

Gapping and Word Order in Turkish

Cem Bozsahin
Middle East Technical University
bozsahin@metu.edu.tr

1. Introduction

Although word order is believed to be pragmatically controlled in Turkish, it nevertheless emerges from syntax. Identifying its origins in the lexicon is crucial, at least for the lexicalist theories of grammar, which would go one step further than phrase structure grammars and look for its source in the lexicon. The search for the ‘basic’ word order in a language must appeal to the grammatical processes that are known to be most dependent on surface word order: extraction and coordination.

Gapping (identical verb deletion under coordination) is a good source to locate word orders for several reasons. It is a phenomenon that exhibits the interaction of these processes. As Ross (1970:251) pointed out, it is susceptible to surface word order (“the order in which gapping operates depends on the order of elements at the time that the rule applies”). It is also notorious for providing surface constituents such as subject-object (SO) or object-subject (OS) that are traditionally considered non-constituents. Since they behave as units in coordination, shown syncategorematically as $X \& X \Rightarrow X$, any difference they show in gapping behaviour requires that they have a distinguishable surface category (hence bear a surface category to begin with). How—and if—this category is derivable differentiates theories of surface grammar from movement-based frameworks. In this paper, we follow the first track and present empirical data on gapping in Turkish. We present the explicit construction of the grammar-lexicon and elaborate on its consequences elsewhere (Bozsahin 2000).

Gapping operates forward if identical elements (the verb) are on the left (cf. type A-B in Table 1). This pattern is a sign of leftward grouping VO (verb-object). It operates backward if V is on the right (cf. type C in Table 1). It is a sign of rightward grouping OV. This is confirmed in most common word orders SOV (subject-object-verb), SVO and VSO (we refer to surface grammatical relations unless otherwise stated):

SOV (Japanese): SO & SOV but *SOV & SO

SVO (English): SVO & SO but *SO & SVO

VSO (Irish): VSO & SO but *SO & VSO

Table 1. Ross's (1970) classification of Gapping patterns.

Type	Base	Pattern
A	SVO	SVO & SVO =>gap SVO & SO (forward gapping)
B	SVO	SVO & SVO =>gap SVO & SO =>scramble SOV & SO
C	SOV	SOV & SOV =>gap SO & SOV (backward gapping)
D	SOV	SOV & SOV =>gap SO & SOV =>scramble *SO & SVO

2. Gapping in Turkish

Studies on Turkish word order, e.g., Enç (1991), Erguvanlı (1984), Erkü (1983), Hoffman (1995), Kılıçaslan (1994), and Turan (1995) reflect the dominance of discourse-functional concerns for identifying the basic word orders. Turkish is generally regarded as a free word order language and all six variations of S, O, V are attested.

Turkish follows the SOV pattern in gapping (type C) as in examples (1a-b), with the added complexity that forward gapping is also possible (1c-d). This was the main reason for Ross to claim that Turkish must have VO grouping and the SVO basic word order.

- (1)a. Adam kitabı, çocuk da dergiyi okudu. (SO & SOV)
 man book-ACC child CONJ mag.-ACC read-PAST
 'the man read the book, and the child, the magazine.'
- b. Kitabı adam, dergiyi de çocuk okudu. (OS & OSV)
- c. Adam kitabı okudu, çocuk da dergiyi. (SOV & SO)
- d. Kitabı adam okudu, dergiyi de çocuk. (OSV & OS)

Table 2. Backward gapping.

SO & SOV	OS & OSV
*SO & OSV	*OS & SOV
*SO & XVY	*OS & XVY for X,Y=S,O
*SO & VXY	*OS & VXY for X,Y=S,O

We conducted an experiment with 45 native speakers, each evaluating around 12 utterances. Table 2 shows the results on backward deletion of identical verbs. The rightmost conjunct in which the undeleted verb appears is verb-final. The first row indicates that SO and OS must be treated as constituents at the surface structure. The

second row shows that their surface categories must differ, otherwise SO & OS and OS & SO coordination would be grammatical.

Table 3. Forward Gapping.

SOV & SO	OSV & OS
SOV & OS	OSV & SO
SVO & SO/OS	?OVS & SO/OS
?VOS & OS/SO	?VSO & SO
??VSO & OS	

Table 3 shows the results on forward deletion of identical verbs. As evident from the first two rows, structural parallelism that we observe in Table 2—which has earlier prompted all too powerful transderivational constraints (Hankamer, 1973)—do not apply in forward gapping. It appears that the judgments can be ranked according to the position of the verb in the leftmost conjunct. All informants found verb-final leftmost conjuncts grammatical, irrespective of the surface structure of the right conjuncts. Verb-medial leftmost conjuncts are second most acceptable with SVO & SO ranking along with verb-final versions. Verb-initial leftmost conjuncts are considered problematic but acceptable by majority of the informants. Judgments improve when examples are provided in context, which we did only for questionable cases. For instance, VSO & OS, which is considered ungrammatical by the majority of the informants in null context, is felicitous in context (2). Providing intonational contours also affect judgments in borderline cases. For instance, informants who considered SVO & SO questionable considered it grammatical with falling intonation on the left conjunct and rising intonation on the right (3).

- (2) Context: a film director wonders whether the producer has read the novel, after the producer claims that he read the screenplay:

Bence	okumuş	yapımcı	filmin	senaryosunu,
I-ADV	read	producer	film-GEN	screenplay-POSS-ACC
romanınıysa		yardımcısı		
novel-POSS-CONJ		assistant-POSS		

‘I think the producer has read the screenplay of the film, and his assistant, its novel.’

- (3) ADAM OKUMUŞ dergiyi, çocuk DA KİTABI
 man read magazine-ACC child CONJ book-ACC
 ‘The man has read the magazine, and the child, the book.’

3. Basic Word Order and Gapping

In the remainder of the paper, we consider three possibilities for basic word order: (1) SOV, (2) SOV and SVO, and (3) SOV and OSV. These alternatives will be tested on the empirical data in section 2, with the following theoretical concerns:

(i) All syntactic rules operate on constituents, that is, their 'input' and 'output' are always constituents. This 'Constituent Condition on Rules' is perhaps the least contentious assumption of Generative Grammar (Chomsky 1975). Coordination being a syntactic rule, it also must operate on constituents, hence X in X & X => X schema must be a constituent.

(ii) The implication of this principle on gapping is that SO and OS must be constituents. In a theory of surface grammar without movement and empty categories, X must be a *surface* constituent because there is no auxiliary operation to take a non-constituent X' and turn it into constituent X.

(iii) Surface categories of SO and OS must be distinguishable. Parallelism in backward gapping (e.g. SO & SOV but *SO & OSV) requires such distinction. It might be argued that different syntactic processes may be at work in backward and forward gapping (Kornfilt p.c.), and that so-called backward gapping is a result of another process such as right-node raising, whereas forward gapping is true gapping. But in a theory without movement, accounts of node-raising also appeal to surface constituency (cf. Abbott's (1976) warning that constituents can be right-node-raised, but right-node raising does not entail constituency). The process is reduced to another case of constituent coordination at surface structure, just like gapping (Steedman 1990). In other words, both gapping and node-raising behaviour originates from the lexicon (assuming that the lexicon specifies combinatory possibilities of lexical items as well, which is a basic tenet of lexicalist grammars). There would be no proliferation of construction-specific category assignment in the verb lexicon either, because all processes are subsumed under a single category, and all are subject to the same constituent condition on rules.

(iv) Since Turkish is a free word order language, the variants of the basic word order(s) must be marked somehow so that pragmatics can do its work by deciding on the discourse felicity of the word order. Having no access to derivational structure, discourse grammar must rely on categories of syntactic derivations, most likely working in parallel with the syntactic component (Steedman 2000).

For example, if we consider SOV as the only basic word order, all word orders in which S or O are to the right of the verb will be marked. Same is true for SOV-and-OSV hypothesis. However, SOV-and-SVO hypothesis would make a weaker claim about discourse sensitivity of the object's position; rightward placement of O is now captured in SVO basic word order, hence it is not marked. Only subject surface extraposition to the right would be marked under this assumption.

Steele (1978) observed over a hundred languages that in VSO languages, VOS usually surfaces as the alternate word order, but not, say, OVS. Similarly, SOV languages have the alternate OSV, not SVO. The main claim of this paper is that gapping data on Turkish also conforms to this pattern and is indicative of two alternate word orders for Turkish. We show this by examining each hypothesis at a closer look with respect to the gapping data.

Directionality of the verb is critical in evaluating the hypotheses. For the purpose of this paper, it suffices to use a simplified Categorical Grammar-like notation to avoid a full exposure to combinatorial and directional alternatives in the lexicon. Let us assume that in $X\{Y\}$, X denotes the result category, and $\{Y\}$ the set of arguments. This is based on Hoffman's (1995) notation, but with the notable absence of slashes (due to lack of space, we concentrate on lexical matters and eschew grammatical inference by slashes). $S\{NP1<,NP2>\}$ denotes the surface category of an item that yields the category S (for sentence) if it finds an NP to the right ($NP2>$) and an NP to the left ($NP1<$). '>' and '<' show the directionality of the argument (right and left). The indices 1 and 2 can be replaced by case decorations to make grammatical relations explicit; $S\{NPnom<,NPacc>\}$ indicates an item that yields a sentence if there is an $NPacc$ to the right and $NPnom$ to the left, i.e., a transitive verb in an SVO language like English. VP constituency in English can be represented by further refining this category as $(S\{NPnom<\})\{NPacc>\}$ in which the object has to combine first to yield VO.

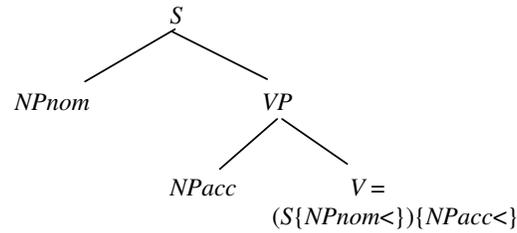
Alternatives for directionality can now be made explicit. If basic word order is SOV only, the surface category of the verb is $(S\{NPnom<\})\{NPacc<\}$. This is because the verb has to find an $NPacc$ to the left *first*, before it can look for $NPnom$ on the left. If basic word orders are SOV and SVO, the verb's surface category is $(S\{NPnom<\})\{NPacc\}$. Although the subject ($NPnom<$) must be to the left, the object ($NPacc$) can be either to the left or right, and $NPacc$ combines first. Similarly, if the basic word orders are SOV and

OSV, the verbal category is $S\{NPnom<,NPacc<\}$; although the arguments are both to the left, the order in which they combine is unspecified. We look at the alternatives below.

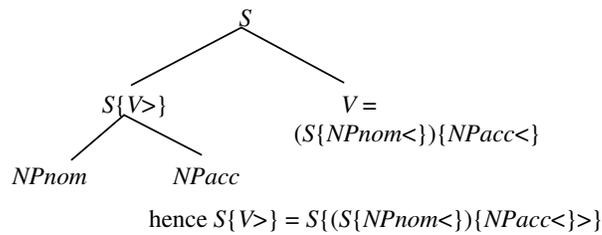
3.1 SOV Hypothesis

The surface category of the verb in SOV hypothesis is $(S\{NPnom<\})\{NPacc<\}$. This is the most common view on Turkish word order (Erguvanlı 1984, Erku 1983, Göksel 1993, Kılıçaslan 1994, Kornfilt 1997). It is clear that this hypothesis can correctly account for backward gapping data in Table 2. SO and OS have different surface categories for the following reason (hence *SO & OSV and *OS & SOV predictions): Phrase structure in (4) shows constituency and order for the assumed SOV basic word order. The structure in (5) is what is needed for SO & SOV coordination as in (6) by the constituent condition on rules (X in $X \& X \Rightarrow X$ is a constituent). $S\{V>\}$ denotes a surface category that yields S when a verb of category V is to the right. This category of the non-traditional bracketing $[NPnom NPacc]$ is obtained in Categorical Grammar by type-raising and composition (cf. Steedman 2000), which are presumed to be part of universal grammar.

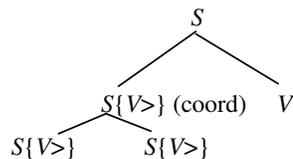
(4)



(5)



(6)



However, the category of OS with the bracketing $[NPacc NPnom]$ is different. Phrase structure in (7) shows what is needed for OS constituency. But the verb's lexical category

(which in our surveys presented strongest agreement among informants) is compromised. This asymmetry cannot be resolved with a single directionality alternative for the verb.

- (10) a. Makaleyi adam okudu, çocuk da kitabı.
 article-ACC man read child CONJ book-ACC
 ‘The man read the article, and the child, the book.’
 b. Adam makaleyi okudu, kitabı da çocuk.

3.2 SOV and SVO Hypothesis

The surface category of the verb in SOV-and-SVO hypothesis is $(S\{NP_{nom}<\})\{NP_{acc}\}$. This is a combination of Ross’s (1970) SVO hypothesis and Koutsoudas’s (1971) SVO-and-SOV hypothesis. Koutsoudas claimed that Turkish must be SOV as well because of split coordination of the type SOV & O (11).

- (11) Ayşe makaleyi de okudu romanı da.
 article-ACC CONJ read novel-ACC CONJ
 ‘Ayşe read the article and the novel.’

This hypothesis can account for backward gapping data the same way as the SOV hypothesis. Addition of SVO basic word order does not provide like categories for SO and OS because in this hypothesis, the subject precedes the verb in basic word order thus OS is still marked and SO is treated as basic.

However, for the same reason as SOV-only hypothesis, forward gapping is compromised. For instance, OSV & SO and SOV & OS cannot be accounted for. In sum, Ross’s motivation for considering SVO as basic was forward gapping. But SVO can account only for parallel structures SVO & SO and OSV & OS.

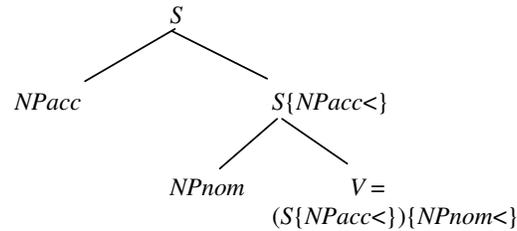
3.3 SOV and OSV Hypothesis

The surface category of the verb in SOV-and-OSV hypothesis is $S\{NP_{nom}<,NP_{acc}<\}$. This category in fact stands for two directionally more specific categories in which the argument NPs are specified one at a time. Unfolding a single category in such a fashion reveals why SO and OS receive different surface categories under this hypothesis as well.

The first verbal category is $(S\{NP_{nom}<\})\{NP_{acc}<\}$. This stands for the basic word order SOV, cf. the verb’s category in section 3.1. All the results of SOV-hypothesis on backward gapping therefore apply to this part of the current hypothesis. This category can maintain the SO/OS surface-categorial distinction.

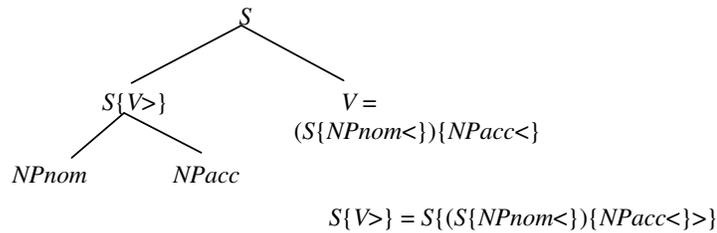
The second verbal category is $(S\{NPacc<\})\{NPnom<\}$. Coming from the lexicon, this category projects the phrase structure in (12); $NPnom$ and V form a constituent by this category because the verb first combines with $NPnom$, and yields the result $(S\{NPacc<\})\{NPnom<\}$.

(12)



OS surface constituency is obtained similar to the process that derives (5) from (4). We get $S\{V>\}=S\{(S\{NPacc<\})\{NPnom<\}>\}$, which is a different category than that of SO (13). Hence SO and OS are distinguishable surface-categorially, and paralelism in backward gapping is accounted for.

(13)



The predictions of SOV-and-OSV hypothesis for forward gapping are as follows:

(i) Since SOV and OSV are basic word orders, forward gapping of the type SOV & SO/OS and OSV & SO/OS would be unproblematic. This is due to the fact that these word orders yield unmarked sentences, and SO and OS are unmarked as well. This prediction is confirmed in Table 3.

(ii) SVO & XY, VSO & XY, VOS & XY and OVS & XY (where X and Y are S or O) would be problematic because non-verb-final word orders SVO, VSO, OVS and VOS are marked under this hypothesis, and XY may be marked or unmarked. This is also confirmed in Table 3. We refer the reader to (Bozsahin 2000) for the description of how the SO/OS constituency in the left conjunct is revealed to the right conjunct so as to maintain constituent condition on rules.

4. Conclusion

SOV-and-OSV hypothesis can account for backward and forward gapping by maintaining the asymmetry between these two patterns of gapping. The asymmetry originates in the lexicon from the verb's directionality alternatives, and is projected to surface structure via a direction-preserving grammar.

One implication of this conclusion is that Turkish is not a scrambling language but an extraposition language (we are of course referring to surface extraposition). Non-verb-final word orders in Turkish are results of extraposition. Scrambling cannot maintain the directional sensitivity that the asymmetry in gapping requires.

I believe the hypothesis also clarifies the grammatical origins of *flexible* word order in Turkish, which has heretofore been formulated in discourse-functional terms. The source is the directional specificity of the verb's lexical category. English has rigid word order because the verb's only directionality in the lexicon is SVO. Turkish has flexible word order because the verb's lexical category corresponds to SOV and OSV.

Acknowledgements Thanks to Jason Baldrige, Jaklin Kornfilt, Mark Steedman and Ümit Turan for comments and suggestions.

5. References

- Abbott, B. (1976). Right Node Raising as a Test for Constituenthood. *Linguistic Inquiry* 7:639-642.
- Bozsahin, C. (2000). Directionality and the Lexicon: Evidence from Gapping. ms.
- Chomsky, N. (1975). *The Logical Structure of Linguistic Theory*. Chicago: Univ. of Chicago Press.
- Enç, M. (1991). The Semantics of Specificity. *Linguistic Inquiry* 22:1-25.
- Erguvanlı, E. (1984). The Function of Word Order in Turkish Grammar. Ph.D. Thesis, UCLA.
- Erkü, F. (1983). Discourse Pragmatics and Word Order in Turkish. Ph.D. Thesis, Univ. of Minnesota.
- Göksel, A. (1993). Levels of Representation and Argument Structure in Turkish. Ph.D. Thesis, SOAS.
- Hankamer, J. (1973). Unacceptable Ambiguity. *Linguistic Inquiry* 4:17-68.
- Hoffman, B. (1995). The Computational Analysis of the Syntax and Interpretation of "Free" Word Order in Turkish. Ph.D. Thesis, Univ. of Pennsylvania.
- Kılıçaslan, Y. (1994). Information Packaging in Turkish. MS Thesis, Univ. of Edinburgh.
- Kornfilt, J. (1997). *Turkish*. London: Routledge.
- Koutsoudas, A. (1971). Gapping, Conjunction Reduction, and Coordinate Deletion. *Foundations of Language* 7:337-386.
- Ross, J.R. (1970). Gapping and the Order of Constituents. In Bierwisch, M. and Heidolph K.E. (eds.), *Progress in Linguistics*. The Hague: Mouton.
- Steedman, M. (1990). Gapping as Constituent Coordination. *Linguistics and Philosophy* 13: 207-263.
- Steedman, M. (2000). *The Syntactic Process*. Cambridge, MA: The MIT Press.
- Steele, S. (1978). Word Order Variation: A Typological Study. In Greenberg, J. (ed.), *Universals of Human Language*. Stanford Univ. Press.
- Turan, Ü. (1995). Null Versus Overt Subjects in Turkish Discourse: A Centering Analysis. Ph.D. Thesis, Univ. of Pennsylvania.