

# Drawing Tree Diagrams: Problems and Suggestions

Qiang Wang

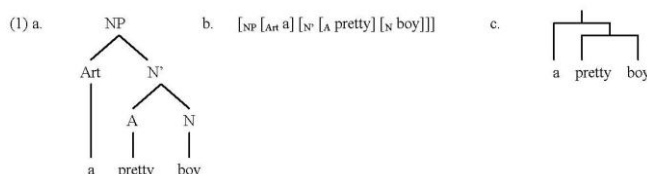
College of Foreign Languages, Chongqing University of Posts and Telecommunications, Chongqing, China  
 Email: qiang\_wong@163.com

**Abstract**—Starting from the theoretical goal of syntax, and taking the difference between pedagogical grammar and syntax into consideration, this article, based on the author's teaching practice and review of exam papers, analyzes common problems in drawing tree diagrams such as ternary branching and improper marking of grammatical categories, followed by proposals for constructing such representations, with a view to helping students to achieve reasonable and elegant drawings, and also to helping teachers to teach syntax more fruitfully.

**Index Terms**—tree diagram, drawing, problems, suggestions

## I. INTRODUCTION

Syntax is the study of the principles and processes by which sentences are constructed in particular languages. Syntactic investigation of a given language has as its goal the construction of a grammar that can be viewed as a device of some sort for producing the sentences of the language under analysis. (Chomsky, 1957, p.11) Put simply, syntax studies how to combine words into bigger linguistic units – phrases or sentences. At present, introductory linguistics courses are generally available to juniors majoring in English at most Chinese universities, where syntax is usually viewed by both teachers and students as one of the most important and difficult chapters of the course. In syntax, tree diagram and labelled bracketing are widely employed to characterize hierarchical structures of phrases or sentences. A tree diagram, or tree, is a two-dimensional diagram used in generative grammar as a convenient means of displaying the internal hierarchical structure of sentences as generated by a set of rules. (Crystal, 2008, p.494) Labelled bracketing, or labelling, is a term in grammatical analysis for the explicit marking of the parts or stages in a structural analysis of a sentence. (*ibid.* p. 263) They are shown in (1a) and (1b) respectively in characterization of the structure of *a pretty boy*.



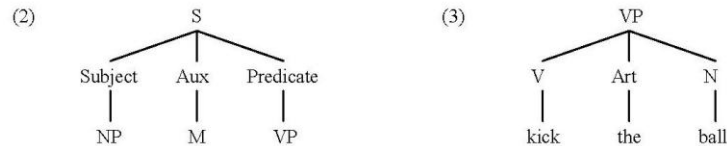
(1a) and (1b) are in contrast to (1c), immediate constituent analysis (IC analysis), also a way of characterizing phrase or sentence structures, proposed by descriptive linguistics, or structural linguistics. IC analysis segments or cuts a phrase or sentence into two and then cut these parts into two until morphemes are obtained. (Gleason, 1961, p.128-148) But it segments structures in form only while marking no lexical and phrasal categories, which is not helpful to understand the internal constituents of a structure, and the semantic relation between constituents in particular. In contrast, tree diagram is intuitive and vivid, capable of explicitly demonstrating the internal hierarchy of a phrase or sentence, and the intimacy of its constituents and their sequence of merge, although it takes up more space. In less space, usually one or two lines, labelled bracketing is not so intuitive. Moreover, both (1a) and (1b) mark categories or nodes, making up for the disadvantage of immediate constituent analysis. In general, tree diagram is the most widely used tool of formalizing phrase or sentence structures in syntax, in case the space is available.

Due to the great difference between tree diagram and proposals of pedagogical grammar and descriptive linguistics, students find it difficult to draw tree diagrams. In English linguistics textbooks currently used in Chinese universities, such as Hu (2001), Dai & He (2002) and Liu & Wen (2006), there are no detailed elaborations on how to draw tree diagrams in chapters of "Syntax". Students cannot find satisfactory solutions for such representations, either, in further readings recommended by teachers, such as Radford (1988), Ouhalla (1999) and Poole (1999). In view of the current situation, this article analyzes ten common problems that students incline to have in drawing tree diagrams for assertives, negatives, interrogatives and some relative clauses, found in practical teaching of syntax for juniors of English major with Chongqing University of Posts and Telecommunications in southwest China and review of their exam papers. It is followed by suggestions, including general idea and procedures, with the aim of helping students with drawing of reasonable and elegant tree diagrams, and to help teachers with more fruitful teaching of syntax.

## II. ANALYSIS OF PROBLEMS IN DRAWING TREE DIAGRAMS

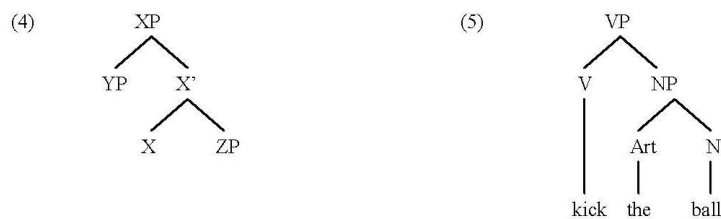
### A. Ternary Branching

Currently available textbooks adopt both ternary and binary branching trees arbitrarily without any comment on their respective merits and demerits. Part of earlier research of syntactic theory also employed ternary branching trees characterizing sentence structures as shown in (2) below (see Chomsky, 1965, p.69), for such sentence as “Sincerity may frighten the boy”. Beginners of syntax tend to draw ternary branching trees as shown in (3) below for the sole purpose of simplicity, as they put all the words in the same level and then merge them once and for all.



All would agree that, the three words in (3) should not be in the same level. Instead, “the” is semantically closer to “ball” than to “kick”. In this way, “the ball” can form a constituent. So the phrase marker (3) is not reasonable.

Ternary branching makes prominent the grammatical functions and categories of constituents of a sentence as well as the linearity of constituents. However, it cannot effectively demonstrate the internal hierarchy of a phrase or sentence, the semantic intimacy of its constituents and their sequence of merge. With the development of syntactic theory, ternary branching has been replaced by binary branching. Derived from immediate constituent analysis, the binary branching method parses a big unit into two smaller units each time in segmentation of structures. Theoretically speaking, binary branching is a fundamental principle of X-bar theory (see Jackendoff 1977) as shown in (4) below and the antisymmetry of syntax (see Kayne 1994). With binary branching adopted, (3) above shall be rewritten as (5) below, where *the* merges with *ball* into an NP which in turn merges with *kick* into a VP.

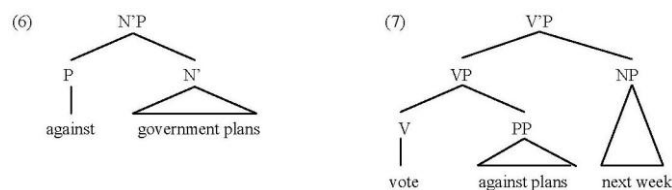


Xiong (2009) argues that the phrase structure in the generative grammar is endocentric to constrain possible phrase structures, and its branch is binary to represent the asymmetry between the constituents in the structure. In modern syntax, binary branching has become a fundamental principle of drawing tree diagrams while merging two grammatical categories that are semantically close to each other into a larger category each time has become a basic method for representation and a core operation in the theory of grammar. Binary branching trees like (5) shall be applicable to most, if not all, characterizations of phrase structure in syntax.

### B. Improper Marking and Representing of Categories

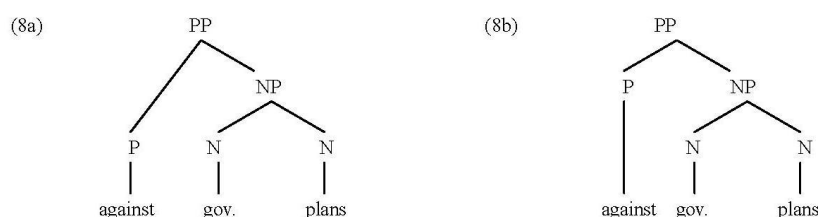
Marking categories is a basic skill for drawing of tree diagrams. Nevertheless, categories in syntax are different from parts of speech, or word classes, in pedagogical grammar. Perfect comparison and contrast between the two can be found in Xu (2007). Some students tend to use parts of speech in pedagogical grammar to mark categories in trees, e.g. marking auxiliaries such as *will*, *would*, *may* and *might* as Aux, a sentence as S, and a subordinate clause as S'. It is known that one of the theoretical goals of syntax is to pursue uniformity and generalization in characterizing phrase structures. In syntax, I(nflection) includes not only Aux but also the infinitival marker *-to* and inflectional morphemes such as *-ed* and *-s*. Hence, *will* shall be marked I. I is usually regarded as the head of a sentence, so a sentence is in fact a projection of I. That is why a sentence shall be marked as IP rather than S. A subordinate clause is usually composed of a complementizer, or C, including *that*, *if*, *whether* and *for* in English, followed by a sentence, or IP. As C indicates whether the subordinate clause is declarative or interrogative, C is usually regarded as the head of a subordinate clause. In other words, a subordinate clause is a projection of C. Hence, a subordinate clause shall be marked as CP rather than S'.

In addition to category marking, problems with representing of categories are also found, as shown in (6) and (7) below.



### C. Ill-formed Structural Hierarchy

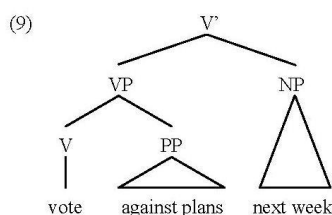
The linear sequence from left to right is a superficial property of sentence, but the structural hierarchy is its nature (Chomsky, 1957, 1975). This point can and must be reflected on the tree. A large unit is composed of smaller units, but the relationship between units is different. A common example is subject-object asymmetry – the relationship of object with predicate verb is quite close, but the relationship between subject and predicate verb is not (Xu, 1988, p.26, 55). The hierarchy highlights the semantic closeness of constituents and the sequence of merge in a tree. When labeling the category for the phrase “against government plans”, some students ignore its hierarchy, as (8a) shows:



“government” and “plans” are in a modifier-head relationship in (8a), and therefore it is right to put the two nodes “N” and “N” in the same level. But putting the three nodes “P”, “N” and “N” in the same level is wrong, because their syntactic positions are different. In the phrase “against government plans”, the preposition “against” is the head, “government plans” acts as the complement of “against”, and “against” merges with the overall complement “government plans”. Therefore, P should be in the same level with NP in the tree. A right practice is to elevate P to the same level as NP, as (8b) shows, demonstrating the structural hierarchy of the phrase and the aesthetics of tree diagram as well. An informal requirement for such hierarchy is to put two nodes only at each level.

### D. Confused Relationship between Intermediate and Maximal Projections

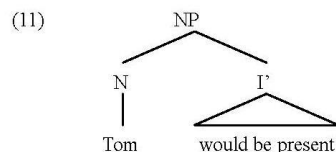
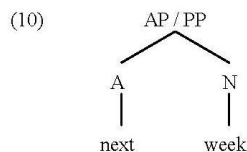
In syntax, the projection higher than lexical categories and lower than maximal projection is called intermediate projection. It is not an ultimate category, that is to say, there are other constituents which have not been merged yet. They are usually written like N', N'', V', V'' with quotation marks. The maximal projection, also called phrasal projection, is a full phrase, with no more constituents to be merged, e.g. NP and VP. Therefore, in any two adjacent levels in the projection system, an intermediate projection must be below a maximal projection. There also exists above-below relationship within an intermediate projection. For example, V' and V'' must be below VP, N'' must be above N' and so on. It's found that some students make mistakes as (9) below:



(9) confuses the relationship between an intermediate projection and a maximal projection. V' must be the intermediate projection, and VP the maximal projection. The intermediate projection V' must be below the maximal projection VP. The right practice is exchanging the positions of V' and VP in (9), with V' merging with NP, projecting into VP, a phrasal category.

### E. Projection from Incorrect Heads

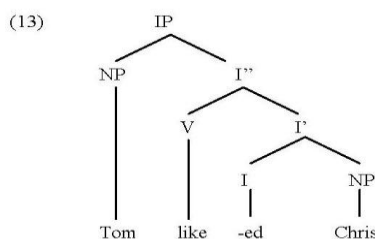
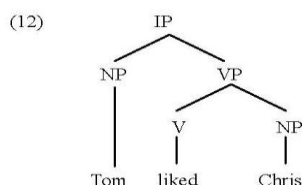
In syntax, any phrasal category comes from projection of its head, and therefore recognizing the head correctly is crucial to determine a phrasal category. Some students fail to locate the head correctly. As a result, the phrasal category has nothing to do with any lexical category in the tree, as (10) and (11) show.



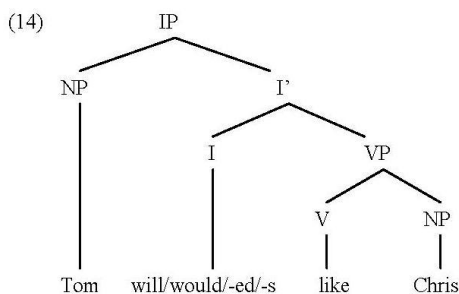
In (10), marking “next week” as AP or PP is wrong. A phrasal category must be marked according to its head. If the head of a phrasal category is N, the phrasal category must be an NP. If it is A, the phrasal category must be an AP. For more information concerning heads in grammatical theory, see Zwicky (1985), Hudson (197) and Corbett et al. (1993). In short, the phrasal category with a head X must be identified as XP that is a fundamental point of view of X-bar theory. Pedagogical grammar tells us that in the phrase “next week”, the head is obviously “week”, and its lexical category is N, so the phrase must be identified as NP, not AP. It cannot be PP either, because neither of the lexical category “next” and “week” is a preposition P. (11) regards *Tom* as head and labels the overall sentence as NP, which is obviously wrong. In pedagogical grammar, the sentence *Tom would be present* can never be a noun phrase NP. Secondly, it does not locate the head of the sentence correctly. In syntax, the head of a declarative sentence must be an I. For (11), *would* must be the head. So the overall sentence should be marked as IP, the maximal projection in the (11), where NP should be revised to IP.

#### F. No Separation of Inflectional Morpheme from Verb and / or Prepositioning of Infl

With regard to sentences such as *Tom liked Chris* in which a verb is attached with an inflectional morpheme, some students do not separate the verb from its inflectional morpheme when they represent the tree derivation, as (12) shows, or separate the verb from its inflectional morpheme, but the order is not right, as (13) shows.



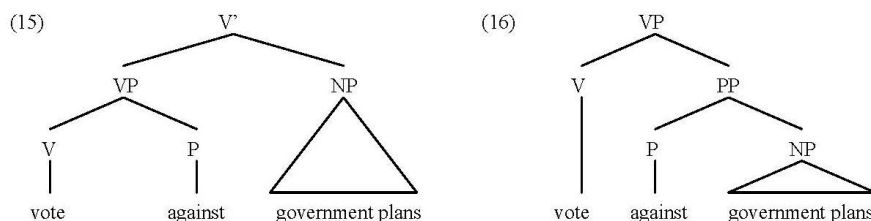
Here we focus on the separation of a verb from its inflectional morpheme and their linear order. In a tree, the reason for the separation of a verb from its inflectional morpheme and put the inflectional morpheme in front of the verb is for the purpose of uniformity in drawing tree diagrams. We know that use of I is based on syntactic analysis of complete or typical English sentence, such as *Tom will like Chris*, i.e. S → NP AUX VP. In syntax, different from pedagogical grammar, the verb “like” is not the head of a sentence. Instead, the auxiliary “will” becomes the head, because if we transform this sentence into its negative and interrogative counterparts, we have to rely on the auxiliary, and the tense is also reflected in the auxiliary. The phrase marker of the sentence is shown in (14) below.



“liked” is one of the inflectional forms of “like”, with an inflectional morpheme *-ed*. In order to unify analysis, in the tree, we separate all the inflectional morphemes from the verb, and place them in front of the verb. The relative position of inflectional morphemes *-ed* and *-s* to *like* is the same as *will* to *like*. The tree of *Tom liked Chris* is also shown in (14).

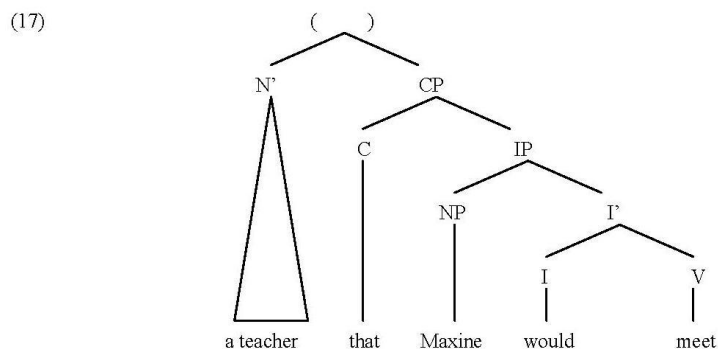
#### G. Wrong Sequence of Merge in Embedded Phrases

For such embedded phrases as “vote against government plans” where a PP is embedded in a VP, some students tend to, following the practice of pedagogical grammar, view “vote against” as a VP like “look after”, by merging “vote” and “against” first and then merging the resulting construction with the NP “government plans”. This leads to a wrong tree diagram as shown in (15) below.



In the first place, a mistake indicated in subsection D above also exists in (15), that is, confused hierarchy between the intermediate projection  $V'$  and the maximal projection VP. Here we focus on the sequence of merge in embedded phrases. In tree diagrams it is not acceptable to merge "vote" and "against" first. It is known that if two constituents can be merged, they must have intimate syntactic and semantic relation – head-complement relation or modifier-head relation, expressing a relatively complete semantic interpretation following merge. In the VP "vote against government plans", "against" is not the complement to "vote". Hence, they cannot be merged. Instead, the complement of "vote" is the PP "against government plans" in which "government plans" and "against" shall be merged first as the former is complement of the latter. Following such analysis, a correct tree is shown in (16) which clearly shows the head-complement relationship between the constituents of the phrase.

In addition, for such phrases embedded with a relative clause as "a teacher that Maxine would meet", some students, again influenced by pedagogical grammar, tend to hold the view that a clause must be larger than a sentence and a VP must be larger than an NP. This leads to a wrong drawing of tree, as shown in (17) below.

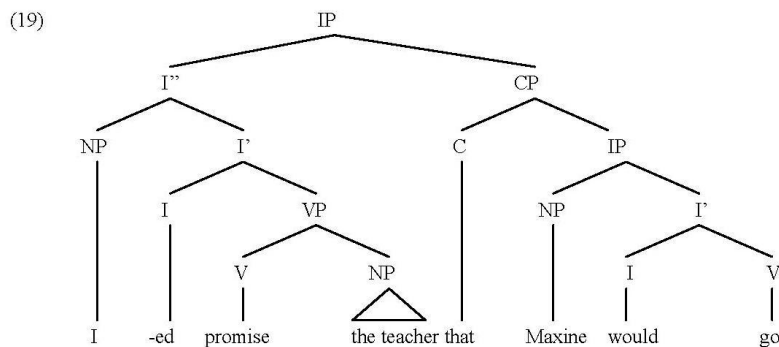
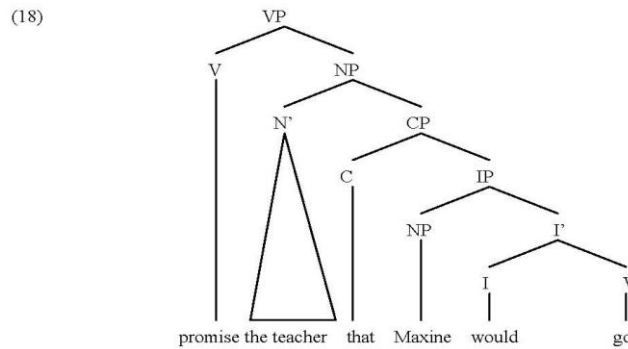


Some students mark the maximal projection, i.e. the place with an empty bracket, as CP, perhaps because they see there is a CP below already, or make no mark at all, perhaps because they are afraid there would be no unit larger than CP or IP. Pedagogical grammar tells us that "a teacher that Maxine would meet" is an NP. What makes it complex is the relative clause following "a teacher". One of the basic tenets of syntax is that language is recursive (see Chomsky, 1957, 1975; Wang, 2008a, 2008b). Clauses can be embedded, so a sentence can be larger than a clause and an NP can be larger than a VP or even a clause. (17) as a whole is an NP which contains an  $N'$ , i.e. "a teacher" and a CP, i.e. "that Maxine would meet". Filling in the blank bracket with an "NP" would make the tree complete and correct.

#### H. Misrepresentation of Different Clauses

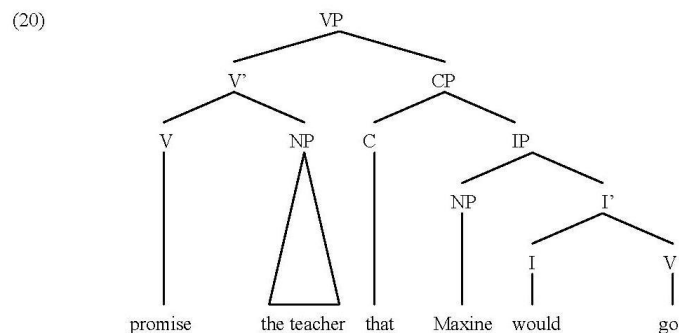
In pedagogical grammar, there are differences among subjects, relative clauses and complement clauses in terms of form and meaning. It is not easy to reflect such differences in a tree. How to represent relative clauses in a tree has already been discussed in the previous subsection, so here we turn to the representation of complement clauses as in "I promised the teacher that Maxine would go", where "the teacher" is an object and *that*-clause is its complement. However, some students are found to mistake this complement clause for a relative clause, as shown in (18) (with I and -ed omitted).

It is wrong to view the *that*-clause as a relative clause and to directly merge it and "the teacher" into an NP. The result of such merge is incomprehensible as "the teacher that Maxine would go" cannot constitute any relatively independent and complete semantic interpretation. A similar mistake shown in (19) is also found.



Such a scheme seems perfect at the first sight, but it is wrong indeed, as obviously the main clause “I promised the teacher” is not in the same syntactic position as the subordinate clause “that Maxine would go” (the complement of “promised”). They are not in a head-complement relationship, hence it is impossible for them to be merged in the same level. Semantic interpretation following such merge is incomprehensible, neither. In (19), the *that*-clause is not like a relative clause, nor a complement clause, in fact.

The right comprehension is that “promise” first merges with “the teacher” and then with *-ed*. (20) shows the right representation (with *I* and *-ed* omitted).

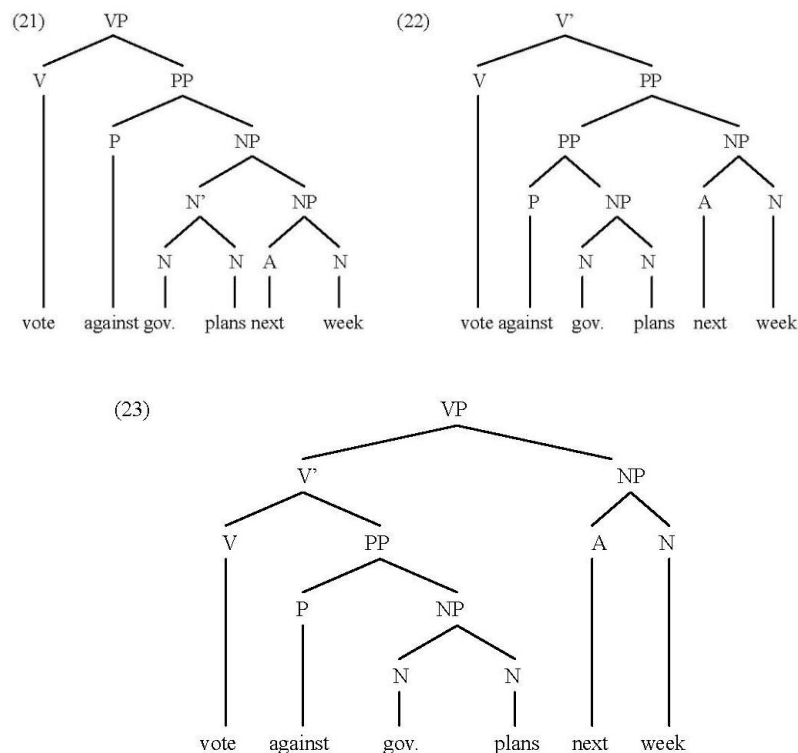


Here, it is reflected that “the teacher” serves as the object of “promise” and “that Maxine would go” as the complement. Semantic interpretation following such merge is comprehensible.

### 1. Mistreatment of Adverbials

It is known that an adverbial can modify a VP or a sentence, but generally not a NP or PP only. For example, in the sentence “They will vote against government plans next week”, in terms of meaning the adverbial “next week” modify “vote against government plans” or “they will vote against government plans”, but not “against government plans” or “government plans” only. Some students are found to make wrong representations such as (21) and (22) (with “they will” omitted).

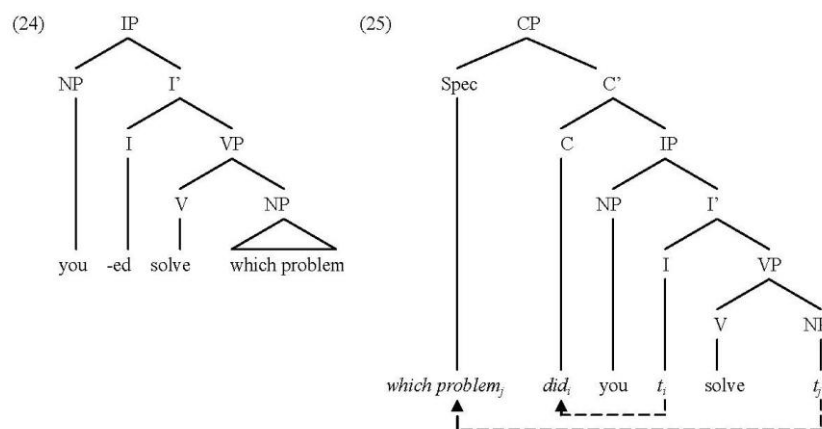
Here we focus on how to deal with adverbials in a tree. (21) is wrong in that the NP “government plans” is directly merged with the adverbial “next week”, which means that “next week” modifies “government plans”. This must lead to semantic incomprehension for lack of semantic basis. The sentence “They will vote against government plans next week” should be partly paraphrased as “vote against government plans when next week comes” but not “vote against government plans to be made next week”. Therefore, “government plans” should not be directly merged with “next week” as they are not in a modifier-head relationship. The right representation is shown in (23) (with “they will” omitted).



This clearly demonstrates the syntactic relation that the adverbial “next week” modifies the VP “vote against government plans”, and correct interpretation of the sentence.

#### J. Misrepresentation of WH-questions with Constituent Movement

Constituent movement may happen in syntax of WH-questions. For example, in “Which problem did you solve?”, the WH-word “which problem” is located behind “solve” as its complement in deep structure, and then moves to the beginning of the sentence, i.e. [Spec, CP], in surface structure. It is also possible that, not strictly speaking, there is no constituent movement in syntax of WH-questions. For example, in “Who taught you?”, the WH-word “who” is located at the beginning of both deep structure and surface structure, without any overt movement. Some students find it difficult to represent WH-questions with constituent movement. Pedagogical grammar tells us that WH-questions are derived from declaratives via yes-no questions with the addition of WH-words. Therefore, we shall start from declaratives to draw trees of WH-questions step by step. For this example, we shall start from characterizing its deep structure “[you [-ed [solve [which problem]]]]”, in which the WH-word “which problem” is located behind “solve” as its complement due to the subcategorization requirement of the verb “solve”, as shown in (24).



First, transform the structure into an interrogative in which -ed, the inflectional morpheme, is moved to a place in front of “you” as a complementizer (C) and changed to “did” in the course of morphological realization as -ed is not a word. Next, move the WH-word “which problem” to a place in front of “did” as a specifier (Spec). The traces for such two movements are shown as  $t_i$  and  $t_j$  in (25). Then, mark the indexes for the moved constituents and their traces in pairs

to show their co-indexation, and indicate the routes of movement with arrowheaded dotted lines. Representation of yes-no question is also shown in the lower part under C' in (25).

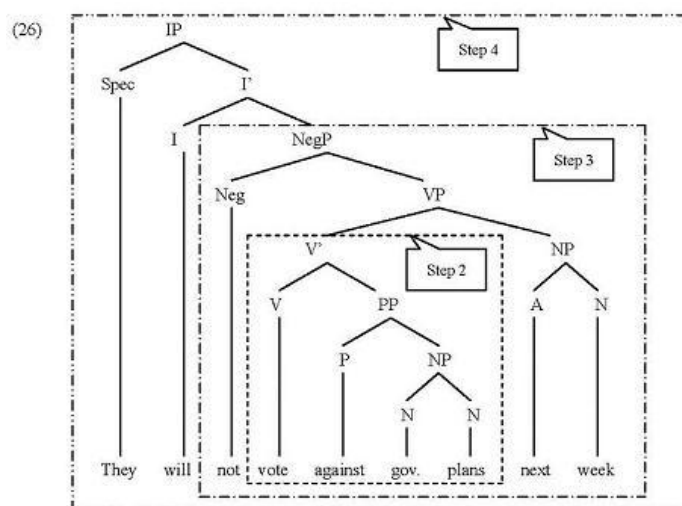
### III. SUGGESTIONS

After analysis of common problems found in the author's teaching practice of syntax and review of exam papers, and now taking a typical English sentence "They will not vote against government plans next week" as an example, we summarize a general idea and steps of drawing tree diagrams as follows.

General idea: Step by step merge phrases from right to left and from bottom to up while maintaining hierarchy, and mark appropriate lexical and phrasal categories until we merge all the constituents and get the maximal projection.

Steps: First, determine the semantic closeness between constituents (i.e. words), and figure out the subject-predicate-object and attributive-adverbial-complement of the phrase or sentence. Next, represent structures of the predicate verb based on head-complement or modifier-head relationship. Then, merge the verb phrase with Neg, adverbial, and complement (if any). Finally, merge the inflectional word I with subject. Consider (26). Step 2 to 4 are shown in frames with different dotted lines.

We know that in this sentence, "they" is the subject (Spec); "will", the auxiliary, is inflection I; "not" is Neg; "vote against government plans" is verb phrase VP; "next week" is adverbial. From right to left, by means of step-by-step merge, we merge "vote against government plans", then merge with the adverbial "next week" into VP, then merge with negative word "not" into NegP, and finally merge with the inflection "will" and the subject (Spec) "they" until we get the maximal projection IP. The second to fourth step can be seen in the figure. Among the four steps of representation, the second step is the key process. Merging the predicate verb phrase often determines whether the final trees are right or not. Many of the problems discussed above are caused by the incorrect merging of the verb phrase. In ordinary circumstances, it requests us to merge phrases from right to left. For more complex sentences, we also need to consider whether there are other special situations to be dealt with.



### IV. CONCLUSION

The article analyses ten common problems in tree diagram drawings and propose a general idea and steps of such representations. Tree diagram is an effective tool to characterize the internal hierarchical structure of natural language. Drawing tree diagrams is a basic skill of syntax studies. It is based on not only the pedagogical grammar's understanding and new thought of syntax as well. We are sure that if we master the methods and frequently practice, we must be able to draw reasonable and elegant tree diagrams in the near future.

### ACKNOWLEDGMENT

This research is sponsored by "Building an English Major with IT Characteristics against the Background of Globalization", a major promotion project of Chongqing University of Posts and Telecommunications 2008, and "A Study of Interface between Syntax and Computational Linguistics", a planned project of philosophy and social science of Chongqing Municipality 2009 (Grant No.: SKZ0909). I am indebted to Fuzhen Si, Eric Chen, Liejiong Xu, Binli Wen, Alexander Grosu, Andrew Radford, and Alec Marantz for helpful comment on earlier drafts of the article.

### REFERENCES



- [1] Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- [2] Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- [3] Chomsky, N. (1975). *The Logical Structure of Linguistic Theory*. Chicago: The University of Chicago Press.
- [4] Corbett et al. (1993). *Heads in Grammatical Theory*. Cambridge: Cambridge University Press.
- [5] Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics* (sixth edition). Oxford: Blackwell.
- [6] Dai, W. & He, Z. (2002). *A New Concise Course on Linguistics for Students of English*. Shanghai: Shanghai Foreign Language Education Press.
- [7] Gleason, H. (1961). *An Introduction to Descriptive Linguistics*. New York: Holt, Rinehart and Winston.
- [8] Hu, Z. (2001). *Linguistics: A Course*. Beijing: Peking University Press.
- [9] Hudson, A. (1987). Zwicky on Heads. *Journal of Linguistics*, 1.
- [10] Jackendoff, R. (1977). *X' Syntax: A Study of Phrase Structure*. Cambridge, Mass: MIT Press.
- [11] Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: The MIT Press.
- [12] Liu, R. & Wen, X. (2006). *Linguistics: A New Coursebook*. Beijing: Foreign Language Teaching and Research Press.
- [13] Ouhalla, J. (1999). *Introducing Transformational Grammar: From Principles and Parameters to Minimalism*. London: Edward Arnold (Publishers) Limited.
- [14] Poole, S. (1999). *An Introduction to Linguistics*. London: Macmillan.
- [15] Radford, A. (1988). *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
- [16] Wang, Q. (2008a). Grammar is Generative rather than Integrative. *Journal of Tianjin Foreign Studies University*, 1.
- [17] Wang, Q. (2008b). Reinterpreting Syntactic Structures (Chomsky 1957) – To Celebrate the 50<sup>th</sup> Anniversary of Syntactic Structures. *Journal of Guangdong University of Foreign Studies*, 3.
- [18] Xiong, Z. (2009). Phrase Structure Theory and Computational Operations in the Generative Grammar. *Journal of Tianjin Foreign Studies University*, 5.
- [19] Xu, J. (2007). A Comparative Study of Categories in Generative Grammar and Parts of Speech in Traditional Grammar. *Research on Chinese as a Second Language*, 3.
- [20] Xu, L. (1988). *Theory of Generative Grammar*. Shanghai: Shanghai Foreign Language Education Press.
- [21] Zwicky, A. (1985). Heads. *Journal of Linguistics*, 1.

**Qiang Wang** was born in Hunan, China in 1979. He received his MA degree in linguistics from Sichuan International Studies University, China in 2007.

He is currently a lecturer in the College of Foreign Languages, Chongqing University of Posts and Telecommunications in southwest China. His research interests include syntax and biolinguistics.