

WORD LENGTH

Abstract

In this paper it is argued that the process of affixation is subject to a morphological constraint which determines the size of a word. Languages vary with respect to the affix space they have, that is, the number of affixes in a word depends on the number of slots available for affixes in that language. This hypothesis is based on certain morphological and syntactic differences between two Turkic languages, Turkish and Yakut. The morphological form of main and embedded verbs, the presence and behaviour of buffer stems, the interaction of affixation and auxiliaries, the cross-linguistic asymmetries in the presence of complementisers and double nominal constructions (possessive phrases and relative clauses) indicate that Turkish has a maximum number of three slots for inflectional affixes whereas Yakut has two. According to this proposal morphological and syntactic processes relating to inflectional morphology are not coextensive, and morphology is a separate component which contains restrictions on word length.

Introduction

One of the central issues regarding morphology is its status within the structure of grammar. This issue becomes particularly relevant when approaching inflectional morphology. In a number of studies, inflection is characterised purely in syntactic terms, by means of representing inflectional affixes in syntax (Pollock 1989, Chomsky 1993). The part morphology plays is reduced to lexical features and their role in deriving syntactic representations. The implicit assumption is that there is not much to be said about inflectional morphemes other than their syntactic manifestation. An alternative route of research highlights the differences between word structure and phrase structure (Di Sciullo and Williams 1987, Anderson 1992, Bresnan and Mchombo 1995, Sells 1995, Di Sciullo 1996, Bach 1996, among others). One common factor in these studies is that, although there may be overlaps, morphological structure is not an extension of syntactic structure, and vice versa.

The proposal that I put forward in this paper supports the view that morphology is a separate component. Its architecture contains, among other things, a condition on word length characterised in terms of morphological slots. In particular, I propose that the number of slots available determines the upper limit of affixation. Based on the differences between verbal forms in main clauses and embedded clauses in Turkic languages, it will be argued that some of the differences in the inflectional paradigm of these languages can only be explained in terms of a cross-linguistic variation in morphological space. In addition, this hypothesis explains certain variations in the clause structure of Turkic languages, it accounts for the defective nature of the tense/aspect paradigm in relative clauses and sheds light on the placement of agreement

morphology, the occurrence of complementisers, the availability of multiple nominal constructions and the presence of buffer stems. It also explains why certain forms which are syntactically allowed do not exist.

Parts 1 and 2 focus on main verbs and embedded verbs in Turkish and Yakut, a Turkic language spoken in north-eastern Siberia. A comparison of the two languages leads to the proposal concerning morphological templates in the third section. The fourth section gives further evidence for morphological templates. Section 5 provides a discussion of morphological slots, followed by the implications of the present hypothesis for syntax in the final section.

1. Turkish

1.1 Main verbs

Inflected verbs in Turkish are minimally composed of a stem followed by a tense suffix and by a subject agreement marker:

- (1) Ben balıĝ-ı ye-di-m.
 I fish-ACC eat-PAST-1
 I ate the fish.

Other suffixes which may occur in the place occupied by the past tense marker are the following:

- (2) Ben balıĝ-ı yi-(y)eceĝ-im FUTURE: I will eat the fish.
 Ben balıĝ-ı ye-r-im. HABITUAL: I eat fish.
 Ben balıĝ-ı yi-yor-um. PROGRESSIVE: I am eating fish.
 Ben balıĝ-ı ye-se-m CONDITIONAL: If I eat the fish
 Ben balıĝ-ı ye-miř-im. HEARSAY: I must have eaten the fish.

An inflected verb may have two temporal/aspectual suffixes:

- (3) a. Ye-r-di-m. b. Yi-yor-muř-um. c. Yi-(y)ecek-se-m...
 eat-AOR-PAST-AGR eat-PROG-HS-AGR eat-FUT-COND-AGR
 I used to eat. I appear to have been eating. If I am going to eat...

The differences between temporal and aspectual suffixes do not have a bearing on the present analysis and will not be discussed here. For practical purposes, I will refer to all tense and aspect suffixes as T/A, and the part of the verb containing T/A suffixes and agreement as the inflectional domain. The inflectional domain follows grammatical function changing suffixes, modals and negation. The data above indicates that the morphological structure of the main verb with respect to inflection is:

- (4) V-T/A-(T/A)-AGR

1.2 Embedded verbs

Verbs in complement clauses and object relative clauses are also inflected for T/A and agreement¹:

¹Other types of embedded clause such as the subject relative clause which is not inflected for T/A or person will not be discussed here.

- (5) a. *Complement clause*
 [Ben-im balıġ-ı ye-diġ-im] ortada.
 I-GEN fish-ACC eat-PAST-C-AGR is obvious
 It is obvious that I eat/ate the fish.
- b. *Object relative clause*
 [Ben-im ye-diġ-im balık] barbunya.
 I-GEN eat-PAST-C-AGR fish red mullet
 The fish that I eat/ate is red mullet.

-diġ often denotes non-future events. It is morphologically replaceable only by one other form, *-(y)eceġ*, which denotes future events:

- (6) a. *Complement clause*
 [Ben-im balıġ-ı yi-yeceġ-im] biliniyor.
 I-GEN fish-ACC eat-FUT-C-AGR is known
 It is known that I will eat the fish.
- b. *Object relative clause*
 [Ben-im yi-yeceġ-im balık] barbunya.
 I-GEN eat-FUT-C-AGR fish red mullet
 The fish that I will eat is red mullet.

In grammars of Turkish both *-diġ* and *-(y)eceġ* are treated as single morphemes, as they are in the generative literature (Özsoy 1988, Kennelly 1992), with the exception of Kural (1993) who analyses *-ġ* in *-(y)eceġ* as a separate morpheme. I suggested elsewhere (Göksel 1997) that this analysis was preferable for morphological reasons and extended it to include *-diġ*, hence *-di+ġ*, and represented *-ġ* as the complementiser in Turkish. The motivation for this analysis is based on the fact that a syntactically discrete structure (subordination) is marked phonologically (*-ġ* being vowel length), shown below:

(7)	Main clause	Subordinate clause
non-future	/yedim/ (I ate)	/yedi:m/ (that I ate)
future	/yiyçæm/ (I will eat)	/yiyçæ:m/ (that I will eat) ²

There are a number of differences between embedded verbs and matrix verbs. Embedded verbs are only marked for future versus non-future specification, and they lack the array of T/A suffixes which occur in matrix verbs. Embedded verbs cannot occur with any of the markers in (2) (other than *-(y)eceġ*). Then, they cannot have more than one T/A which is possible in matrix verbs:

²See Göksel 1997 for a discussion of the dialectal variations of this form.

The morphological structure of the embedded clause is:

(16) V-T N-AGR

3. A comparison of Yakut and Turkish morphology

Yakut and Turkish are similar in terms of the unavailability of some T/A affixes in embedded clauses. However, there are non-trivial differences between the two. The inflectional domain of the main verb in Turkish contains three suffixes as opposed to two in Yakut. The number of suffixes is again three in Turkish embedded verbs whereas in Yakut embedded constructions there is one suffix each on the verb and on the head of the nominal construction, summarised below:

(17)	<i>Turkish</i>	<i>Yakut</i>
Main verbs	V-T/A- (T/A) - AGR	V-T/A-AGR
Embedded verbs	V-T/A- C -AGR	V-T/A N-AGR
Complementiser	✓	×

I suggest that these dissimilarities are the manifestation of a difference in the morphological structure of Turkish and Yakut. Turkish has three slots available within the inflectional domain whereas Yakut has two slots, illustrated below:

(18)	<i>Turkish</i>	<i>Yakut</i>
	V..... X ¹ X ² X ³	V..... X ¹ X ²

where X stands for any inflectional morpheme. The difference is thus reducible to the space that is available in each language: Turkish has a morphological template that has three slots for inflectional affixes, whereas the morphological template in Yakut has two slots.

The existence of morphological templates explains the absence of a second T/A marker in embedded verbs in Turkish, this position being occupied by C. It also explains why the agreement marker in Yakut has to occur on the noun rather than on the verb in an embedded clause. If it were to occur on the verb it would require a nominaliser, as it is a nominal agreement marker. However, the presence of a nominaliser would itself have the effect of filling up the morphological space. An additional suffix, in this case agreement, would violate the constraint that no more than two suffixes are allowed in Yakut.

4. Further evidence for the presence of morphological templates

The presence of morphological templates explains a number of constructions in Turkish and Yakut, such as the occurrence of buffer stems, the structure of auxiliaries and possessive constructions.

4.1 Buffer stems

Both Turkish and Yakut use buffer stems as a strategy for further affixation. As explained in 1.2, the embedded verb in Turkish does not have the array of T/A affixes found in main verbs. A verbal form such as (19a) does not have a corresponding embedded form (19b), which is ungrammatical for the reasons explained above. Instead, (19c) corresponds to (19a).

- (19) a. Oku-muş-tu-n
 1 2 3
 read-HS-PAST-AGR
 You had read.
- b. *oku-muş-tu-ğ-un kitap
 read-HS-PAST-C-AGRbook
 1 2 3 4
 (Intended interpretation: the book that you had read)
- c. oku-muş ol-du-ğ-un kitap
 read-HS AUX-PAST-C-AGR book
 1 1 2 3
 the book that you had read.

(19c) contains the copular verb *ol* 'be' which, here, serves no other purpose than to act as a buffer stem to which additional affixation can be attached. Yakut also uses a copular verb as a buffer stem. A form such as (20a) which is semantically possible does not exist, as the morphological template of Yakut has two slots for the inflectional domain. Instead the copular verb *olor* 'be' is used solely for reasons of morphology.

- (20) a. *Si-i-bit-im.
 1 2 3
 eat-AOR-PAST-AGR
 (Intended interpretation: I had been reading.)
- b. Si-i olor-but-um.
 eat-AOR AUX-PAST-AGR
 1 1 2
 I had been reading.

4.2 Modals

Modal auxiliaries exhibit similar properties. In constructions with the modal verb *buol* 'can/may' in Yakut, the two stems share the inflectional affixes. Again, as expected, neither stem has more than two suffixes, and the person marker can either appear on the main verb as in (21b) or on the auxiliary, as in (21c):

- (21) a. *Si-i(r)-i-im buol
 eat-AOR-FUT-AGR may
 1 2 3
 (Intended interpretation: I will probably eat.)
- b. Si-i(r)-im buol-uox
 eat-AOR-AGR may-FUT
 1 2 1
 I will probably eat.

c. Si-i(r)	buol-uoy-um
eat-AOR	may-FUT-AGR
1	1 2

I will probably eat.

4.3 Double possessives

Embedded clauses in Turkish and Yakut are nominal constructions which have similar properties to possessive phrases. In the possessive construction in Turkish, the possessor is in the genitive case, as is the subject of an embedded clause. Both constructions have nominal agreement:

(22)	<i>embedded clause</i>		<i>possessive phrase</i>
	ben-im dik-ti-ğ-im elbise		ben-im elbise-m
	I-GEN sew-PAST-C-AGR dress		I-GEN dress-AGR
	the dress that I sewed		my dress

The position of the agreement marker in embedded clauses makes it possible to have double possessive constructions in Turkish, where the relative clause is embedded in the possessive construction:

(23)	ben-im şu sen-in dik-ti-ğ-in elbise-m
	I-GEN this you-GEN sew-PAST-C-AGR2 dress-AGR1
	this dress of mine which you sewed

In Yakut, embedded clauses are also similar to possessive phrases:

(24)	<i>embedded clause</i>		<i>possessive phrase</i>
	min si-ebit balı-ım		min balı-ım
	I(NOM) eat-PAST fish-AGR		I(NOM) fish-AGR
	the fish that I ate		my fish

However, an embedded clause cannot occur within a possessive phrase:

(25)	a. *min en si-ebit balı-ım
	I you eat-PAST fish-AGR1
	b. *min en si-ebit-in balı-ım
	I you eat-PAST-AGR2 fish-AGR1
	c. *min en si-ebit balı-ım-ın
	I you eat-PAST fish-AGR1-AGR2

(Intended interpretation: the fish of mine which you ate)

Although morphologically well-formed, (25a) is ungrammatical because of the absence of an agreement marker. The occurrence of second person agreement marking on the embedded verb, as in (25b) results in ungrammaticality as well, this time for both morphological and syntactic reasons. Although complying with the requirements of the morphological template, it is ungrammatical as a result of the presence of nominal agreement on the verb, as AGR2 is a member of the nominal paradigm. If it were to appear, it would require a nominaliser. However, the presence of a would violate the template, as explained above, in section 3. The ungrammaticality of (25c) arises from

two agreement markers occurring one after the other, a violation of slot-(affix) type mismatches.

5. The nature of morphological slots

The discussion so far has centered around motivating morphological templates and how they provide the slots for hosting affixes. A question that arises at this point concerns the pairing of affixes and slots. The fact that the order of affixes is not free and that there are restrictions on the co-occurrence of certain affixes indicate that affixes and slots must match, an issue discussed in Sells (1995). There it is argued that the occurrence of certain affixes in certain positions, as well as the non co-occurrence of particular affixes are due to their being typed to occur in a particular slot.

The ungrammaticality of (25c) can then be explained in terms of slot-type mismatches, where, for two affixes of the same type (AGR1 and AGR2), there is only one slot. Similarly, there is no apparent reason for certain sequences in Turkish to be ungrammatical, except for slot-type mismatches, as in (26):

- (26) *Oku-yor-acak-sın.
read-PROG-FUT-AGR
(Intended interpretation: You will be reading.)

The only explanation for the ungrammaticality of (26) is that the progressive and the future markers are typed to occur in the same slot. This suggestion is supported by the grammaticality of (27) where both affixes occur in the first slot after the stem:

- (27) Oku-yor ol-acak-sın
read-PROG AUX-FUT-AGR
You will be reading.

6. Implications for syntax

Certain aspects of the proposal regarding morphological templates are relevant to the structure of syntax. Firstly, the present proposal assigns inflection to the morphological component. This is not, in itself, at variance with having a syntactic reflex for each inflectional affix. Morphological templates and slot-type matching requirements could determine certain aspects of inflectional morphology, and other aspects could be represented in syntax. The question is, of course, whether one component replicates the effects of the another component.

In current analyses in syntax, inflectional affixes are posited as heads of functional projections. Following Pollock (1989) the earlier version of this research programme involves generating each functional category as a head. Words are then formed in the syntax as a result of head movement. In the Minimalist Programme (Chomsky 1993) words are formed in the lexicon but they still have to go through the process of head movement for feature checking. In both versions, the order of the morphemes is determined by the selection properties of heads.

There is no a priori reason to object to such a design. In fact, it is even desirable to have a maximal projection for each functional head for its specifier and complement positions, if it can be shown that the presence of these positions account for the facts. In the literature there are a number of convincing arguments in favour of positing each functional category as a syntactic head (Pollock 1989, Ouhalla 1991, Bobaljik and Jonas

1996, among others). As is well known, one of the motivations behind the proliferation of functional projections within inflection is the contrast between word order variations in French and English, which is argued to be an outcome of a difference in the occurrence of adverbs in different specifier positions (Pollock 1989). Ouhalla (1991) derives the difference between VSO and SVO languages from the difference between the ordering of tense and agreement morphology, and crucially, from the specifier positions that are available because of the status of these morphemes as syntactic heads.

As for Turkish, there is compelling evidence regarding specificity that there are two specifier positions, one related to VP, the other to IP (Kennelly 1993). However, it has to be seen whether the proliferation of functional categories serves a purpose in Turkish. To my knowledge, there is no evidence that Turkish requires the specifier positions associated with other inflectional elements.³ The role of the availability of various specifier positions in determining word order, as summarised above, does not apply to Turkish because these arguments are based on the interference of the position of the specifier with the path of head movement, both of which are on the left (for the languages mentioned in Pollock 1989 and Ouhalla 1991). As Turkish is an SOV language, it has rightward head movement, whereas specifier positions are on the left. Therefore arguments based on the interference of specifier positions with head movement are irrelevant for Turkish.⁴ If the structure of a morphologically complex word is derivable by mechanisms specific to the morphological component, there is no reason to represent the complexity of the same word in syntax, unless there are independent reasons for doing so.

Conclusion

Certain differences in the morphology of Turkish and Yakut can be explained in terms of morphological templates which determine the length of a word. The data indicates that Turkish has three slots in its inflectional domain and Yakut has two. Language internal and cross-linguistic variations in verbal and nominal morphology which are a consequence of this constraint are summarised below:

³It has recently come to my attention that Tosun (forthcoming) suggests that scope properties of certain adverbs might require an analysis where they are generated in the specifier of TP. At the time of writing this paper there wasn't conclusive evidence for this suggestion.

⁴ It might be tempting to argue, along the lines of the antisymmetry of syntax hypothesis (Kayne 1994) that Turkish is also left-headed, ie. SVO, in which case it would have to be shown that the arguments above apply to Turkish as well. Keleşir (1996) has shown that positing an underlying SVO structure for Turkish creates insurmountable problems for the representation of clause structure.

(28)	<i>Turkish</i>	<i>Yakut</i>
Main verb	V...T/A-T/A-AGR	V...T/A-AGR
Embedded verb	V...T/A-C-AGR	V...T/A
Embedded clause	V...T/A-C-AGR N	V...T/A N-AGR
Buffer stems/ Auxiliaries	V...T/A V-T/A-(T/A)-AGR V...T/A V-T/A- C -AGR	V...T/A V-T/A-AGR V...T/A-AGR V-T/A
Complementiser	✓	×
Double Possessives	V...T/A-C-AGR _x N-AGR _y	×
Structure	V..... X ¹ X ² X ³	V..... X ¹ X ²

The hypothesis regarding morphological templates is a morpheme based approach which supports the presence of an independent morphological component. Although certain aspects of morphology overlap with phonology and syntax, and although certain properties of grammar may apply to all components, a morphological template is definable only in morphological terms. It does not seem to be reducible to phonological structure since morphemes which have one or more syllables behave similarly with respect to morphological slots. It is not reducible to syntax either, as there seem to be purely morphological reasons for some inflectional phenomena to be as they are.

One of the outcomes of this proposal is that it provides a theoretical explanation for morphological typology. If this proposal is correct, the description of the morphological structure of languages as isolating, synthetic and agglutinating can be recast partly in terms of the morphological space they have.

References

- Anderson, S. (1992) *A-morphous Morphology*. Cambridge: CUP.
- Bach, E. (1996) "On the Grammar of Complex Words", in A.-M. Di Sciullo (ed.) *Configurations; Essays on Structure and Interpretation*. Somerville: Cascadilla Press.
- Bobaljik, J.D. & D. Jonas (1996) "Subject Positions and the Roles of TP." *Linguistic Inquiry* 27/2.
- Bresnan, J. and S. Mchombo (1995) "The Lexical Integrity Principle: Evidence from Bantu." *Natural Language and Linguistic Theory* 13/2.
- Chomsky, N. (1993) "A Minimalist Program for Linguistic Theory", in K. Hale and J. Keyser (eds.) *The View from Building 20*. Cambridge, Mass.: MIT Press.
- Di Sciullo, A.-M. (1996) "Atomicity and Relatedness in Configurational Morphology", in A.-M. Di Sciullo (ed.) *Configurations; Essays on Structure and Interpretation*. Somerville: Cascadilla Press.
- Di Sciullo, A.-M. and Williams, E. (1987) *On the Definition of Word*. Linguistic Inquiry Monograph 14. Cambridge: MIT Press.
- Göksel, A. (1997) "Morphological Asymmetries between Turkish and Yakut", in K. İmer and N.E. Uzun (eds) *Proceedings of the VIIIth International Conference on Turkish Linguistics*. Ankara: Ankara Üniversitesi Basımevi.

- Kayne, R. (1994) *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Kelepir, M. (1996) The Implications of the Antisymmetry Theory of Syntax on Turkish Word Order. Unpublished MA dissertation, Bosphorus University.
- Kennelly, S. (1992) "Turkish Subordination: /-CP, -Tense, +Case/. Talk presented at the 6th Conference on Turkish Linguistics, Eskişehir.
- Kennelly, S. (1993) "Non-specificity in Turkish: Inherent Case Revisited". Ms.
- Kural, M. (1993) "V-TO(I-TO)-C in Turkish." *UCLA Occasional Papers in Linguistics*, 11.
- Ouhalla, J. (1991) *Functional Categories and Parametric Variation*. London: Routledge.
- Özsoy, S. (1988) "On Complementaton in Turkish," in Koç, S. (ed.), *Studies on Turkish Linguistics*. Ankara: Middle East Technical University Press.
- Pollock, Y.-J. (1989) "Verb Movement, Universal Grammar, and the Structure of IP." *Linguistic Inquiry* 20/3.
- Sells, P. (1995) "Korean and Japanese Morphology from a Lexical Perspective". *Linguistic Inquiry* 26/3.
- Tosun, G. (forthcoming) Functional Projections in Turkish.

This research has been funded by the School of Oriental and African Studies (grant XG15) and the British Academy (grant BA-AN1438/APN1654). I am grateful to Lena Lukchina, Vassily Klimovich Pavlov, Vladislav Terekov and Svetlana Yegorova who worked as Yakut informants in the project.