşam-sefa-(*sú)-s  
  neighbor-GEN evening-joy-SIN-SIN  
  ‘the neighbor’s four-o’clock’

-(s)I(n) cannot co-occur with a possessive marker. Any given account of NN compound derivation in Turkish must explain why this is the case.

#### 2.3 Premise 3: Possessive-free genitives

In Turkish, the possessive markers in a possessive phrase do not always surface. (See Öztürk, Eruğvantı Taylan and Zimmer, this volume.) Crucially, in cases where the possessor of a possessive/genitive phrase is occupied by a NN compound, the -(s)I(n) suffix obligatorily surfaces, as exemplified in (7).

(7)  
- bizim sokak kedi-*(sú)  
  our street cat-SIN  
  ‘our street cat’
What example (7) shows is that even in certain types of possessive/genitive phrases, -(s)I(n) is required if the possessive marker of the selecting PossP is lacking.

2.4 Premise 4: Suspended affixation
Turkish has the property of what has come to be known as “suspended affixation” (Lewis 1967), which is found in coordinated constructions where a grammatical ending shared by two or more conjuncts is suspended to the last conjunct. Consider the following example, where a possessive marker is suspended:

(8) **anne ve baba-mız**

‘our mother and father’

Coordination of NN compounds constitutes another puzzling case that warrants an explanation. When two such compounds are coordinated under the scope of a possessive phrase, such that the possessive marker is suspended, the initial conjuncts bear the -(s)I(n) suffix while the final conjunct does not:

(9) **gaz dedektör-ü ve hırsız alarm-imız**

‘our gas detector and anti-theft alarm’ (Kharytonava 2011:232)

2.5 Premise 5: m-reduplication
Let us finally look at a peculiar case of reduplication, m-reduplication (henceforth, m-red) in Turkish and how it interacts with Turkish NN compounds. This kind of reduplication in Turkish targets nouns and leads to the semantic interpretation of “and the like”:

(10) a. böcek mőcek
    b. art mär
    ‘bug(s) and the like’
    ‘bee(s) and the like’

In (11), we can see that both constituents of a compositional NN compound as well as the compound in its entirety are legitimate targets of m-red. Our theory of NN compounds must also account for these observations:

(11) a. **sokak kedi-si**
    street cat-SIN
    ‘street cat (=stray cat)’
    b. **sokak mokak kedi-si**
    street M-RED cat-SIN
    ‘cat(s) on streets and the like’
    c. **sokak kedi-si medisi**
    street cat-SIN M-RED
    ‘cat(s) and the like on streets’
    d. **sokak kedi-si mokakkedi**
    street cat-SIN M-RED
    ‘Cinderella and the like’

(12) a. **Kül kedi-si**
    ash cat-SIN
    ‘Cinderella’
    b. *Kül mülä kedi-si
    ash M-RED cat-SIN
    c. **Kül kedi-si medisi**
    ash cat-SIN M-RED
    ‘Cinderella and the like’
    d. **Kül kedi-si mülkedisi**
    ash cat-SIN M-RED
    ‘Cinderella and the like’
3. Analysis

3.1 -s(s)n as a pre-syntactic morpheme

The Lexicalist Hypothesis is based on the assumption that word formation and phrase formation belong to two independent components of grammar and that there is a strict division of labor between them (Chomsky 1970, 1995; Lieber 1980; Williams 1981; Anderson, 1992; Lapointe, 1981; Di Sciullo and Williams 1987). It comes in a strong and a weak variety. The strong variety assumes that all kinds of words are formed in a component preceding syntax while the weak variety assumes that some words, either those that are idiosyncratic, unproductive or those that are only derivational, are formed pre-syntactically.

The strong version of the Lexicalist Hypothesis makes a strict distinction between so-called words and phrases. Words, in this version, regardless of whether they contain an inflectional morpheme or a derivational morpheme, are claimed to be derived in a component that provides input to the syntax, which cannot manipulate these words, see their derivational histories or refer to them.

Premise 1, which is based on the observation that the plural marker has to precede the -(s)s(n) suffix, as shown in examples (3) and (4), constitutes a challenge for strong lexicalism. First of all, it is a challenge for any structural analysis as it is not clear why a plural marker that has scope over the entire compound noun precedes the linker.

Ralli (2008) points out that everything that can follow -(s)s(n) (i.e. case) is a contextual inflection while the plural marker, which precedes -(s)s(n), is an inherent inflection in the sense of Booij (2005). The former type is a kind of inflection that is dependent on the syntactic context, such as an accusative case requirement of a direct object, while the latter type is freely available to the speaker and is not dependent on any syntactic requirement. However, -(s)s(n) can be preceded by a type of plural marker that is actually contextually required, such that it has to agree with the possessor. Consider the following pair:

(14)  a. [on-lar-m] [araba-st]\n    3-PL-GEN car-SIN
    ‘their car’

b. [pro [araba-lar-t]]
   car-PL-3.POSS
   ‘their car’

In Turkish possessive phrases, the possessor can be pro-dropped. If it is a 3rd person plural possessor, then the number feature is represented by the plural marker. When this possessor is pro-dropped, then the plural marker is lowered as in (14b). In such examples, the plural marker that precedes -(s)s(n) is a type of contextual inflection as it is part of possessor-possessee agreement morphology. This shows that an element that precedes -(s)s(n) can potentially be of the contextual/syntactic type too, leaving the plural puzzle unresolved.

Premise 2, which shows that -(s)s(n) cannot co-occur with possessive markers, as exemplified in (5-6), is probably the biggest challenge to the Lexicalist Hypothesis among the five premises. Whether -(s)s(n) is deleted entirely in that context or replaced by the possessive marker is not relevant; these examples show that syntax can manipulate a linguistic unit that is lexically-derived.

Premise 3, the observation about possessive-free genitives, such that -(s)s(n) has to surface when there is no agreement marker as in (7), is actually what strong lexicalism would predict for all cases of possessive phrases with NN compound possessees. Premise 3, in this sense, does not cause an immediate challenge to strong lexicalism but, of course, this flimsy character of -(s)s(n), which we also observe in Premise 4, needs to be explained. We
have seen that under suspended affixation, where the possessive marker is suspended, the final conjunct’s -(s)l(n) cannot co-occur with the possessive marker while the non-final conjuncts all retain -(s)l(n). From this perspective, strong lexicalism faces a similar problem since there is at least one conjunct whose compound marker is manipulated.

The final premise, *Premise 5*, regarding *m*-reduplication, shows that both constituents of an NN compound are permissible targets for reduplication as well as the compound noun in its entirety, as shown in (11). However, we have seen that there is at least one restriction. If an NN compound is semantically transparent, then the previous statement is true; however, if we have a semantically opaque compound as in (12), then the second constituent (head noun) and the entire compound are possible targets of *m*-reduplication, while the first constituent (non-head) is not. Note that the reduplicated forms include -(s)l(n) in the relevant examples in (11) and (12). If -(s)l(n) is part of the lexical derivation, then this reduplication can potentially happen anywhere in the grammar architecture, either pre-syntactically or syntactically, or possibly post-syntactically. In this regard, a pre-syntactic analysis of NN compound formation can make the correct predictions about *m*-reduplication provided we assume that such reduplication happens pre-syntactically. But our theory would then need to stipulate why (11b) is allowed while (12b) is not.

The weak version of lexicalism would face the same problems as strong lexicalism, as a linguistic unit that is part of the lexical derivation of compounds -(s)l(n) is subject to syntactic manipulation in the first four of the premises discussed so far. *Premise 1* would especially be more problematic for weak lexicalism since the plural marker would in this framework belong to the syntax, and yet it can precede a lexical, derivational suffix. The version of weak lexicalism that considers non-compositional, idiosyncratic formations to be lexical, and compositional formations to be syntactic would also run into problems since, as we have seen in section 2, both of these types of compounds are equally subject to the five premises we have discussed so far. As for *Premise 5*, the kind of stipulation we have made for strong lexicalism above will not hold because *m*-reduplication is a productive type of inflection that must belong to the syntax.

### 3.2 -(s)l(n) as a syntactic morpheme

Let us now assume that -(s)l(n) is an agreement marker and it is derived entirely by syntactic means. There are at least two types of analysis of agreement in syntax in the literature. In the first type, agreement is represented categorically in the syntax, such that an AgrP is headed by an Agr that bears some agreement features. This view of agreement was widely accepted after Pollock’s (1989) split-INFL proposal that T(ense) and Agr(eement) are separate projections. The second type became widely accepted after Chomsky (1995) dispensed with Agr phrases on grounds that only those functional projections whose interpretable features are visible to LF (e.g. CP, TP, vP) must have categorical representations, and Agr phrases, which bear uninterpretable features, are not among them. In this view, agreement morphology is realized by means of a checking, or AGREE relationship between a lexical head bearing interpretable φ features (e.g. number, person, gender) and their uninterpretable counterparts in a minimal domain.

Let us assume that the structure of an NN compound is follows, leaving aside the differences between NPs and DPs as well as the further internal structure of this NP/DP. In this structure, the possessor NP moves to spec, AgrP and checks its number and person features in a spec, head relationship before moving to the specifier of the PossP. The
possessee, on the other hand, head-moves to Agr and receives the agreement morphology before moving to the position where it heads the compound NP(PossP).

\[(15)\]

This structure would face a series of challenges, too. First of all, the Agr head in this structure must be dominated by a NumP for the structure to receive a plural marker and for that marker to precede the -(s)I(n) suffix. Cross-linguistically, this would result in an unusual order of morphemes. Considering a number of languages, Alexiadou, Haegeman and Stavrou (2007) presents the following extended order of phrases in the nominal domain:

\[(16)\]  
\[
\text{DP} > \text{AgrP} > \text{Num(ber)P} > \text{Gen(der)P} > \text{nP} > \text{NP}
\]

If the hierarchy in (16) is correct, then -(s)I(n) has to precede the plural, but the examples in (3-4) indicate that it is the other way around.

**Premise 2**, which shows that -(s)I(n) cannot co-occur with possessive phrases, posits a challenge for an AgrP analysis of compounds, too. If the structure in (15) is selected by a possessive phrase as a possessee, there would be nothing in the structure that disallows structures such as in (17) and (18).

\[(17)\]  
\[
*\text{benim masa-örtü-sün-m} \\
\text{my table-cloth-SIN-INSG.POSS}  \\
\text{‘my table cloth’}
\]

\[(18)\]  
\[
*\text{onun akşam-sefa-sı-sı} \\
\text{his/her evening-joy-SIN-SIN}  \\
\text{‘his/her four-o’clock’}
\]

**Premise 3** shows that, in cases where the NN compound possessee lacks a possessive agreement marker, then -(s)I(n) obligatorily surfaces. Öztürk, Erguvaş Taylan and Zimmer (this volume) (henceforth ÖEZ) provide a syntactic analysis and show that there are, as it happens, two types of such. (Instead of AgrP, they propose a complex D+P head.) For this study, their Type I is relevant. Let us now assume their structure and locate an NN compound in the possessee position, where the FP represents a discourse phrase similar to a topic phrase. In this architecture, the possessor first moves to Spec, DP to acquire a referent status before checking the relevant discourse features in Spec, FP, as shown in (19). If the lack of possessive morphology in such constructions, despite the fact that the possessor NP moves to Spec DP, is properly motivated, then the structure proposed by proposed by ÖEZ correctly predicts the observation described in **Premise 3**. The structure in (19) is adapted for example (7):
Premise 4 is a challenge to the AgrP analysis of -(s)İ(n), too. In fact, any theory failing to account for Premise 2 will also fail to explain Premise 4 since both of these include a compound-final -(s)İ(n), which cannot co-occur with a possessive marker. Premise 5 can be explained by putting forward some kind of a ‘vagueness’ feature on N that is realized at PF by reduplication. This feature must be interpretable since its semantic consequences (e.g. kedi medi, “cats and the like”) are also relevant to LF. However, this analysis runs into a problem when we compare (11) to (12). The challenge would be to account for the observation that the first constituent of (12) cannot be reduplicated while the second one can be, despite the fact that they are both semantically opaque. One can argue that, since the second noun is the head, and its reduplication can have semantic consequences over the entire compound, it can be reduplicated unlike the first constituent. However, this analysis must then assume that an idiosyncratic compound, such as (12), is syntactically derived, so that its constituents are visible to syntax, an assumption not typically made in the GB/Minimalism literature, which assumes some version of weak lexicalism or another.

The AGREE analysis of -(s)İ(n) appears to be a better alternative then and AgrP analysis since, for instance, we can easily account for Premise 1 if checking of agreement features takes place before the external merge of the NumP for the plural. However, Premise 2 and Premise 4 would still cause some challenge because the possessor NP and the higher constituent of the NN compound would be carrying two different sets of interpretable features, both required to check against their uninterpretable counterparts. In this framework, a set of interpretable features can check against multiple sets of uninterpretable features (e.g. Romance adjective phrases, where the interpretable features on the nominal head can be checked against uninterpretable features on multiple adjectives). However, in the Turkish case, if either one of the interpretable sets of features are checked, then the other set will be left unchecked, and therefore, the derivation crashes. Premise 3, on the other hand, will have the same problem discussed for the AgrP analysis above, such that the possessor moving to FP without checking its agreement features at Spec DP must be satisfactorily explained.

We have seen that both pre-syntactic and syntactic analyses of -(s)İ(n) runs into a set of problems. The logical conclusion would be then to consider the third alternative, which considers agreement to be a post-syntactic phenomenon. Let us consider DM, a theory, which assumes a post-syntactic morphological component in the architecture of grammar.
3.3 Proposal

DM (Halle and Marantz 1993, 1994; Harley and Noyer 1999) is a piece-based theory of grammar, which rejects the Lexicalist Hypothesis (e.g. Di Sciullo & Williams 1987). In this theory, syntax manipulates Roots (primitives with some semantic and phonological content) and Abstract morphemes (functional morphemes), but vocabulary insertion takes place only after these syntactic operations are completed (i.e. “Late insertion”). (20) shows the grammar architecture commonly assumed in the DM literature:

(20) Syntax (move, merge, copy)       Certain morphemes that do not 
Morphology                                    figure in syntax proper, such as case 
PF                                             (Marantz 1991), can be inserted at the 
     LF                                        Morphology component after syntax, but 
                                                  before any morpheme receives phonological form at PF (Harley and Noyer 1999).

Bobaljik (2008) proposes that agreement morphemes are a type of dissociated morpheme. Since agreement is triggered by morphological case (m-case), which is a kind of dissociated morpheme, agreement must be realized post-syntactically, too. Using examples from several languages like Icelandic and Hindi, he shows that there is a hierarchy of m-case, which can show different accessibility properties in different languages. In Turkish, the only accessible case for agreement in the verbal domain is the nominative (hence only subject-verb agreement).

I propose that we can extend Bobaljik’s (2008) work to the nominal domain for languages with nominal agreement. In the Turkish nominal domain, it is the genitive that triggers agreement. Possessors and subjects of nominalized clauses bear a genitive marker, which can also take a null form depending on certain semantic and morphosyntactic constraints on the genitive-marked NP (see Kornfilt 2008 for a detailed analysis of these). Bobaljik (2008) shows that, in cases where there are more than one accessible NPs marked with the same case, it is always the higher NP that becomes the controller of agreement.

Bearing this in mind, we can propose that -(-s)İ(n) and other agreement markers cannot co-occur because they are of the same type (both are agreement markers), and in cases such as Premise 2, where an NN compound is the possessee of a PossP, they would be competing for the same position of exponence. Following Bobaljik (2008), we can predict that in such cases, the highest accessible NP is the controller of agreement. This is borne out by examples like (5) and (6), where the possessee agrees with the highest NP, benim, “my” and komşunun, “neighbor’s,” respectively. Note that, in this analysis, the -(-s)İ(n) suffix of the NN compounds is not deleted under a PossP; in fact, it never gets realized. Because there is only one position of exponence where the morphology component can insert an agreement marker, there is no stage at the derivation where -(-s)İ(n) and another agreement marker co-occur. (There is independent evidence that this analysis better captures the psychological reality of -(-s)İ(n), such that aphasics speakers of Turkish do not make any errors in which -(-s)İ(n) and an agreement marker co-occur. See Tat (upcoming) for details.)

In suspended affixation examples of Premise 4, only the last conjunct is in a local domain to the possessor, and therefore, in non-final conjuncts, the controller of agreement would be the first constituent of the NN compound. To account for such locality, we would
have to reject a flat representation of coordinate structures, and instead, assume a binary, left-branching structure where the highest conjunct is the most local one to the PossP.

*Premise 1*, the observation that the plural marker precedes -(s)İ(n), can also be accounted for in this post-syntactic analysis of agreement. In the grammar architecture assumed here, post-syntactic morphology, by nature, follows syntactic representations.

*Premise 5*, showing that both constituents of a compositional NN compound and only the second constituent of an idiosyncratic compound can be targets of m-red, can also be accounted for within DM. In this theory, access to idiomatic meanings, termed as “Encyclopedia” (Harley and Noyer 1999), is accessible only after all syntactic and morphological operations are completed. For example, an idiom, such as *kick the bucket*, which means “die,” is syntactically compositional, such that the pieces of this idiom are independently represented in syntax, but its idiomatic meaning is only available, when these pieces are in each other’s context. By the same token, we can say that, the idiomatic meaning of *Külekedi* “Cinderella” is still available when we m-reduplicate the second constituent as *Kül kedisi medisi* because *kül* would still be in the context of *kedisi*, resulting in the idiosyncratic meaning. On the other hand, m-reduplicating the first constituent will destroy the context for the idiomatic reading since the m-reduplicated piece would be in-between the first constituent and the second constituent.

The only challenge to the post-syntactic analysis of -(s)İ(n) comes from the observation in *Premise 3*. When the possessor moves to the specifier position of an FP, the NP marked for genitive, would still be an accessible trigger for agreement. This premise still awaits an explanation. However, if the analysis in (19) is correct, the NN compound possessee would receive -(s)İ(n) since another NP that can trigger agreement would not be available in the first place.

4. Conclusion
We have seen that both a pre-syntactic and a syntactic analysis of -(s)İ(n) run into a series of problems considering the premises we have listed in section 2. It appears that a post-syntactic analysis of -(s)İ(n) has a better explanatory power. Extending the work of Bobaljik (2008) to agreement in the Turkish nominal domain shows that the NP/DP robustly correlates with CP in this language, an issue that warrants a thorough investigation.

5. References


Kamali, B. & Ikızoğlu, D. (this volume) Against compound stress in Turkish.


---

1 There is a small set of exceptions to this, which consists of a vanishingly rare number of examples: *ayakkabı*, ‘shoe,’ *hamameli*, ‘honeysuckle,’ *divealtı*, ‘submarine,’ *onbaşı, yüzbaşı, binbaşı*, ‘corporal, captain, major.’ These are maximally lexicalized forms that could be analyzed as Roots in DM.