Inflectional morphology in Turkish VP-coordination

Antske Fokkens[♣] **Laurie Poulson**[♢] **Emily Bender**[♢]
[♣] Department of Computational Linguistics, Saarland University
[♢] Department of Linguistics, University of Washington

1 Introduction

This paper presents an analysis of the interaction between verbal morphology and VP coordination in Turkish. The interaction is of interest because of three general properties of the language: (i) VP conjuncts must agree in their tense and aspect features, (ii) tense/aspect (TAM) morphemes may appear only on the final conjunct, either obligatorily (in the *-ip* construction) or optionally (in *ve* coordination), and (iii) verbal affixes expressing necessity and possibility also take scope over the whole coordinated structure while attaching morphologically only to the final conjunct. Generalizations (i) and (ii) are shown to be straightforwardly analyzed in terms of local constraints on identity of TAM features. Generalization (iii) is more surprising, and would seem to point to the necessity for a notion of phrasal affixes. In fact, however, it is amenable to a constructional analysis along the lines of (Tseng, 2003).

The analyses proposed in this paper are implemented in a grammar fragment derived from the LinGO Grammar Matrix (Bender et al., 2002), through its web-based customization system.[1] Our grammar fragment focuses on capturing syntactic and semantic constraints on verbal morphology in coordination and uses semantic representations in the format of Minimal Recursion Semantics (Copestake et al., 2005, MRS). Like all Matrix-derived grammars, it is compatible with the LKB (Copestake, 2002).

The implementation confirms that the analyses manage to account for all examples presented below. It also shows the cross-linguistic applicability and the practical utility of the Grammar Matrix, because we were able to implement and test the analyses presented in this abstract quickly.

2 Verbal Morphology in Turkish

2.1 Properties of Turkish Verbs

This section describes some basic properties of verbal morphology in Turkish. We provide an overview of morphemes that may be added to the stem and present conditions of completeness and well-formedness of the verbs. The description is based on, among others, Kornfilt (1997) and Lewis (2000) and we base our interpretation of the data partially on the descriptions provided by Sezer (2001) and Kabak (2007).

Table 1: Morpheme slots of inflectional morphemes

1	2	3	4	5
-DI def. witn. past *-sE* subj. cond. *-mIş* inferential pres. perf *-Iyor* continuous *-yEcEG* future *-Ir/-Er* aorist *-mEli* neccesitive *-mEkte* continuous	*-(i)DI* def. witn. past *-(i)sE* ind. cond. *-mIş* inferential	*-(i)sE* ind. cond *-(i)mIş* inferential	AGR	*-DIr*

The distinction between derivational and inflectional morphemes is not clear cut in Turkish. Traditionally, morphemes that can be followed by the infinitive marker *-mEk* are considered derivational. Derivational morphemes are *-DIr/t* (causative), *-Il* (passive), *-(y)A* (abilitative), *-mA* (negation) and *-(y)Abil* (potentiality). In addition to the derivational morphemes, there are five slots that may host an inflectional morpheme. The inflectional morphemes are presented in Table 1. In the representation, capital letters represent phonemes whose realization depends on vowel harmony. A finite verb must bear an inflectional marker from slot 1 and an agreement marker. At least one inflectional marker must be phonologically overt (Kabak, 2007).[2]

Turkish has two paradigms of agreement markers: the *k*-paradigm for definite past and conditional (*-DI* and *-sE*, respectively) and the *z*-paradigm for all other TAM[3] morphemes. Which paradigm is used depends on the last TAM morpheme attached to the verb.

[1] http://www.delph-in.net/matrix/customize/matrix.cgi; accessed on 2/25/09

[2] Some linguists assume that secondary tense markers are hosted by an auxiliary suffix *-i/(y)* (Lees (1962) and Sezer (2001), among others), though this suffix has also been analyzed as a phonological element (Erguvanlı-Taylan, 1999). Our analysis is compatible with both views.

[3] Throughout the paper, we use the term TAM morphemes to refer to all inflectional morphemes in slots 1-3.

2.2 Verbal Morphology with Lexical Rules

The analysis we propose makes use of the morphotactic infrastructure added to the Matrix customization system by O'Hara (2008), which provides implementations for some wide spread phenomena in morphology. The grammar created this way only requires minor changes for the basic morphology to work.

The morphotactic support allows the definition of multiple morphological "slots" for each stem type. It provides implementations for optional and obligatory morphemes that may add syntactic and semantic features to the derived form. It also allows lexical rules to require earlier, or force later, slots as well as to forbid other slots from appearing. These properties are enforced by binary features on the verb that are related to specific morphological slots. For instance, if optional morpheme2 requires morpheme1 in order to be licensed, bare verbs will carry a feature [MORPHEME2 −]. The lexical rule associated with morpheme1 turns this value into +, which allows the (otherwise prohibited) morpheme2-rule to apply.

The morphotactic infrastructure in the customization system does not provide an analysis for the two agreement paradigms. In this case, we have an obligatory slot of morphemes that interact in a different ways with other morphemes. In order to account for the different agreement paradigms, we created two subtypes of the agreement-lexical-rule, and distinguished them with the binary feature AGR-PARADIGM. The three inflection-rules have two subtypes as well: one supertype of the so-called "true" tenses *-DI* and *-sE*, and one supertype for the other morphemes appearing in the same slot. Rules inheriting from the former supertype turn AGR-PARADIGM to +, whereas rules inheriting from the latter assign it the value −. The value of AGR-PARADIGM controls which agreement rule applies.

The analysis described above ensures that the right morphology is present on independent finite verb forms. In the next section, we present two structures that more or less correspond to VP-coordination in English. In these structures, the morphological requirements on a non-final conjunct differ from those on independent verbs. We propose an analysis along the lines of the basic morphological rules presented above.

3 Coordinated VPs

3.1 Suspended affixation and the "-ip" structure

Turkish has several structures that correspond largely to VP-coordination in English. Namely, simple juxtaposition, the coordination word *ve*, the coordination clitic *de*, and the suffix *-ip*.[4] These structures interact in different ways with the morphology of their conjuncts. We focus on the structures with the suffix *-ip* and the word *ve*, presented in examples (1) and (2).

(1) Çocuk-lar film izle-**yip** pizza
child-PL movie watch-COORD pizza
yi-yor-dı-lar.
eat-CONT-PAST-3PL
"The children were watching a movie and eating pizza."

(2) Çocuk-lar film izli-yor **ve** pizza
child-PL movie watch-CONT and pizza
yi-yor-dı-lar.
eat-CONT-PAST-3PL
"The children were watching a movie and eating pizza."

Both coordination structures share the property that all conjuncts must have the same tense, aspect and mood, even though they may not be overtly marked on non-final conjuncts. The difference lies in the morphological requirements of the first conjunct: the verb marked with *-ip* does not bear any other markers, whereas the progressive marker *-yor* is (obligatorily) repeated in the *ve* structure. In fact, adding or inserting any inflectional marker on a verb bearing *-ip* renders the sentence ungrammatical. In (2), reflecting the phenomenon of "suspended affixation", two of the three suffixes are only marked on the final verb. Additional inflection markers may be present on the preceding conjunct, as long as they are also found on the following conjunct.

Kabak (2007) poses as the primary condition for suspended affixation that the verb must end in a terminal morpheme.[5] Agreement morphemes are terminal, as well as TAM morphemes from the first inflectional slot, except for *-DI* and *-sE*. This difference between the "true tense" morphemes *-DI* and *-sE* on the one hand, and aspect morphemes such as *-Iyor* on the other hand is illustrated in the example below:

(3) Film izle-di-∅ ve pizza ye-di-m
movie watch-PAST-3SG and pizza eat-PAST-1SG
"(S)he watched a movie, and I ate pizza."

(4) Film izle-yor ve pizza yi-yor-um.
movie watch-CONT and pizza eat-CONT-1SG
"I am watching a movie and eating pizza."

[4] Some linguists consider verbs marked by *-ip* "converbs" (Tikkanen, 2001), though in descriptive literature (Lewis, 2000; Kornfilt, 1997) it is generally treated as a coordination marker. Empirical studies have, to our knowledge, not been able to settle the matter up to date. A converb would require a similar treatment when scopal affixes are at play, hence our decision to only present the coordination analysis in this abstract.

[5] A detailed discussion on suspended affixation in *ve* coordination can be found in Kabak (2007). For reasons of space, we limit our presentation to facts necessary to understand the analysis, which is compatible with all observations made by Kabak (2007).

In (3), suspended affixation is not possible. It can only be interpreted as two coordinated sentences. In example (4), on the other hand, both verbs are understood to have the same subject.

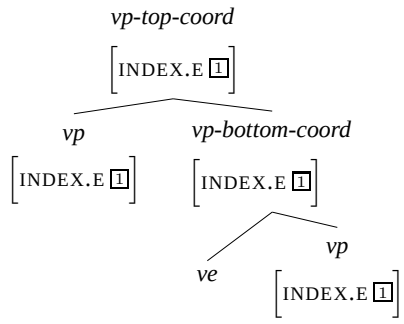
3.2 Analyzing coordinated VPs

The previous section explained that in the *-ip* structure, the verb bearing *-ip* receives the information expressed by inflectional morphemes from the conjunct that follows it. Coordinated VPs within the *ve* structure are also interpreted as if both had the inflection of the final conjunct, even if this may not be overtly present on the preceding verb. On our analysis, the *-ip* suffix takes the same slot as primary inflection morphemes. Evidence for this assumption comes from the fact that all derivational morphemes can precede *-ip*, but *-ip* cannot co-occur with any of the inflectional morphemes.

We assume that the required identity of tense morphemes is a semantic constraint (i.e. coordinated VPs must express events that take place in the same time, with the same mood, aspect, etc.), and implement this constraint via a sharing of semantic features (see below). The requirement that morphemes on a conjunct must be a subset of those on the following conjunct(s) is treated by morpho-syntactic constraints.

Just as for the verbal morphology, the coordination analysis here builds upon the implementation of coordination in the Matrix customization (Drellishak and Bender, 2005). A coordinated structure consists of a bottom-coord-phrase combining the coordination marker with the right element of the coordination, and a top-coord-phrase that adds the left conjunct as specifier, as in (5). Alternatively, when the coordination mark is inflection on the nonfinal conjuncts, the bottom-coord-phrase is a unary rule, and the top-coord-phrase joins appropriately inflected left conjuncts to bottom-coord-phrases, as in (10). In the Matrix definition of basic coordinated verb phrases, the TAM features of the coordinated phrase are identical to those of the right conjunct. Semantically ill-formed structures (i.e. structures in which left and right conjunct have a different TAM interpretation) can easily be excluded by sharing the TAM features of the left conjunct as well. Unification now fails when left and right conjunct provide conflicting semantics.

(5)



This enforces the right semantics, but optional morphemes can still be placed on either conjunct. Additional constraints are necessary to make sure all morphemes present on the first conjunct, are also found on the following one(s). Moreover, we need to make sure that the non-final conjunct ends in a terminal morpheme. For this purpose, we introduce a special lexical rule for nonfinal conjuncts. It takes a verbal form ending in a terminal morpheme as its daughter and creates a word that must be the left daughter of a coordinated structure. The rule is presented in reduced form below. The specification on DTR refers to the lexical rule that adds terminal morphemes to the stem.

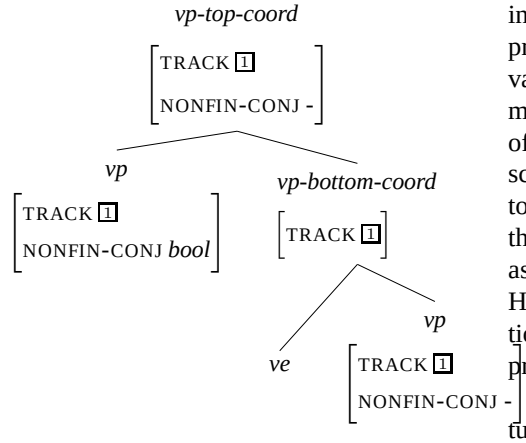
(6)
$$\begin{bmatrix} \text{TRACK} \begin{bmatrix} \text{2ND-INFL } \textit{luk} \\ \text{AGR } \textit{luk} \end{bmatrix} \\ \text{NONFIN-CONJ } + \\ \text{DTR } \textit{term-morph-infl-lex-rule} \end{bmatrix}$$

Here, we use the boolean feature NONFINAL-CONJ to make sure that appropriately inflected constituents appear as the left-hand conjunct in coordinated structures (with *ve*), and not elsewhere. When the nonfinal conjunct rule applies, NONFIN-CONJ is set to +. Subject-head phrases cannot have a head daughter whose NONFINAL-CONJ value is +. Coordinated phrases also require their right hand daughters to be [NONFINAL-CONJ −], but do not constrain this value on the left conjunct. The entire coordinated phrase has value [NONFINAL-CONJ -]. This way, verbs ending in a terminal morpheme that are not fully inflected can only be part of a well-formed structure if they appear as a left conjunct.

What remains is to ensure that the overtly expressed morphology on the left-hand conjunct is a subset of that on the right. To do this, we employ the substructure under the feature TRACK, posited by (O'Hara, 2008) to model dependencies between morphemes. Because we need three situations ("allowed", "prohibited", "applied") for some of the slots, we make use of the type *luk* (subtypes *na* and *bool(ean)*) for their value. When a lexical rule applies, its associated feature receives the value *na*.

The nonfinal conjunct lexical rule changes all values back to *luk*, as shown above. In the coordinated phrase, the TRACK features of both conjunct daughters must unify. If suspended affixation has applied, the *luk* values on the left daughter will unify, regardless of the values found on the right conjunct. Otherwise, the structure is only permitted if the values on both conjuncts are identical. Note that the NONFINAL-CONJ feature cannot be part of the TRACK features, since its value changes in the coordination structure.

(7)



This analysis has one draw-back: in the standard morphotactic analysis provided by the Matrix customization system, TRACK is part of words and lexemes only. In order to use it in phrases, we needed to upgrade it to signs. It seems, however, a property of the data that morphological properties have an impact on well-formedness at a phrasal level. In that sense, the expansion of the features to phrases reflects exactly what makes Turkish coordinated phrases so interesting.

4 Necessity and Ability

4.1 The scope of -(y)Abil and -mEli

Consider the following examples:

(8) Çocuk-lar film izle-yip pizza yi-meli-ler.
child-PL movie watch-COORD pizza eat-NEC-3.PL
"The children must watch a movie and eat pizza."

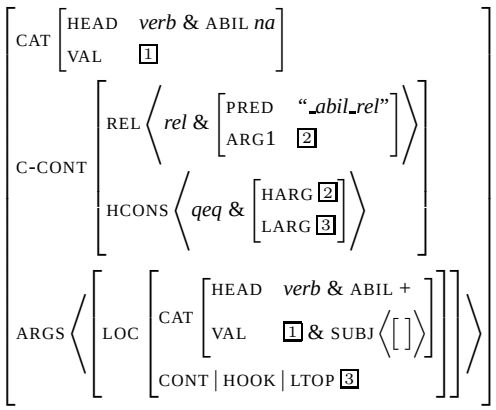
(9) Çocuk-lar film izle-yip pizza
child-PL movie watch-COORD pizza
yi-yebil-ir-ler.
eat-ABIL-AOR-3PL
"The children can watch a movie and eat pizza."

On our analysis, matching of TAM information is handled because these features of verbs are recorded as "variable properties" on their event variables. Following general practice in MRS (Copestake et al., 2005), the event variable of the elementary predication introduced by a verb is also "published" through the verb's INDEX value. Furthermore, this INDEX value is shared with larger constituents that are projections of that verb, and thus the coordination construction has access to the information it needs to ensure matching across conjuncts. However, the derivational affixes *-mEli* and *-(y)Abil* apparently contribute information that is usually handled in terms of (scopal) elementary predications: neccessity and ability. Thus it is more surprising to see this information shared across conjuncts.[6]

4.2 A Constructional Analysis

If we treat *-(y)Abil* and *-mEli* as predicate introducing morphemes, we cannot obtain the correct interpretation of coordinated VPs by simply sharing the value of both events. Nor can we just allow the semantics of these morphemes to attach "low"; Instead of (merely) the second verb, the suffixes must have scope over the entire coordinated VP.[7] This seems to suggest that these affixes attach to phrases rather than words, but "phrasal affixes" would violate the assumption of lexical integrity, generally held in HPSG. Instead, we propose a constructional solution, in the spirit of the analysis that Tseng (2003) proposes for apparent phrasal affixes in French.

Both *-(y)Abil* and *-mEli* contribute a HEAD feature, which is referenced by a special construction that takes a VP daughter and adds the potential or necessity semantics, respectively. The AVM below represents a simplified representation of the unary "ability"-phrase.



The ability predicate is added to the phrase through the C-CONT value.[8] The daughter VP is the first argument of the introduced predicate

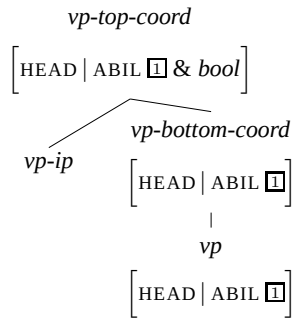
[6] Other derivational morphemes seem not have this property. According to Lewis (2000), the negation morpheme *-mA* also has wide-scope in the *-ip* structure, but none of the speakers we consulted got this reading.

[7] Furthermore, rather than an event variable, the relevant argument position takes a *handle* as its value, which is equal modulo quantifiers to the label of the coordinated VP (cf Copestake et al. (2005)).

[8] Per standard MRS solutions, the CONT value of a phrase is constructed from the CONT values of the daughter(s) and the C-CONT value of the rule licensing the phrase.

and falls under the predicate's scope constraints in HCONS. The construction assigns the value *na* to the HEAD features associated with the *-(y)Abil* inflection. When a phrase has an ability or necessity feature with value *na*, it cannot be the right daughter of an *-ip* coordination. On the other hand, head-subj phrases require that their head daughter's potential or necessity feature is *na* or -. These constraints ensure that the associated constructions can only apply after the coordinated VP has been formed, but must apply before the subject is added to the phrase if *-(y)Abil* or *-mEli* was present on the verb. The tree below represents the ABIL feature in an *ip*-coordination.

(10)



5 Conclusion

This paper presents an analysis for Turkish verbal morphology. The first part of the paper discusses an analysis for basic properties of Turkish inflectional morphemes, which can be implemented with help of the Matrix customization system with only minor changes.

The second part of the paper addresses the morphology of coordinated VPs. We show that both on a syntactic level, as well as on a semantic level, morphological processes partially operate on a phrasal level.

For the syntactic properties at hand (verification whether certain processes took place on each conjunct), passing features that keep track of morphological processes up to the phrase manage to accurately account for the data. But the scopal properties of suffixes *-(y)Abil* and *-mEli* seem to suggest that these affixes attach to phrases rather than words. This property would violate HPSG assumptions on lexical integrity. We show, however, that the data can be analyzed with the help of a construction.

The broader implications of this project are twofold: First, it can be seen as a case study in the cross-linguistic applicability and practical utility of the Grammar Matrix, which has allowed us to quite quickly produce a grammar fragment with which to test these ideas. Second, we have shown that the approach to apparent phrasal affixes of Tseng (2003) is not idiosyncratic to French, but also quite applicable to the unrelated language Turkish.

References

Bender, Emily, Flickinger, Dan and Oepen, Stephan. 2002. The Grammar Matrix: An Open-Source Starter-Kit fir the Rapid Development of Cross-linguistically Consistent Broad-Coverage Precision Grammars. In *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th Conference on Computational Linguistics*, pages 8 – 14, Taipei, Taiwan.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. Stanford, CA: CSLI Publications.

Copestake, Ann, Flickinger, Dan, Pollard, Carl and Sag, Ivan. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3(4), 281–332.

Drellishak, Scott and Bender, Emily. 2005. A Coordination Module for a Crosslinguistic Grammar Resource. In Stefan Müller (ed.), *Proceedings of the HPSG05 Conference*, Stanford: CSLI Publication.

Erguvanlı-Taylan, Eser. 1999. Review: Turkish by Jaklin Kornfilt. *Anthropological Linguistics* 41(2), 253 – 258.

Kabak, Baris. 2007. Turkish suspended affixation. *Linguistics* 45(2), 311–347.

Kornfilt, Jaklin. 1997. *Turkish*. New York: Routledge.

Lees, Robert B. 1962. A compact analysis for the Turkish personal morphemes. In Nicholas Poppe (ed.), *American Studies in Altaic Linguistics*, volume 13 of *Indiana University Uralic and Altaic Series*, The Hague, the Netherlands: Mouton.

Lewis, Geoffrey. 2000. *Turkish Grammar*. New York: Oxford University Press, second edition.

O'Hara, Kelly. 2008. *A Morphosyntactic Infrastructure for a Grammar Customization System*. Masters Thesis, University of Washington.

Sezer, Engin. 2001. Finite inflection in Turkish. In Eser Erguvanlı Taylan (ed.), *The Verb in Turkish*, Chapter 1, pages 1 – 46, Philadelphia: John Benjamins.

Tikkanen, Bertil. 2001. Converbs. In Martin Haspelmath and Wulf Oesterreicher (eds.), *Language Typology and Language Universals: An International Handbook*, volume 2, Chapter 83, Berlin, Germany: W. de Gruyter.

Tseng, Jesse. 2003. Phrasal Affixes and French Morphosyntax. In G. Jäger, P. Monachesi, G. Penn and S. Wintner (eds.), *Proceedings of Formal Grammar 2003*, pages 177–188, Stanford: CSLI Publications.