TBTA Tutorial Lesson 4:
Spellout Rules and Phrase Structure Rules

([video: TBTA-4A Spellout Rules](#))
([video: TBTA-4B Spellout Rules and Phrase Structure Rules](#))
([video: TBTA-4C Table Spellout Rules](#))

4.0 Introduction

In the previous lesson you created a database for your tutorial language, and you entered target stems into several syntactic categories. You also specified features and forms for the nouns and verbs. The forms were generated by rules which you wrote. Those rules are called Lexical Spellout rules because they generate (spell out) target forms in the lexicon.

In this lesson we're going to look at two types of rules in TBTA's synthesizing grammar[1]. The first type of rule is almost identical to the Lexical Spellout rules that you developed in the previous lesson, but these rules are slightly different because they're in the synthesizing grammar rather than the lexicon. To see TBTA's grammar, go to the main menu and choose Generator, Lexicon and Grammar Development. This is the screen where you'll do the vast majority of your work. At this screen you see one verse; its reference is in the upper left corner. If you click on the Reference button, you'll see a dialog that will let you select any verse from any of the analyzed source texts (semantic representations). The grammar tree is in the upper right corner of this screen and is shown below in Figure 4-1. As you can see in the tree, the grammar consists of two major components: the transfer grammar, and the synthesizing grammar. Linguists have been developing synthesizing grammars for decades, so we'll begin by learning how to use the rules in that grammar first. Subsequent tutorials will discuss the various rules in the transfer grammar.

There are eight types of rules in the synthesizing grammar: 1) Feature Copying rules, 2) Spellout rules, 3) Clitic rules, 4) Constituent Movement rules, 5) Phrase Structure rules, 6) Pronoun and Switch Reference rules, 7) Word Morphophonemic rules, and 8) Find/Replace rules. This lesson will focus on the Spellout rules and the Phrase Structure rules. However, we're first going to make a brief excursion so that you can link concepts in the analyzed source texts to target language entries.

---

[1] Remember that TBTA has three grammars: 1) the Lexical grammar, which you studied in the previous tutorial, 2) the Synthesizing grammar which performs the synthesis step of the translation process, and 3) the Transfer grammar which performs the transfer step of the translation process.
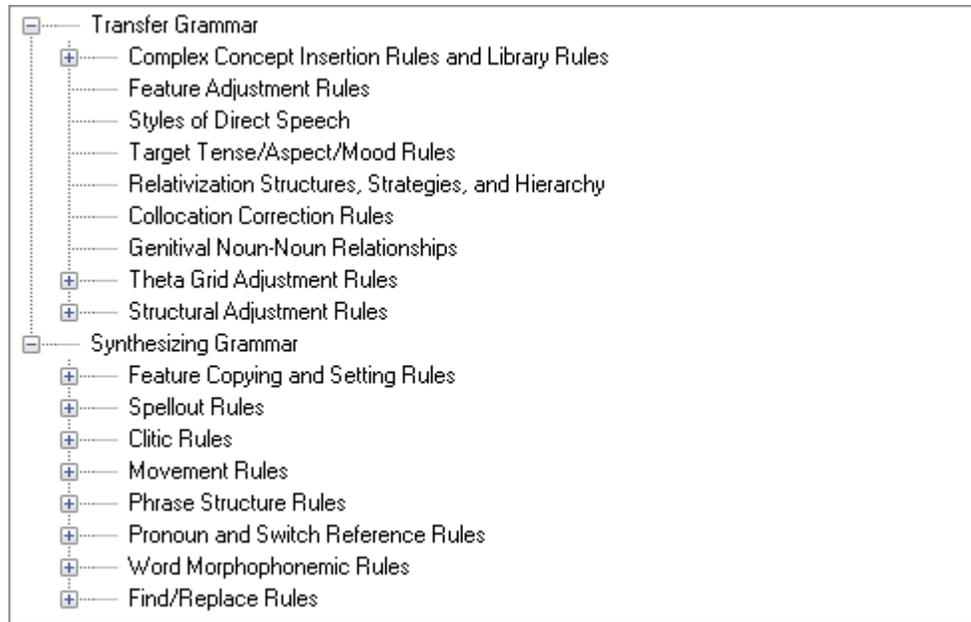
```
├─ Transfer Grammar
│    ├─ Complex Concept Insertion Rules and Library Rules
│    ├─ Feature Adjustment Rules
│    ├─ Styles of Direct Speech
│    ├─ Target Tense/Aspect/Mood Rules
│    ├─ Relativization Structures, Strategies, and Hierarchy
│    ├─ Collocation Correction Rules
│    ├─ Genitival Noun-Noun Relationships
│    ├─ Theta Grid Adjustment Rules
│    └─ Structural Adjustment Rules
├─ Synthesizing Grammar
│    ├─ Feature Copying and Setting Rules
│    ├─ Spellout Rules
│    ├─ Clitic Rules
│    ├─ Movement Rules
│    ├─ Phrase Structure Rules
│    ├─ Pronoun and Switch Reference Rules
│    ├─ Word Morphophonemic Rules
│    └─ Find/Replace Rules
```

Figure 4-1. The Grammar Tree

4.1 Linking Concepts to Target Words

First make sure that the currently selected verse is Nouns 1:1. If this isn't the current verse, select it by clicking on the Reference button in the upper left corner, choose English Grammar Introduction as the source text, and then select Nouns 1:1. There are two concepts in this verse, MAN-A and WALK-A. Neither of these concepts are currently linked to target entries, so '???' is displayed below both concepts. To link a concept to a target entry, right click on the concept. You'll see a popup menu appear, and the first menu option is 'Map this Concept to Target Words'. Rest the cursor on MAN in the large analyzed verse window, right click and select the first menu option. You'll see a dialog listing all of the target nouns that you entered in the previous tutorial. At the top of this dialog is a table with one row, and it contains the concept MAN-A. This table has a column labeled Mappings, and that cell is currently empty. This dialog lets you link the concept MAN-A to as many target entries as you'd like. Let's link the concept MAN-A to the target noun stem *man* that you entered in the previous lesson. Scroll the target table downward to *man* and click on its row. You'll see the Stems cell go white for *man*, and you'll see the row number for *man* appear up in the Mappings cell for MAN-A. If you rest your cursor on the Mappings cell, a popup will show you the target entries you've linked to that concept.

You may link as many target words or entries to each concept as you like, but TBTA will use only the first one unless you write a rule that specifies otherwise; this will be discussed in a subsequent lesson. After you've linked MAN-A to *man*, close the Mapping dialog and you'll see *man* appear below MAN-A in the analyzed source verse. Repeat this linking process so that WALK-A is mapped to *walk*. You've now linked all of the concepts in this verse to target words.

Rest the cursor on the features under MAN-A in the semantic representation. You'll see a popup explaining the meaning of the features associated with MAN-A; this popup is shown below in Figure 4-2. First the popup lists the feature values that are included with the semantic representations. These features include Number, Participant Tracking, Polarity, Proximity, etc. Then you'll see a dashed line, and below that line are the features that you've defined for your target nouns.
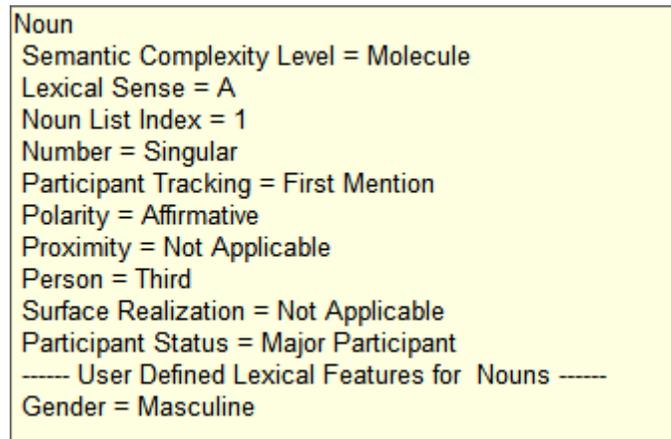
```
Noun
Semantic Complexity Level = Molecule
Lexical Sense = A
Noun List Index = 1
Number = Singular
Participant Tracking = First Mention
Polarity = Affirmative
Proximity = Not Applicable
Person = Third
Surface Realization = Not Applicable
Participant Status = Major Participant
------ User Defined Lexical Features for  Nouns ------
Gender = Masculine
```

Figure 4-2. Popup Explaining Source and Target Features for MAN-A

If you rest the cursor on the features under WALK-A, you'll see a similar popup.

Now rest the cursor on the concept WALK-A. You'll see a popup with the concept's definition, the number of times that the concept occurs in the semantic representations, the verb's theta grid and its argument structure. If you rest the cursor on the target word *man* below MAN-A, the popup will list all of the target nouns linked to the concept MAN-A. Right now you only have one target word linked to MAN-A, so you'll only see *man* listed in the popup.

4.2 Synthesizing Grammar Spellout Rules

You're now ready to write your first grammar spellout rule. If you rest your cursor on the features under MAN-A, you'll see that the Number value is Singular and the Participant Tracking value is First Mention. When Participant Tracking is First Mention, the participant is being introduced onto the discourse stage. In English when a participant is singular and first mentioned, we introduce him with the indefinite article *a*, as in *A man walked down the street*. So let's write a spellout rule that will add the Article *a* to nouns when the nominal's Number is Singular and the Participant Tracking value is First Mention. Click on the button labeled Grammar Wizard which is below the grammar tree. Grammar Wizard works the same way here that it did in the lexicon, but the template list box contains many more examples. The reason for this is because in the lexicon you're only able to write lexical spellout rules, but here Grammar Wizard will help you write all of the types of rules in the synthesizing grammar. Look at each of the rule templates listed in Grammar Wizard shown below in Figure 4-3.
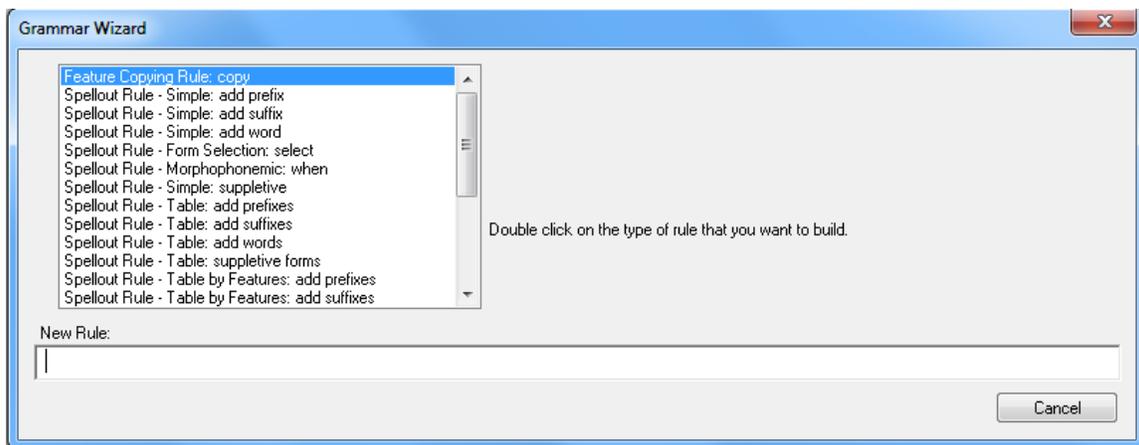
Figure 4-3. Grammar Wizard in the Synthesizing Grammar

The indefinite article *a* is a word rather than a prefix or suffix, and we want to add this word to nouns, so the fourth template is the closest to what we're looking for. Double click on that item in the list box. For the word enter ***a*** and push the Tab key. For the word's tag enter **Article** and push the Tab key. Double click on Nouns in the list box, and then you'll see a spellout rule that is almost identical to the spellout rules you wrote in the lexicon. As in the lexicon, we need to specify a group. Right now there are no groups for noun spellout rules, so push the New Group button and enter **Articles**. You'll see that a comment has automatically been entered for you in the Comment box down at the bottom of the rule. This comment is the reference of the verse where you're writing this rule, and the English translation of this verse. You should always enter comments into your rules so that later you'll know why you wrote them. We want this rule to apply only to nouns that have a Number value of Singular, and a Participant Tracking value of First Mention, so click the Features button and select Singular and First Mention. As always, we need to assign this rule to one or more topics; assign this rule to both Number and Participant Tracking. Your rule should now look like that shown in Figure 4-4.
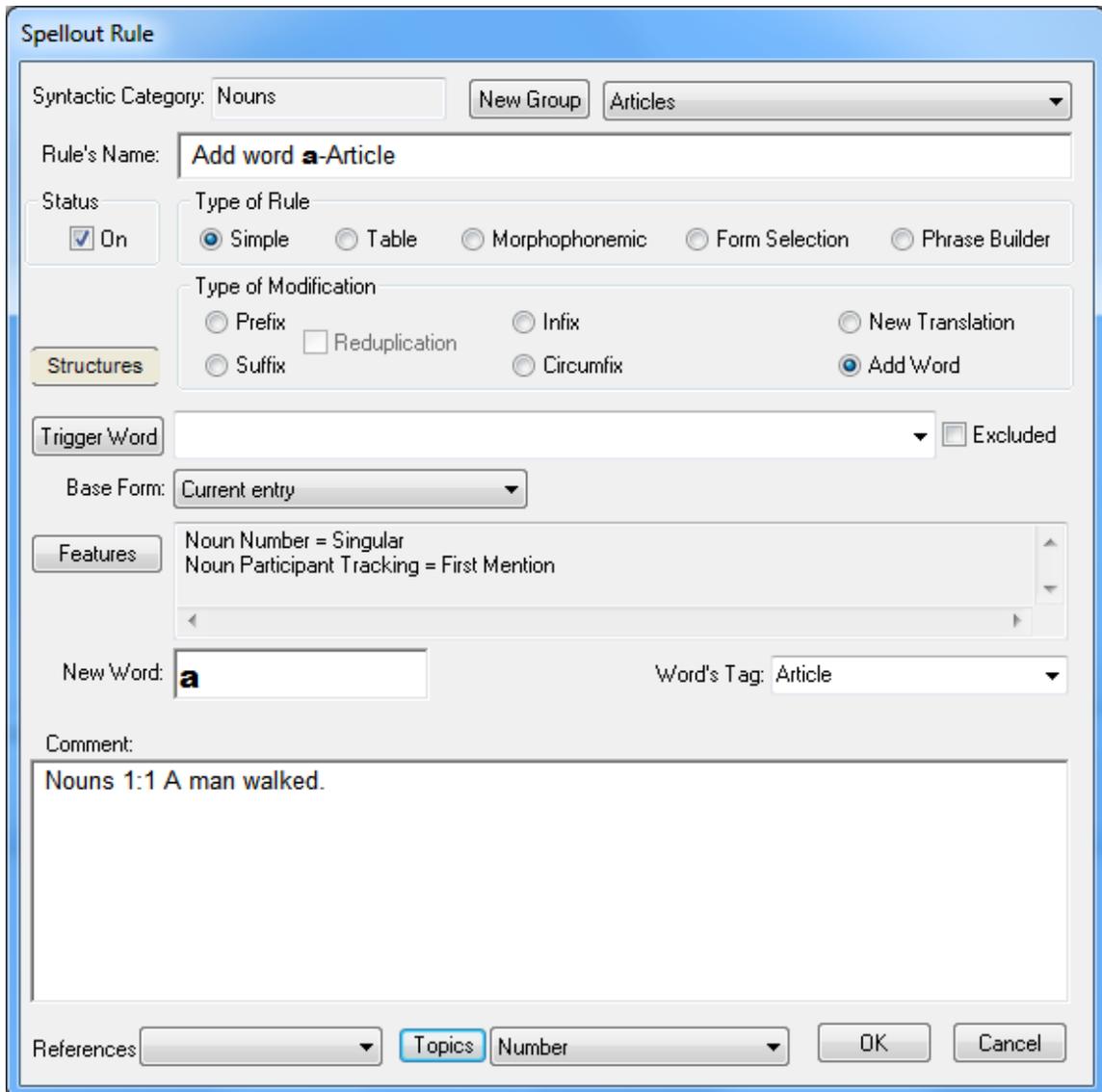
Figure 4-4. Spellout Rule for adding the Article *a*

Save the rule by clicking the OK button. The program will immediately detect that you've just inserted a new constituent into noun phrases that hasn't yet been accounted for in the first NP phrase structure rule. So it will ask you if you want to edit that rule now. Answer Yes, and you'll see the general purpose NP phrase structure rule that was automatically built by TBTA when you created your target database. When TBTA built this rule, the rule contained only a Noun. When you see the rule now, it has two constituents because your new Article has been inserted at the top of the tree in the rule. If you click on the Article, you'll see a popup showing you an example, in this case the word *a*. Drag the Article down the tree to the Noun and release it. You'll see that it's then positioned after the Noun in the tree. English articles precede their nouns, but save the rule just so that we can see the effect of putting the Article after the noun. The phrase structure rule should now look like that shown in Figure 4-5.
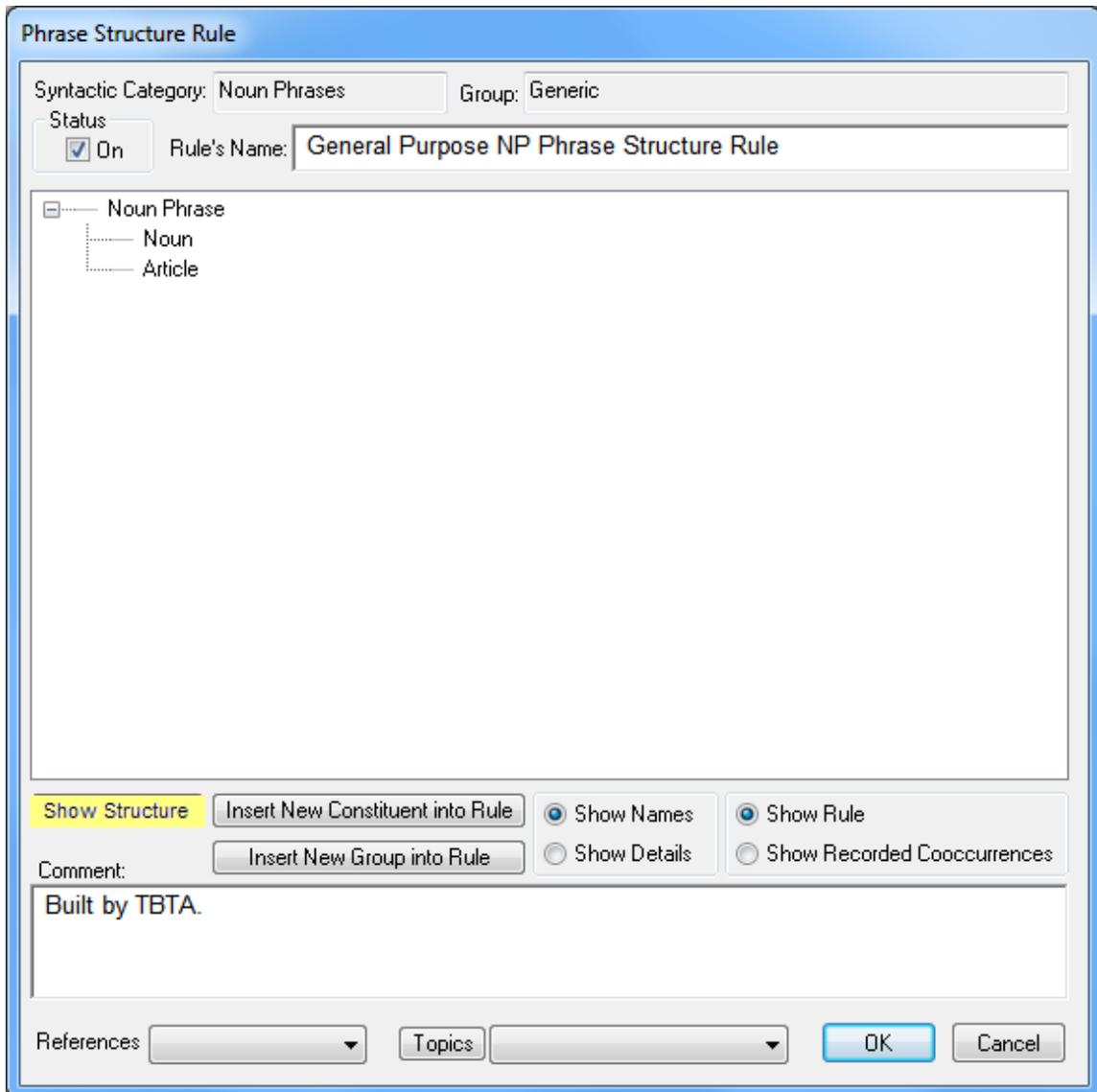
Phrase Structure Rule

Syntactic Category: Noun Phrases   Group: Generic

Status
☑ On   Rule's Name: General Purpose NP Phrase Structure Rule

⊟――― Noun Phrase
     ――― Noun
     └――― Article

Show Structure   Insert New Constituent into Rule   ◉ Show Names   ◉ Show Rule
                  Insert New Group into Rule   ○ Show Details ○ Show Recorded Cooccurrences
Comment:
Built by TBTA.

References [▼]   [Topics] [▼]   [OK] [Cancel]

Figure 4-5. NP Phrase Structure Rule with Article following Noun

Close the phrase structure rule by clicking OK.

Now let's find the spellout rule that you just entered in the grammar tree. Open the Spellout Rules node in the grammar tree, and then open the Nouns node. You'll see the group called 'Articles' that you just added; open its node and you'll see the name of the rule you just built. Double click on the rule's name and the rule will be displayed. As mentioned earlier, this rule looks virtually identical to the spellout rules you entered in the target lexicon during the previous lesson. The primary difference between these spellout rules and the lexical spellout rules is that the lexical spellout rules generate forms in the lexicon; the spellout rules here in the grammar are executed only when you push the Generate button in the upper left corner of the screen. Another significant difference between the lexical spellout rules and the grammar spellout rules is the features that are involved. Click the Features button on the spellout rule

dialog to look at the features you've specified for this rule that inserts *a*. When you specify features for rules here in the grammar, you'll see the noun features that are included with the semantic representations in addition to the features that you defined for nouns in the target lexicon. So you'll see columns for Number, Participant Tracking, Polarity, etc. Scroll the features to the right and you'll see the Gender feature that you defined in the previous lesson. The features that you defined in the lexicon are displayed with a yellow background. When you entered lexical spellout rules in the previous lesson, those rules looked at only the features that you had defined in the lexicon for that syntactic category; those rules don't include features from the semantic representations such as Number, Participant Tracking, etc. Spellout rules here in the grammar look at the features in the semantic representations as well as the target features you've defined for that syntactic category. This distinction will become clearer in subsequent lessons.

In the lower left corner of the features dialog are three buttons labeled Word, Phrase and Clause; these buttons weren't present when you were writing lexical spellout rules. Right now the Word button is depressed indicating that you're looking at word features. If you click the Phrase button, you'll see the features for noun phrases because nouns are always in noun phrases. If you click the Clause button, you'll see the features for clauses. None of these other features are relevant for this rule, so you can close this features dialog. However, suppose you wanted to specify that the article *a* should only be inserted in imperatives. You could do that by clicking the Clause button on the features dialog, and then selecting the Illocutionary Force value of Imperative. So in this features dialog you may specify word features, phrase features, and clause features. Now close the rule dialog by clicking the Cancel button. In general, when you look at a rule but don't change anything in the rule, it's a good idea to close the rule by clicking the Cancel button rather than the OK button. Whenever you click the rule's OK button, the rule will be saved even if you didn't intentionally make any changes to it. After closing the rule, let's test it by clicking the Generate button in the upper left corner of the screen.

In the small window below the Generate button you'll immediately see *Walk man a*. In the large window you'll see that the Article *a* has been inserted into the noun phrase with *man*. If you rest your cursor on *a*, you'll see a popup showing the name of the rule that inserted it and also the two[2] phrase structure rules that positioned it. If you right click, the popup menu will now have the option 'Display List of Rules that applied during Generation Process'. Select that menu option and a dialog will appear at the bottom of the screen with these three rules listed in it; this dialog is shown below in Figure 4-6. In the Type of Rule column, if you double click on 'Spellout', you'll be able to see and edit your rule that inserted the Article.

---

[2] There are four phrase structure rules listed in the popup, but that's because the phrase structure rules are executed twice for reasons which will be explained later. So even though you see four phrase structure rules in the popup, there are only two phrase structure rules that applied.
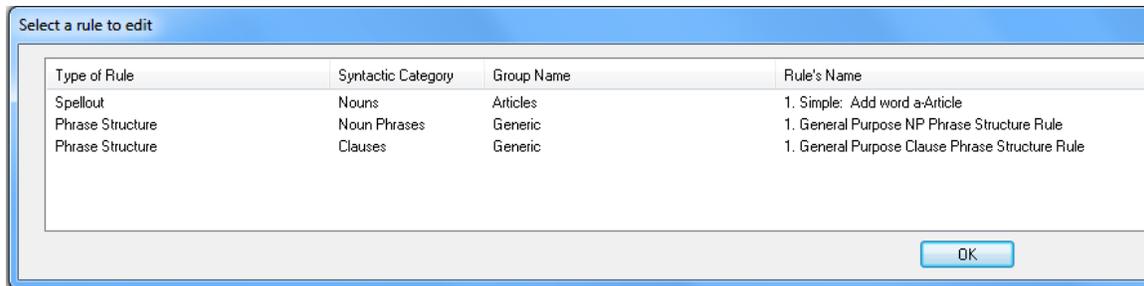
Figure 4-6. Selecting a Rule to Edit after a verse has been Generated

In general, after a verse has been generated, you can rest the cursor on a constituent in the large window, and the popup will list all of the rules that affected that constituent during the generation process. This makes it very easy for you to determine which rules are affecting the constituent, and, if one of the rules isn't working the way you intended, you can easily edit that particular rule. In this case let's edit the Phrase Structure rule for Noun Phrases. So double click on the first 'Phrase Structure' in the Type of Rule column, and you'll see the NP phrase structure rule. In English we want the Article to precede the Noun, so drag the Article up to the Noun and release it; you'll then see it appear in the tree above the Noun. We've intentionally edited this rule, so close the rule dialog by pushing the OK button. Now close the Select a Rule to Edit dialog, push the Initialize button in the upper left corner of the screen, and then push the Generate button again. This time you'll see the Article appear before the noun, just as you specified in the NP phrase structure rule.

4.3 Phrase Structure Rules

The phrase structure rules always determine the final order of the constituents in the target sentences. When you began this language project, TBTA built some basic phrase structure rules for you. The phrase structure rule that you just finished editing is one of those rules.

The text being generated right now is *Walk a man*. In English we want the verb phrase to follow the subject NP. We haven't yet written rules to generate grammatical relations from semantic roles, so for now we'll edit the clause phrase structure rule and specify that NPs with a Semantic Role of Agent precede the VP. Expand the Phrase Structure Rules node in the grammar tree, and then expand the Clauses node. You'll see one group called Generic, and it has one rule called General Purpose Clause Phrase Structure Rule. Both this rule group and the rule were built automatically by TBTA when you created your target database. Open this phrase structure rule and you'll see just a VP in the clause. To insert another constituent into the rule, click on the button labeled 'Insert New Constituent into Rule'. A new dialog will appear with a Constituent dropdown at the top. Select Noun Phrases in the dropdown, and you'll then see the features that are associated with Noun Phrases. Select the Most Agent-like value in the Semantic Role column, and then close the dialog. You'll be asked if you want to enter a name; you can answer No. You'll then see 'NP – Most Agent-like' added to the rule at the top of the tree. Since we want the Agent NP to precede the VP,

we'll leave it there at the top. Save this rule by clicking the OK button. Then click on the Initialize button at the top of the screen to initialize the verse, and then push the Generate button. The target text is now *A man walk*. So you can see that the clause phrase structure rule positioned the NP-Agent before the VP just as you instructed.

4.4 More about Synthesizing Grammar Spellout Rules

If you rest the cursor on the features under WALK in the semantic representation, you'll see that the Time value is Discourse. English speakers tell narratives using the past tense. In the previous lesson we generated a lexical form for verbs called Past Tense, and we wrote several rules to generate the past tense form from the verb stems. Those rules generated a list of past tense verbs, but the words in that list aren't linked in any way to the semantic representations. Similarly for nouns we wrote rules in the previous lesson to generate the plural form from the stems, but again, the words generated by those rules are in a list in the lexicon. In order to use the words in those lexical lists, we have to write rules here in the grammar that instruct TBTA when to use a word from a particular lexical list. For example, we want to use the Past Tense form for each verb whenever a verb is tagged with a Time value of Discourse. TBTA will always use the stem for each target word unless we write a rule in the grammar instructing it to use another form of that word in a particular environment. So in order to select the Past Tense form from the lexical list, we need to write another spellout rule for verbs here in the grammar tree. We want to write a rule that will select the Past Tense form of a verb from the lexicon if a verb in the semantic representation has a Time value of Discourse. Look at the templates in Grammar Wizard and you'll see that the fifth template selects forms from the lexicon. Double click on the fifth template, and then double click on 'Past Tense form for verbs.' You'll then see a spellout rule, and its Type is Form Selection. We need to specify a group so click the New Group button. Enter a group name such as **Tense**. For the features, specify a Time value of Discourse. Assign this rule to the Tense topic and click OK. Your final rule should look like the one shown in Figure 4-7.

Figure 4-7. Spellout Rule to select the Past Tense form from the Lexicon

After you've entered this rule, close the rule dialog, click the Initialize button, and then click the Generate button. The target text is now *A man walked*. If you rest the cursor on the word *walked* in the large window, the popup will tell you that it was changed from *walk* to *walked* by the rule that you just entered.

Let's save the result of this verse by clicking the Save Verse button at the top of the screen, and then click Next Verse. The reference is now Nouns 1:2. This verse looks identical to the previous verse, but one of the features on the noun MAN-A is different. If you rest the cursor on the noun features, you'll see that the noun Number is Plural. You can always see the English text generated by a particular verse by looking in the small window directly beneath the grammar tree. For this verse, that small window says, "Nouns 1:2 Some men walked." Since this noun is tagged as Plural and First Mention, this verse should generate *Some men walked*. If you click the Generate button right now, the

result will be *Man walked*. The reason for this is because the noun's Participant Tracking value is again First Mention, but the rule which we wrote to insert the Article *a* specified that the Noun's Number value must be Singular. Since this noun's Number value is Plural, that rule won't apply to this noun.

In order to generate *Some men walked*, we need to write two rules; one rule will select the Plural form of a noun from the lexicon when a noun is tagged as Plural in the semantic representations, and another rule will add the Article *some* when a plural noun is tagged as First Mention. Use Grammar Wizard and write these two rules yourself. The first rule which selects the Plural form from the lexicon will be very similar to the rule we wrote above which selected the Past Tense form of a verb from the lexicon. Put this rule into a new group for Nouns that you call **Number**. When writing this rule, be sure to include the Number values Dual, Trial, Quadrial, Paucal, and Plural because English will use the plural form for all of these values. Then write another rule which will insert the Article *some* whenever a plural noun is tagged with First Mention. Put that rule in the group called Articles with your other rule that adds *a* to singular first mention nouns. After you've written these two rules, click the Generate button, and you should see *Some men walked*.

Now turn off the two simple spellout rules which insert the Articles *a* and *some*, and then use Grammar Wizard to write a table spellout rule[3] that will accomplish the same thing. That table spellout rule will add a word to nouns, and it should have one row and two columns. The tag for this rule should again be Article. To specify features for a column, double click on the column's header cell. For the first column, specify Singular and First Mention, and give it a name such as 'Singular First Mention'; for the second column, specify Plural and First Mention, and give it a name such as 'Plural First Mention.' Remember that you can make columns wider by resting the cursor at the right edge of a column header, and then dragging the column's edge to the desired width. It's helpful to enter comments for each cell in a table. To enter a comment for a particular cell, right click on the cell. You'll then see that the cursor is in the comment box at the bottom of the rule. Enter a comment such as 'Nouns 1:1 A man walked.' for the first cell. After you've finished entering the comment, push the Enter key. The comment will then be cleared from the comment box, but if you rest the cursor on the cell, you'll see a popup showing your comment. You should try to enter comments for every cell in a table spellout rule. To check your table rule, generate both Nouns 1:1 and Nouns 1:2 again to see if both verses are being generated correctly. After you've generated Nouns 1:2 correctly, be sure to save the target text by clicking the Save Verse button.

Go to Nouns 1:3 and rest the cursor on the features under MAN-A. You'll see that the Participant Tracking is Routine. Look in the window below the grammar tree and you'll see that this verse produces *The man walked*. Modify your table spellout rule for nouns that inserts Articles so that it inserts *the* whenever a noun's Participant Tracking value is Routine. Generate this verse,

---

[3] The process of building table spellout rules was described in Exercise Set 3. If you skipped those exercises, you should skim through them now to learn how to build table spellout rules.

and when you get the right results, click the Save Verse button. At this point your table rule for Articles should look something like that shown in Figure 4-8.
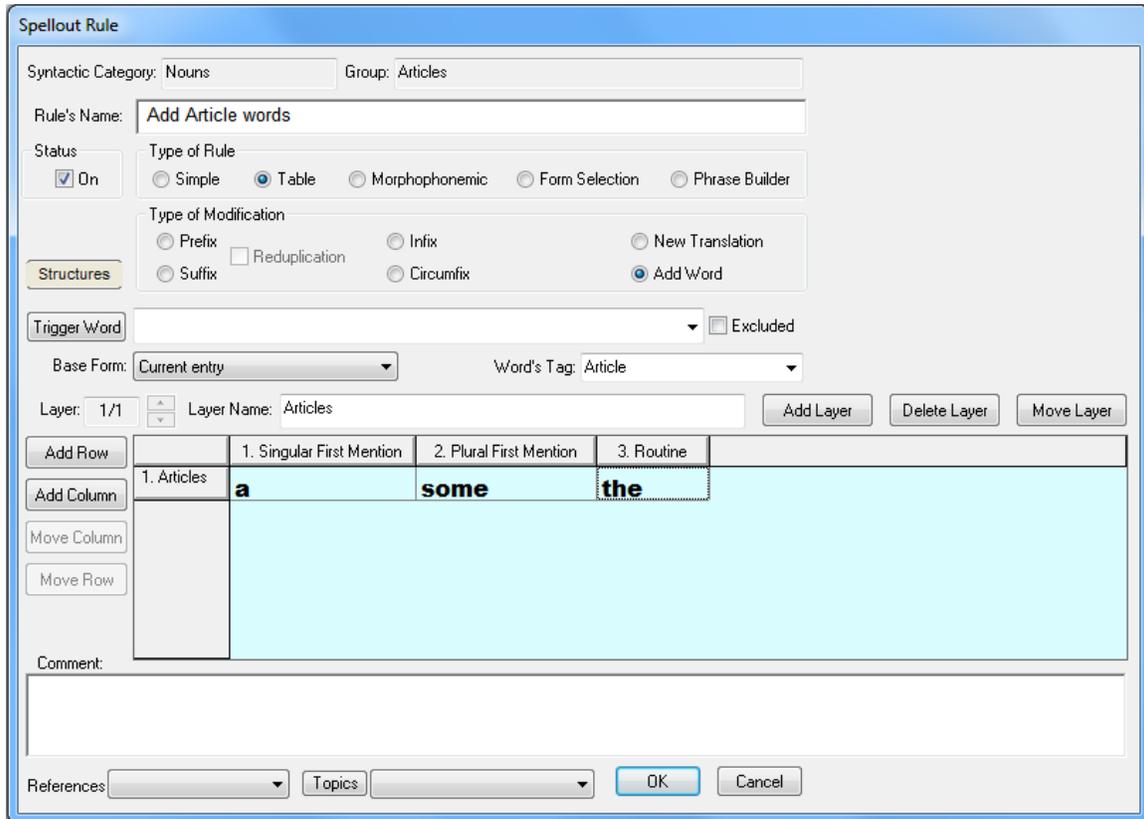

Figure 4-8. Table Spellout Rule for English Articles

We're going to skip Nouns 1:4 because it contains several complexities which we're not yet ready to deal with, so go to Nouns 1:5. The window below the grammar tree indicates that this verse will generate *This man walked*. The features under MAN-A indicate that its Number is Singular, Participant Tracking is Routine, and Proximity is Near Speaker and Listener. If you push the Generate button now, you'll get *The man walked*. You may either revise your table spellout rule which inserts Articles so that it also inserts *this* when the Proximity is Near Speaker and Listener, or you may want to write a completely different rule that inserts Demonstratives. If you write a new rule that inserts Demonstratives, you'll need to modify the rule that inserts Articles so that it only applies when Proximity is Not Applicable; you'd probably want to specify that feature in the upper left corner of the table for articles. Choose whichever of these two options appeals to you the most, and generate the text. When you get the right results, save the verse and go to Nouns 1:6.

In Nouns 1:6 the small window below the grammar tree shows that this verse generates *This man walked*. The Proximity feature under MAN-A is Near Speaker. Revise your rule that inserts *this* so that it also inserts *this* when the Proximity is Near Speaker. When you generate the right results, save the verse. Repeat this process for Nouns 1:7.

4.5 Checking your Target Grammar

Periodically you need to make sure that all of your grammar rules are working properly and that you're generating the desired text. To check your grammar, select Generator in the main menu and then select Generate Text. At this screen you'll generate many verses of text rather than just one at a time. Right now we want to compare the generated text with the text that you've previously saved for each verse to make sure that all of the rules are working together properly. In order to compare the text generated now with the saved text, click the Setup button in the upper center of the screen. You'll see a dialog with a frame called Text Output Instructions. Select the option labeled Compare New Text with Saved Text. This instructs TBTA that when you're at this screen, it should compare the newly generated text with the text that you saved in the database when you were generating one verse at a time and clicking the Save button. Close the Setup dialog and you'll now see two tables, one labeled Saved My English (or whatever you entered for your target language name), and the other one is labeled New My English. Click the Text button in the upper left corner and tell it to generate Nouns 1:1 through Nouns 1:7. Then click the Generate button. You'll see the saved text appear immediately in the table on the left, and the newly generated text will appear in the table on the right. Whenever you run this generation process, if the new results differ from the saved results, TBTA will show you the differences highlighted in red so that you can make the necessary corrections. Another tutorial will discuss this screen in more detail; for now, just make sure that all of the newly generated text is correct. After Nouns 1:1 through Nouns 1:7 have been generated, only the verses that have differences will be displayed. If one or more verses have differences, go back and make the necessary corrections at the Lexicon and Grammar Development screen. Then come to this screen again and generate all of the text again to make sure that your grammar is working properly. Note that since we skipped Nouns 1:4, no text has been saved for that verse. Therefore the newly generated text will be different from the saved text, so that verse will be displayed in red. That's ok; let's leave it that way for now.

4.6 Conclusion

In this lesson you've learned about two of the eight types of rules that are in TBTA's synthesizing grammar. You've written Spellout rules for Nouns and Verbs, and you've edited a couple of Phrase Structure rules. You've also learned how to compare text that you've saved with newly generated text at the Generator, Generate Text screen. As you work through additional text, it will be important for you to occasionally return to the Generate Text screen and make sure that your grammar is generating the intended results. As you write new rules and edit existing rules, sometimes the previously generated text will no longer be generated properly. So you must periodically regenerate all of your text to make sure that the entire grammar is working correctly.

Another important point to remember from this lesson is that the lexical spellout rules which you wrote during the previous lesson generated lexical forms in the lexicon. Those lexical forms are lists of words in the lexicon. For example, you have a list of nouns called Plurals, and you have a list of verbs called Past Tense. But those are just lists of words in the lexicon. You must write additional grammar rules that instruct TBTA when to use words from those lexical lists. So you must write additional spellout rules in the synthesizing grammar to instruct TBTA when to use the various forms defined in the lexicon.