Inflectional morphology in Turkish VP coordination

Antske Fokkens
Universität des Saarlandes

Laurie Poulson
University of Washington

Emily M. Bender

Proceedings of the 16th International Conference on
Head-Driven Phrase Structure Grammar

Georg-August-Universität Göttingen, Germany

Stefan Müller (Editor)

2009

Stanford, CA: CSLI Publications

pages 110–130

Fokkens, Antske, Laurie Poulson & Emily M. Bender. 2009. Inflectional morphology in Turkish VP coordination. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, Georg-August-Universität Göttingen, Germany*, 110–130. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2009.6.

Abstract

We address three properties of Turkish morphology and VP coordination: the identification of tense and aspect values across conjuncts, the optional omission of affixes on non-final conjuncts coordinated with the word *ve* and the obligatory sharing of scopal modals across conjuncts in coordination structures with the affix *-ip*. For the modals in an *-ip* structure, we propose an analysis that uses syntactic features to trigger the application of a construction at the level of the coordinated VP introducing the scopal predications. Our analysis is implemented in a small HPSG grammar and tested against datasets confirming the functionality and consistency of the analysis.

1 Introduction

This paper presents an analysis of the interaction between verbal morphology and VP coordination in Turkish. There are three properties of Turkish VP coordination of particular interest: the identification of tense, aspect and modality values across the conjuncts, the phenomenon of suspended affixation wherein affixes may be dropped from earlier conjuncts, and a coordination structure that seems to require an analysis in terms of phrasal affixes and thus seems to challenge the notion of lexical integrity. This phrasal affixation is illustrated in example (1), where the meaning of the sentence, with *-mEli* 'must' taking wide scope over the coordination, seems to suggest that *-mEli* is attached to the whole coordinated VP.

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

This paper is also an example of grammar engineering for linguistic hypothesis testing (Bender, 2008), in the sense that we have built a grammar fragment for Turkish that encodes our analyses and verified its behavior over a group of testsuites. These testsuites contain 163 examples, including 96 culled from the literature and an additional 67 we developed and checked with 2-5 native-speaker consultants. The grammar was developed on the basis of the LinGO Grammar Matrix customization system (Bender et al., 2002; Bender and Flickinger, 2005; Drellishak and Bender, 2005),[1] and both the grammar and the testsuites are available for download.[2] Consistent with other Matrix-derived grammars, our grammar

[†]We would like to thank Cagatay Demiralp, Engin Ural and Huseyin Mergan, as well as two additional consultants for their help with the data and their patience. Anonymous reviewers and the audience at the HPSG 2009 conference provided useful comments, which helped to improve this paper. Naturally, all remaining errors are our own. We also thank the IRTG and PIRE for funding a two month stay at the University of Washington. This material is based upon work supported by the National Science Foundation under Grant No. 0644097. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[1]http://www.delph-in.net/matrix/customize/matrix.cgi

[2]http://www.delph-in.net/matrix/turkish

fragment for Turkish produces semantic representations in the format of Minimal Recursion Semantics (MRS, Copestake et al. 2005) and is compatible with the LKB (Copestake, 2002).

This paper tests the following hypotheses:

(i) Obligatory matching of tense and aspect between VP conjuncts can be modeled through structure sharing of features on the event variable.

(ii) The same structure-sharing plus a lexical rule licensing the partially-inflected forms and additional constraints on the coordination rules can account for most suspended affixation facts.

(iii) When scopal affixes (of necessity and ability) are shared among conjuncts, a constructional account along the lines of Tseng 2003 can resolve the apparent violation of lexical integrity.

(iv) The above hypotheses can be implemented in a mutually consistent fashion, which is furthermore consistent with analyses of word order and other phenomena required to parse the sentences in the testsuite.

§2 provides background information on verbal morphology in Turkish and the set of morphological rules we created using the Matrix customization system. §3 describes *ve* coordination, the tense and aspect matching that it requires, and the phenomenon of suspended affixation, along with our analysis of these facts. §4 describes our analysis of another coordination construction, this time marked by an affix *-ip* on the verb of each non-final conjunct. This second construction is of particular interest because it includes apparent phrasal affixes, our analysis of which is given in §4.3. §5 situates our analyses with respect to related work, including Broadwell's (2008) LFG analysis of related facts in Turkish and Tseng's (2003) analysis of apparent phrasal affixes in French.

2 Verbal Morphology in Turkish

2.1 Properties of Turkish Verbs

This section presents an overview of morphemes that may be added to the stem and presents conditions on completeness and well-formedness of the verbs in order to provide background for the analysis of suspended affixation and inflectional marking of coordination in §§3-4. The description is based on, among others, Kornfilt 1997, Lewis 1967, Sezer 2001 and Kabak 2007.

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. According to this definition, Turkish has the following derivational morphemes: *-DIr/t* (causative), *-Il* (passive), *-mA*

Table 1: Inflectional Morpheme Slots

1		2		3		4
-DI	direct past	*-(i)DI*	direct past	*-(i)sE*	conditional	AGR-*k*
-sE	conditional	*-(i)sE*	conditional	*-(i)mIş*	reported past	AGR-*z*
-mIş	reported past/ present perfect	*-mIş*	reported past			
-Iyor	continuous					
-yEcEG	future					
-Ir/-Er	aorist					
-mEli	necessitive					
-mEkte	continuous					

(negation), -(*y*)*A* (abilitative) and -(*y*)*Abil* (abilitative).[3] In addition to the derivational morphemes, there are four slots that may host an inflectional morpheme. The inflectional morphemes are presented in Table 1. A finite verb must bear an inflectional marker from slot 1 and an agreement marker (slot 4). At least one inflectional marker must be phonologically overt (Kabak, 2007).[4,5]

Turkish has two paradigms of agreement markers: the *k*-paradigm which co-occurs with definite past and conditional (*-DI* and *-sE*, respectively) and the *z*-paradigm which co-occurs with all other TAM[6] morphemes. Which paradigm is used depends on the last TAM morpheme attached to the verb. §2.2 describes the morphological analysis that we obtained from the Matrix customization system and how we adapted this analysis in order to accommodate the selection of the different agreement paradigms.

2.2 Verbal Morphology with Lexical Rules

The analysis of basic Turkish morphology 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 with the Matrix customization system only requires minor changes for the basic morphology to work.

[3]We adopt the convention of using capital letters to represent phonemes whose realization depends on vowel or consonant harmony.

[4]Some linguists assume that secondary tense markers are hosted by an auxiliary suffix *-i*/(*y*) (see 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 either view.

[5]We noticed in our data that the plural morpheme does not always follow the order of the slots presented above, though we have not found mention of this in the literature. For present purposes, we assume that this variability in morpheme ordering is a morphophonological property, and we abstract away from it in our implementation; our testsuites regularize examples to follow the canonical order as presented in Table 2.1.

[6]Henceforth, the term TAM morphemes refers to all inflectional morphemes in slots 1-3.

The morphotactic infrastructure allows the grammar engineer to define multiple morphological "slots" for each stem type or set of stem types. 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 preceding slots or to force following 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 and registered under the feature TRACK. TRACK is appropriate for lexical rules and lexical items, but not for phrases.

These binary features work as follows. If, for instance, an 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.

When filling out the Matrix customization questionnaire, we defined nine morphological slots for verbs: five slots for derivational morphemes, three slots for TAM-markers and a slot for agreement markers. In the current version of the grammar, the derivational slots are placeholders, providing only the form of the morphemes and not the associated morphosyntactic or semantic constraints. This is because the Matrix customization system does not currently support the morphosyntax and semantics of causatives or other morphemes that add predicates, nor can it handle negative affixes that are not word-final. These facts could of course be handled by extending the starter grammar. However, because most of the derivational affixes do not have an impact on our analysis, we decided to leave the implementation of these morphemes for future work. The only exception is the derivational morpheme -(*y*)*Abil*: its behavior in -*ip* coordination forms one of the main points of discussion in this paper, and we implemented our analysis of it as an extension to the grammar produced by the customization system. This analysis is discussed in detail in §4.

The morphotactic infrastructure in the customization system does provide most necessary features to implement the inflectional morphology in our verb forms. The library permits the association of features related to tense, aspect and mood as well as subject agreement on verbs. The only phenomenon that is not supported by the current customization system is the interaction of the two agreement paradigms with different inflectional morphemes. In this case, we have morphemes which fill the same obligatory slot but which interact in different ways with preceding morphemes. In order to account for the different agreement paradigms, we created two subtypes of *agreement-lexical-rule*, and distinguished them with the binary feature AGR-PARADIGM, which we added to TRACK. The morphemes in each TAM-slot have two subtypes as well: one for the so-called "true" tenses -*DI* and -*sE*, and one for the other morphemes appearing in the same slot. Rules inheriting from the former type turn AGR-PARADIGM to k, whereas rules inheriting from the latter assign it the value z. 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 what follows, we present two structures that

correspond to VP coordination in English. In these structures, the morphological requirements on a non-final conjunct differ from those on independent verbs.

3 Coordination with *ve*

Turkish has several structures that correspond largely to VP coordination in English. Namely, the suffix *-ip*, the coordination word *ve*, the coordination clitic *de*, and simple juxtaposition (Lewis, 1967). In this paper, we consider the structures with the suffix *-ip* and the word *ve*, as in examples (2) and (3).

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

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

According to the native Turkish speakers consulted, both of these coordination structures share the property that all conjuncts must have the same tense, aspect and mood even though they may be only overtly marked on final conjuncts. The difference between these two structures lies in the morphological requirements on the first conjunct. The verb marked with *-ip* in example (2) may not bear any other markers. On the other hand, the progressive marker *-yor* is obligatorily repeated in the *ve* structure. In example (3), 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. This reflects the phenomenon often referred to as "suspended affixation". In the rest of this section, we provide a more detailed description of VPs coordinated with *ve*, and propose an analysis for suspended affixation. We take up *-ip* coordination in §4.

3.1 Shared TAM Features

As mentioned above, speakers reject expressions where VPs are coordinated that do not have the same tense, as in example (4). If tense and aspect marking is the same, any two VPs can be coordinated using *ve*.[7]

(4) * Çocuk-lar film izli-yor-du ve pizza yi-yecek.
child-PL movie watch-CONT-PAST and pizza eat-FUT
"The children were watching a movie and will eat pizza" (intended)

[7] The data presented in examples (5) and (6) was provided to us by a native speaker, and rated as acceptable by two others. One of the native speakers we consulted, however, did not accept any of these examples, stating that the plural agreement marker is missing on the verb. See §3.2 for more remarks on the subject.

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

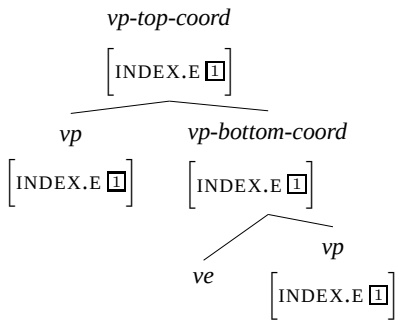
(6) Çocuk-lar film izli-yecek ve pizza yi-yecek.
child-PL movie watch-FUT and pizza eat-FUT
"The children will watch a movie and eat pizza"

We assume that this required identity of TAM morphemes is a semantic constraint (i.e. coordinated VPs must express events taking place in the same time, with the same mood, aspect, etc.), and implement it via a sharing of semantic features.

Just like our analysis of verbal inflection, the coordination analysis here builds upon the implementation of coordination defined through the Matrix customization system (Drellishak and Bender, 2005). Through the customization system, we derived an implementation of polysyndetic coordination, with coordination marker *ve*. This was later manually extended to also include the customization system's implementation of monosyndetic coordination, in order to account for some of the examples found in Kabak 2007.

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 of event features across conjuncts. The Matrix coordination analysis assumes that 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 in (7). 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. With this additional constraint, unification fails when left and right conjunct provide conflicting semantics. The tree in (7) provides a simplified example of a VP *ve* VP coordination.

(7)



3.2 Suspended Affixation

In §3.1, we saw that verbs must bear the same tense and aspect markers in order to form a coordinated VP. However, if we look at (3), repeated as (8) below, it is possible to coordinate the forms *izli-yor* and *yi-yor-lar-dı*, despite the fact that only the last form bears a past tense marker.

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

We see the sharing of tense and aspect information in (8) as well. Here, *izli-yor* is interpreted as if it also bore the past tense and agreement markers visible on the second form. If the past tense marker is only present on the first verb of the VP coordination, the sentence becomes unacceptable, as in example (9):

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

Since Lewis 1967, this phenomenon has been known as "suspended affixation". Suspended affixation also occurs in nominal coordination where case and number marking are shared. Even though only VP coordination is discussed in this work, the proposed analysis easily extends to NP-coordination.

In verbal structures, suspended affixation does not allow arbitrary strings to be omitted. Rather, as argued in Kabak 2007, a form exhibiting suspension of affixes is acceptable only if it constitutes a morphological word, i.e., a word able to stand in isolation. According to Kabak, morphological words end in "terminal morphemes"; agreement morphemes and aspect and modality morphemes are "terminal".[8] These terminal aspect and modality morphemes are all of the slot 1 morphemes in Table 2.1 except *-DI* and *-sE*.

For instance, in example (10), suspended affixation is not possible. It can only be interpreted as two coordinated sentences. Interpreting the first verb with no agreement marking, i.e. without a null 3SG morpheme, is not possible as the verb must end in a terminal morpheme and so cannot end in *-DI*. In contrast, in example (11), the first verb is interpreted as undergoing suspended affixation since *-yor* is a terminal morpheme. Therefore both verbs are understood to have the same subject.

(10) 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."

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

[8] The affix *-ip*, discussed in §4 also functions as a terminal morpheme.

Speakers have a strong preference for coordinated VPs over coordinated sentences with pro-drop. Example (10) was judged "not nice" and one of our speakers even rated it "ungrammatical". This preference may explain why none of the speakers consulted could interpret *izli-yor* as a fully inflected form of third person singular in example (11).

To our knowledge, the work presented by Kabak (2007) provides the most detailed and precise description of suspended affixation available. In the data we collected from native speakers, however, another issue emerged that was not evident in Kabak's data. Three of our four native speakers accepted the example in (12).

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

(12) is an apparent counter-example to Kabaks's generalization about the forms that can appear with suspended affixation, as it ends with *-DI*. However, these speakers appear to treat the -∅ marker as unmarked for number, even in non-coordinated contexts, like (13).[9] Thus Kabak's generalization can be maintained.

(13) Çocuk-lar film izle-r-∅
Child-PL movie watch-AOR-3
"The children watch a movie"

One puzzle remains, however, and is illustrated in (14). The speakers we consulted interpreted this example as having two distinct subjects, but if the -∅ third-person marker is underspecified for number, a same-subject reading should be available.

(14) Çocuk-lar film izli-yor-lar-dı ve pizza yi-yor-dı-∅
Child-PL movie watch-CONT-3PL-PAST and pizza eat-CONT-PAST-3
"The children were watching a movie and he was eating pizza."

Perhaps it is possible to account for this with an appeal to pragmatics, where the marking on the first conjunct is taken as contrastive. Alternatively, a syntactic account in terms of including a feature [LAR *luk*] registering presence of overt plural markers could account for this data. The coordination construction can then exclude structures where the left-hand daughter is [LAR +] and the right-hand daughter [LAR −]. This analysis works in similar ways as that of multiple suspended affixation explained in §3.4, but is relatively inelegant. We leave the resolution of this issue to future work.

3.3 Analysis of Suspended Affixation

The analysis of coordination presented in §3.1 does not accommodate suspended affixation, since only verbs bearing agreement markers are considered words. In

[9] The one speaker who rejected (12) also rejected (13).

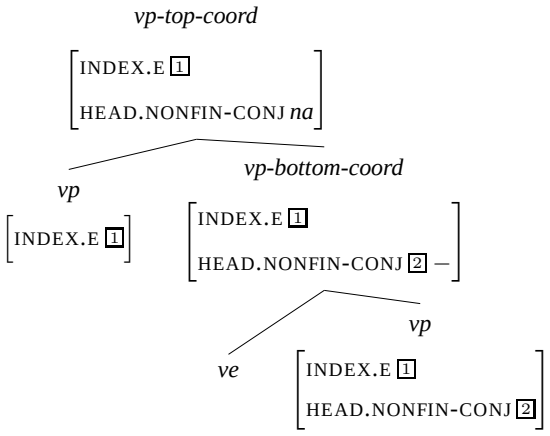
order to account for examples such as (11), we introduce a lexical rule, called the *non-final-conjunct-rule*, that changes verbs bearing a morpheme from slot 1 into words, without adding any further inflection. It takes a verbal form ending in a terminal TAM morpheme as its daughter and creates a word that must be the left daughter of a coordinated structure. The rule sketched in (15) below.

(15) $\begin{bmatrix} \text{SYNSEM.LOCAL.CAT.HEAD.NONFIN-CONJ} + \\ \text{DTR } \textit{term-morph-infl-lex-rule} \end{bmatrix}$

The constraint on the DTR value ensures that this rule may only take as input forms ending with a slot 1 morpheme other than *-DI* and *-sE*; the type *term-morph-infl-lex-rule* is a supertype to all lexical rules that introduce such slot 1 morphemes. When the rule in (15) applies, it creates a word which is underspecified for TENSE and ASPECT, making it compatible with values for these features "unified in" from the right hand conjunct in a coordination structure. The other rules that take slot 1 morphemes as input are the ordinary rules for slots 2 and 3. When these rules apply, the resulting form is not restricted to be a left conjunct and it is given specific values for TENSE and/or ASPECT. In this way, we capture Kabak's generalization that there are two paths for a lexeme to become a well-formed morphological word, through the *non-final-conjunct-rule* or through the slot 2 and 3 rules.

As shown in (15), we posit a head feature NONFIN-CONJ, which takes values of type *luk*.[10] *Luk* is a supertype of *boolean* and *na* (*not-applicable*). The *na* value allows us to distinguish coordinated structures from non-coordinated structures, and facilitates the analysis of suspended affixation in multiple coordination (§3.4). The subtypes of *boolean* are used to distinguish verbs that are marked as non-final conjuncts ([NONFIN-CONJ +]) and exclude them from the head daughter position of subject-head phrases ([NONFIN-CONJ *na-or-*−]) and the right conjunct of coordinated structures ([NONFIN-CONJ −]). The value of left conjuncts in these structures is unrestricted, since suspending affixes is an optional process. (16) illustrates this analysis of binary VP coordination with *ve*.

(16)



[10] In using the type *luk*, we follow the English Resource Grammar (Flickinger, 2000).

3.4 Suspended Affixation with Multiple Conjuncts

In addition to the data presented in Kabak (2007), we looked at structures with more than two conjuncts. In this case, suspended affixation can apply as long as the verb is not preceded by a fully inflected verbal form that is part of the same VP coordination. In fact, speakers prefer expressions where suspended affixation has applied to all but the last verb. The examples below illustrate cases of well- and ill-formed structures with multiple conjuncts.

(17) Çocuk-lar kitap oku-yor-lar-di, film izli-yor-lar-di ve
child-PL book read-CONT-3PL-PAST movie watch-CONT-3PL-PAST and
pizza yi-yor-lar-di.
pizza eat-CONT-3PL-PAST
"The children were reading a book and watching a movie and eating pizza"

(18) ? Çocuk-lar kitap oku-yor, film izli-yor-lar-di ve pizza
child-PL book read-CONT movie watch-CONT-3-PL-PAST and pizza
yi-yor-lar-di.
eat-CONT-3PL-PAST
"The children were reading a book and watching a movie and eating pizza"

(19) Çocuk-lar kitap oku-yor, film izli-yor ve pizza
child-PL book read-CONT movie watch-CONT and pizza
yi-yor-lar-di.
eat-CONT-3PL-PAST
"The children were reading a book and watching a movie and eating pizza"

(20) * Çocuk-lar kitap oku-yor-lar-di, film izli-yor ve pizza
child-PL book read-CONT-3PL-PAST movie watch-CONT and pizza
yi-yor-lar-di.
eat-CONT-3PL-PAST
"The children were reading a book and watching a movie and eating pizza"

The data above suggest the following generalizations:[11]

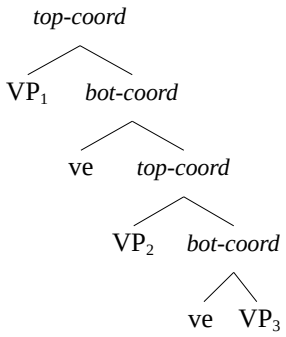
(i) The final VP of a coordinated structure must be fully inflected.

(ii) Fully inflected VPs may not precede VPs that exhibit suspended affixation within a coordinated structure.

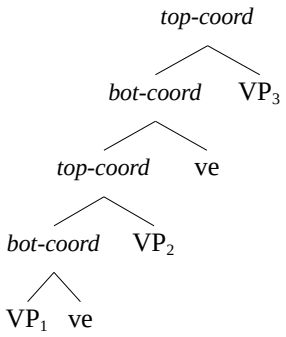
Coordination structures provided by the Matrix customization system are right-branching. This is problematic for generalization (ii) above. Consider the right-branched structure in example (21). The data shows that if affixes on VP_2 are suspended, VP_1 may not be fully inflected, but we cannot pass the value of NONFIN-CONJ from VP_2 to the coordinated VP above it, because outside of multiple coordination, that coordinated VP behaves as if it is [NONFIN-CONJ *na*].

[11] Speakers have different intuitions on this data. Some only accept (19) and (17). Others say that none of the examples is "completely ungrammatical". All speakers agree, however, that the order of acceptability is clear: (19)> (17) > (18) > (20)

(21)



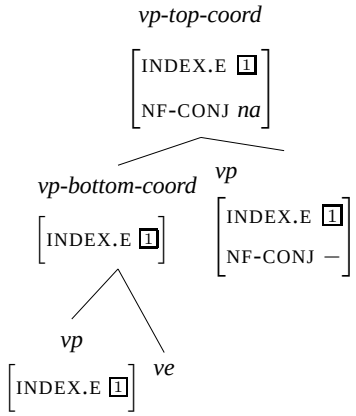
(22)



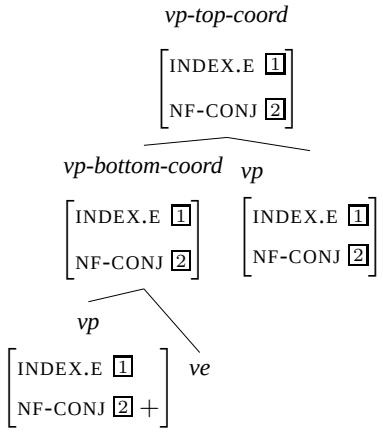
A more natural approach may be to assume that the morphology of VP_1 may pose restrictions on following conjuncts. Compare the structure in example (22) to the one represented in example (21). In (22), the right daughter of a well-formed embedded VP can determine restrictions on the rest of the structure. This allows us to impose the restriction that a VP that has suspended affixes may only serve as the right conjunct if the left conjunct has suspended affixes as well. The resulting VP coordination of two such VPs bears the value [NONFIN-CONJ +] and must occur as the left daughter of a coordinated VP itself. A fully inflected VP, on the other hand, may always be the right conjunct in a coordinated VP. Because the resulting coordination is [NONFIN-CONJ *na*], it can never become left conjunct when the right conjunct exhibits suspended affixation.

The analysis we assume requires two coordination constructions: one for left conjuncts that exhibit suspended affixation, and one for left conjuncts that do not. The trees in (23) and (24) represent the two constructions.

(23)



(24)



Changing VP coordination to a left-branching structure seems natural for Turkish, since it is a language that generally prefers left-branching structures. It also provides further insight in typological properties of coordination structures. Drellishak and Bender (2005) assume that a cross-linguistic analysis of coordination could make do with right-branching structures only, and suggest that the only struc-

tures a right-branching approach would exclude are unattested examples such as "*conj* A B C" (*Ibid.*, p.18). Multiple *ve*-coordination reveals an unforeseen case where left-branching seems required. This is because of the double role suspended affixation and complete inflection play in the well-formedness conditions of the complete coordination. On the one hand, the presence of full inflection on the final conjunct is a well-formedness condition that must be encoded on the final structure, so that the coordinated VP can be combined with other elements in the sentence. On the other hand, this same property poses restrictions internal to the coordinated VP, which requires this information to be shared among the (non-final) conjuncts. In a right-branching structure, the final conjunct is the most embedded phrase within the coordination. Relevant information must thus be passed up through the entire coordination construction in order to appear on the resulting coordinated VP. This makes it impossible to share information between phrases that are added to the coordination structure later on, if they appeal to the same feature. When using left-branching coordination, on the other hand, this problem is avoided: relevant information can be passed up directly from the VP that was added to the structure last, allowing the final conjunct to provide relevant information concerning the entire VP. At the same time, restrictions that are internal to VP coordination can be handled by the interaction between *vp-top-coord* and *vp-bottom-coord*.

3.5 Summary

This section has presented an analysis of *ve* coordination and suspended affixation. The analysis accounts for the matching of tense, aspect and modality features across the conjuncts in *ve* coordination structures as well as the potential for affixes to be "dropped" from left-hand conjuncts. In addition, our analysis extends to coordination of more than two conjuncts with *ve* and captures the facts about the distribution of suspended affixation in these constructions.

4 The *-ip* Structure

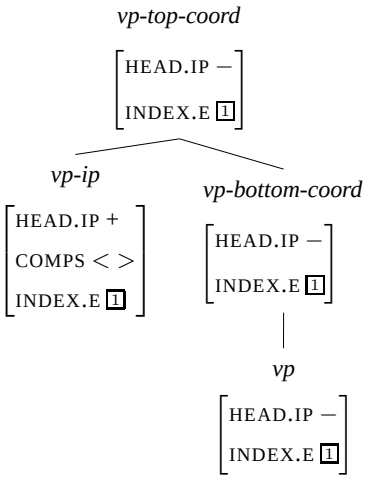
In this section, we discuss the other coordination structure of interest: coordination marked with the suffix *-ip*. As with *ve* coordination, the semantics associated with the inflection marked on the final conjunct are shared with any other conjuncts. In contrast to the *ve* structure, *-ip* is a suffix on the verb in the left conjunct and it cannot co-occur with any inflectional morphology. This section provides a description of our analysis of the *-ip* structure as a coordination relation. In addition, we provide a brief discussion of the consequences for this approach of an alternative analysis of *-ip* as a "converb" marker.

4.1 Affixal VP-Coordination

In order to implement the *-ip* coordination relation, we follow Drellishak and Bender's (2005) analysis for the Trans-New Guinea language Ono wherein a feature registers the presence of marking that is relevant for VP coordination. This feature allows the VP to become part of a coordinated structure through a unary rule. The Turkish *-ip* suffix is, similarly, a VP coordination marker attached to a word. While this option is not directly provided by the Matrix customization system, the relevant constraints can be added to an analysis provided by the customization system in a straight-forward way. In Ono, the marked form was the right conjunct. In our case, where the left conjunct is marked, we needed to change coordination into a left-branching structure, as was done for *ve*-coordination.

The suffix *-ip* cannot occur with any other inflectional morphemes but can be added directly to the stem or to derivational morphemes. According to our analysis, it is therefore added to the verb at the first slot for inflectional suffixes, creating a word. This lexical rule changes the value of a feature IP to $+$. In all other cases, this feature will have the value $-$. The coordination structure that creates *-ip* coordination only takes left daughters that are VPs and marked [IP $+$], the resulting structure is [IP $-$] again, as illustrated in (25). Note that values related to tense, aspect and mood are shared among the conjuncts, just as for *ve* coordination.

(25)



4.2 Converb Marker

Some linguists consider verbs marked by *-ip* "converbs" (Tikkanen, 2001), though in descriptive literature (Lewis, 1967) it is generally treated as a coordination marker. Empirical studies have, to our knowledge, not yet settled this matter; the definition of "converb" is not clear-cut and the importance of the "modifying" character of converbs is debated. Johanson (1995) argues that there are both modifying and non-modifying converbs in Turkish, where non-modifying converbs are distinguished semantically in that they depict "events of equal narrative status" (*Ibid.*,

p.322). The difference between these and coordinated clauses relates to the information structure of the clause; converbs may express information that is in focus.

Another difference between modifying and non-modifying converbs in Turkish lies in their interaction with the scope of the main verb. Whereas modifying converbs fall outside of the scope of tense, aspect and modal markers of the main clause, these do have scope over non-modifying converbs that precede them, as illustrated in (26) from Johanson 1995, p.323.

(26) Herkes çik-ip 'Ben Türk-üm di-yebil-meli
everybody come.out-CONV I Turk-COP.1.SG say-POSSIB-NEC.3SG
"Everybody should be allowed to step forth and [should be allowed to] say that he is a Turk."

According to Johanson's definition, the *-ip* structures discussed in this paper should be considered non-modifying converbs. This would mean that their interpretation would be that of events with narrative status equal to that of main verbs. This is exactly what the coordination analysis above provides. The only difference between a coordinated structure and a non-modifying converb structure is the subordinate character of the latter. However, because this is not represented in the final semantic interpretation of the sentence there does not seem to be a reason to propose an analysis that is radically different from the one that is proposed above, except for perhaps changing the names of the phrases used to *converb* rather than *coordinated*. One could also extend the analysis to incorporate the correct information structure, though this is beyond the scope of the present study.

In sum, whether one considers the *-ip* structure as a converbial structure or as a coordinate structure depends on the criteria that are used to distinguish the two. We take the final semantic representation, which is compatible with the coordination account, as the primary consideration and use it as the basis for our analysis.

4.3 Shared Scopal Morphemes

Whereas the question of whether *-ip* marks converbs or coordinated structures is, in our opinion, not of crucial importance, one observation mentioned by Johanson (1995) is particularly relevant here: Verbs bearing *-ip* fall under the scope of the verb they precede. Though we are not aware of accounts that discuss this matter in detail, this property is mentioned by several authors of Turkish grammars. Our data does confirms this observation concerning the wide scope of the suffix *-mEli*. In addition, we found that the suffix *-(y)Abil* has scope over the entire coordinated structure when it appears only on the right conjunct. Consider (27) and (28):

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

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

The analysis of *-ip* structures described above handles part of the shared interpretation between the verbs: information regarding tense, aspect and mood are stored as features that are part of the verb's event variable, which is identified across conjuncts. However, *-mEli* and *-(y)Abil* contribute information that is usually handled in terms of (scopal) elementary predications: necessity and ability, respectively. Thus, it is more surprising to see this information shared across conjuncts.[12] In §4.4, we demonstrate that a constructional analysis can provide the right semantics for *-ip* structures in which these scopal morphemes occur.

4.4 A Constructional Analysis

If we assume that *-(y)Abil* and *-mEli* are scopal and treat them as predicate introducing morphemes, we cannot obtain the correct interpretation of coordinated VPs by simply sharing the value of TAM features across both events. Nor can we just allow the semantics of these morphemes to attach "low"; instead of merely the second verb, the suffixes too must have scope over the entire coordinated VP. This seems to suggest that these affixes attach to phrases rather than words, but "phrasal affixes" would violate the assumption of lexical integrity, which is 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, each of which is referenced by a special construction that takes a VP daughter and adds the appropriate semantics. The AVM in (29) below provides a simplified representation of the unary *ability-phrase-rule*.

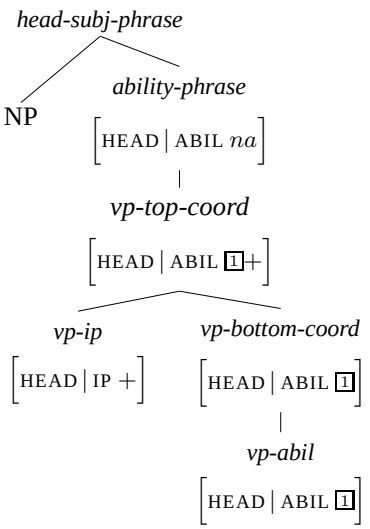
(29)
$$\left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \left[\begin{array}{ll}\textit{verb} & \\ \textsc{abil} & \textit{na}\end{array}\right] \\ \textsc{val} & \boxed{1}\end{array}\right] \\
\textsc{c-cont} & \left[\begin{array}{ll}\textsc{rels} & \left\langle\left[\begin{array}{ll}\textit{rel} & \\ \textsc{pred} & \textit{"_abil_rel"} \\ \textsc{arg1} & \boxed{2}\end{array}\right]\right\rangle \\ \textsc{hcons} & \left\langle\left[\begin{array}{ll}\textit{qeq} & \\ \textsc{harg} & \boxed{2} \\ \textsc{larg} & \boxed{3}\end{array}\right]\right\rangle\end{array}\right] \\
\textsc{args} & \left\langle\left[\textsc{loc}\left[\begin{array}{ll}\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \left[\begin{array}{ll}\textit{verb} & \\ \textsc{abil} & +\end{array}\right] \\ \textsc{val} & \boxed{1}\left[\textsc{subj}\left\langle\,[\,]\,\right\rangle\right]\end{array}\right] \\ \textsc{cont}\,|\,\textsc{hook}\,|\,\textsc{ltop} & \boxed{3}\end{array}\right]\right]\right\rangle
\end{array}\right]$$

This non-branching construction licenses a VP node over any VP with the feature declaration [ABIL +], and its purpose is to insert the *_abil_rel* predication into

[12] Other derivational morphemes seem not have this property. According to Lewis (1967), the negation morpheme *-mA* also has wide-scope in the *-ip* structure, but none of the speakers we consulted got this reading.

the semantics. This predicate is specified in the C-CONT (construction content) feature of the construction, following standard MRS practice for semantically contentful constructions. It further specifies that the local top handle of the daughter VP is the argument of the introduced predicate.[13] In order to ensure that this construction only applies outside (and not within) VPs coordinated with *-ip*, the mother is marked [ABIL *na*] and the rule licensing the right-hand daughter of an *-ip* structure requires [ABIL *bool*]. A similar construction is posited for *-mEli*, with an associated feature NECESS, subject to analogous constraints. We ensure that the relevant construction fires if the morphology is present by requiring the value *na* for both of these features in the *head-subj-phrase*. The tree below illustrates the workings of the ABIL feature in an *-ip* coordination.

(30)



4.5 Summary

This section has presented an analysis of *-ip* coordination. Our analysis handles the following facts: In *-ip* coordination, non-final conjuncts must be marked with *-ip*, which is incompatible with any other inflectional morphology. Information expressed by inflectional morphemes on the final conjunct (including tense and aspect information) is interpreted as shared with all conjuncts. Our analysis handles this sharing through the same identification of TAM features as in *ve* coordination. In addition, when the final conjunct bears the affixes *-(y)Abil* or *-mEli*, these are interpreted as taking wide scope over the whole coordinated VP. We assume these affixes correspond to scopal elementary predications in the semantics and we propose an analysis where the affixes contribute only syntactic features, which then trigger the application of a construction at the level of the coordinated VP introducing the scopal predications into the semantics.

[13] This argument relation is mediated by the "equal modulo quantifiers" (*qeq*) handle constraint, to allow quantifiers to scope in between, while maintaining the scopal relationship where the *_abil_rel* outscopes the verb's (or verbs') predication(s).

5 Related Work

In this section, we situate our analysis with respect to related work. First, in §5.1, we contrast our analysis to the LFG account of Broadwell (2008). Then, in §5.2, we describe how our account of these Turkish facts is broadly similar to Tseng's (2003) account of a very different phenomenon in French.

5.1 Suspended affixation in LFG

To our knowledge, the only other formal account of suspended affixation in Turkish is the LFG account of Broadwell (2008). Broadwell applies Westcoat's (2002) notion of "relaxed lexical integrity" which allows a single (morphological) word to represent two adjacent c-structure nodes, even if the c-structure nodes do not form a constituent. On this analysis, the affixes that are shared between two or more conjuncts represent independent c-structure nodes attaching to the entire coordination. They are associated with special "instantiation" rules which allow them to be co-instantiated with the final word of the nearest conjunct.

Broadwell considers an analysis similar to ours as an alternative to the "co-instantiation" approach. On this alternative analysis, the affixes are part of the final conjunct, which bears special functional equations propagating its values for the features expressed in the affixes to the coordinate structure as a whole. Broadwell argues against this analysis on the basis that it requires the stipulation that the special annotations appear on the rightmost conjunct. On the "co-instantiaton" analysis, the location of the affixes within the coordinate structure can be seen to follow from the general head-final property of Turkish.

However, we argue that lexical integrity is not something to give up lightly. Furthermore, our analysis allows us to capture the similarity between required matching of tense and aspect morphology when it is overt and required matching of tense and aspect values when the morphology is not present on a non-final conjunct. In addition, we note that Broadwell proposes a flat (symmetrical) structure for coordination, whereas we follow a binary-branching analysis. On a binary branching analysis, it is less surprising that one conjunct should have special properties. Finally, Broadwell notes that his syntactic analysis cannot capture Kabak's morphological generalization about which affixes can be suspended, and appeals instead to an external morphological filter. We conclude that our account seems preferable in that it allows us to handle the data in more detail while simultaneously preserving lexical integrity.

5.2 Phrasal affixes in French

Tseng (2003) posits a very similar solution to ours for what appears to be a very different problem. In particular, he is addressing the apparent contradiction between the phonological and syntactic status of the formatives *le*, *de* and *à* in French. These elements (one determiner and two prepositions) are functors, which we would ex-

pect to combine respectively with an N′ or an NP, but phonologically (and, Tseng argues, morphologically) they combine with the first word in that N′ or NP. Since this first word need not be the head, the syntactic and semantic information that the functors require is not available to them locally.

Tseng's solution is to more-or-less freely attach *le, de* and *à* as prefixes. The morphophonological information associated with these morphological rules handles contextual variation in the form of the prefix, while the morphosyntactic effect of the rule is to encode information about the affix in an EDGE feature. The EDGE feature is propagated up the periphery of the constituent and is finally interpreted by a unary rule which builds an NP out of an N′ or a PP out of an NP, according to the information stored in EDGE.

The analysis proposed in this paper of *-(y)Abil* and *-mEli* in *-ip* coordination differs from Tseng (2003) in that it does not refer to the feature EDGE and the values of our phrasal affix features are less complex: they merely register presence of particular morphemes. It would be possible to adapt this analysis and make it more similar to Tseng's account of phrasal affixes in French, with the additional advantage that we would only use one feature (EDGE) rather than two (ABIL) and (NECESS). Fundamentally, however, our analysis is exactly parallel to Tseng's, in positing a pair of rules, one morphological and one syntactic, in order to handle apparent phrasal affixes without sacrificing lexical integrity. The fact that the same analytical device can handle such superficially different phenomena speaks to its generality while also raising interesting questions about the typology of phrasal affixes. When are such paired rules required, and why are they not more common?

6 Conclusion

This paper presented three phenomena related to the morpho-syntax of Turkish VP coordination. First, our data showed that tense, aspect and modality marking on coordinated VPs must be identical. We proposed an analysis that models this by sharing the value of event semantics on both VPs.

The second phenomenon we discussed is that of suspended affixation. This paper introduced new observations related to plural markers and coordination with multiple coordinands. §3.2 and §3.4 presented analyses for binary and multiple coordination, respectively. The latter showed that the restrictions on multiple coordinands require left branching coordination structures, contra the claim in Drellishak and Bender 2005.

The only other formal analysis of suspended affixation that we are aware of is described in Broadwell (2008). In §5.1, we discussed this alternative account and argue that our proposal is superior because (i) it can account for the morphosyntactic properties of the phenomenon as described by Kabak and (ii) it respects lexical integrity.

Finally, we discussed coordinated structures that make use of the suffix *-ip*. The alternative view that *-ip* is a converb marker was discussed, and it was argued

that treating these verbs as converbs does not simplify the analysis, nor lead to a more accurate semantic representation of the sentence. We presented data that shows that markers on the verb that follows the VP marked with *-ip* scope over both VPs and therefore seem to attach to a phrase rather than a word. This would violate HPSG assumptions on lexical integrity. We show, however, that the data can be analyzed with the help of a construction.

All the analyses presented in this paper have been implemented in a small grammar fragment. In addition to presenting the phenomena and their analyses, we also indicated how the analyses were implemented with help of the Matrix customization system. This had two main benefits: First, it allowed us to test both the accuracy of our analyses and whether they could be implemented in a mutually consistent fashion. Second, it allowed us to test the cross-linguistic applicability and utility of the Grammar Matrix. On the one hand, the Grammar Matrix customization system supported the creation of this paper: it allowed us to quite quickly produce a grammar testing our hypotheses, which confirmed its applicability and utility. On the other hand, our implementations pointed to a typological fact that had not been foreseen in building the coordination library of the Matrix: namely that morphological properties may require left-branching coordination structures.

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.

Bender, Emily M. 2008. Grammar Engineering for Linguistic Hypothesis Testing. In Nicholas Gaylord, Alexis Palmer and Elias Ponvert (eds.), *Proceedings of the Texas Linguistics Society X Conference: Computational Linguistics for Less-Studied Languages*, pages 16–36, Stanford: CSLI Publications.

Bender, Emily M. and Flickinger, Dan. 2005. Rapid Prototyping of Scalable Grammars: Towards Modularity in Extensions to a Language-Independent Core. In *Proceedings of the 2nd International Joint Conference on Natural Language Processing IJCNLP-05 (Posters/Demos)*, Jeju Island, Korea.

Broadwell, George Aaron. 2008. Turkish Suspended Affixation Is Lexical Sharing. In *Proceedings of the LFG08 Conference*.

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 Publications.

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

Flickinger, Dan. 2000. On Building a More Efficient Grammar by Exploiting Types. *Natural Language Engineering* 6 (1) (Special Issue on Efficient Processing with HPSG), 15 – 28.

Johanson, Lars. 1995. On Turkish Converb Clauses. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in Cross-linguistic Perspective*, pages 313 – 349, Berlin: Mouton de Gruyter.

Kabak, Bariş. 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. 1967. *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.

Westcoat, Michael T. 2002. *On Lexical Sharing*. Ph. D.thesis, Stanford University.