Indonesian Phrase Structure

Lexicon:

N: orang (person), perempuan (woman), guru (teacher), rumah (house), jalan (road), restoran (restaurant), teh (tea)
D: ini (this), itu (that)
A: kaya (rich), tua (old), senang (happy), besar (big), baru (new)

Phrase structure rules:

By observing the data set, we note the following:

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) orang</td>
<td>N [person]</td>
<td>(a) person</td>
<td>We can immediately develop the following:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NP → N</td>
</tr>
<tr>
<td>(2) orang ini</td>
<td>N [person] D [this]</td>
<td>this person</td>
<td>The determiner, if present, follows the noun:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NP → N (D)</td>
</tr>
<tr>
<td>(5) orang tua itu</td>
<td>N [person] A [old]</td>
<td>that old person</td>
<td>The adjective follows the noun but precedes the determiner:</td>
</tr>
<tr>
<td></td>
<td>D [that]</td>
<td></td>
<td>NP → N (A) (D)</td>
</tr>
<tr>
<td>(13) *perempuan itu senang</td>
<td>woman that happy</td>
<td>(that happy woman)</td>
<td>The ungrammaticality of the phrase sustains the suggested rule that the determiner must follow the adjective</td>
</tr>
<tr>
<td>(17) rumah besar</td>
<td>house big</td>
<td>(a) big house</td>
<td>This reinforces the fact that the adjective must follow the noun.</td>
</tr>
</tbody>
</table>

The rule that we derive using this data set is

\[
NP \rightarrow N \ (A) \ (D)
\]

We note that this is the exact inverse of English, where NPs abide by a rule which in its simplest form derived in class is

\[
NP \rightarrow \ (D) \ (A) \ N
\]

Indefinite determiner
Updated Lexicon:

N: orang (person), perempuan (woman), guru (teacher), rumah (house), jalan (road), restoran (restaurant), teh (tea)
D: ini (this), itu (that, the)
A: kaya (rich), tua (old), senang (happy), besar (big), baru (new)
V*: duduk (sit), minum (drink), lihat (see)
P: di (in), dekat (near)

*: Since all verbs in this data set are all in the past tense and no distinction can be made with regards to tense, the verbs have been listed in their imperative form. Additional research showed that in Indonesian, verbs are not marked for tense, and are instead indicated by auxiliary time markers.\(^1\)

Updated phrase structure rules I:

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(25) Orang kaya itu duduk.</td>
<td>NP[Person rich the] VP[sat].</td>
<td>The rich person sat.</td>
</tr>
</tbody>
</table>

Since we can divide the phrase into two constituent parts based on our previous rules, we derive the following sentence structure:

\[ S \rightarrow NP \ VP \]

The verb phrase here consists only of the verb, so we can only write the following:

\[ VP \rightarrow V \]

---

By looking at the literal translation we note again that we can divide the phrase into two
categories, with a NP nested in a VP as is done in English. The NP inside the VP is the
direct object, and therefore can act as a unit with the VP because the noun is receiving the
action. We therefore revise our rule to include a NP in a VP:
\[
\text{VP} \rightarrow \text{V (NP)}
\]

(32) Perempuan itu duduk
di rumah itu

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(28) Guru itu minum teh.</td>
<td>NP[Teacher the] VP[drank NP[tea]].</td>
<td><em>The teacher drank tea.</em></td>
</tr>
<tr>
<td>(32) Perempuan itu duduk di rumah itu</td>
<td>NP[Woman the] VP[sat PP[in NP[house the]]].</td>
<td><em>The woman sat in the house.</em></td>
</tr>
</tbody>
</table>

To account for the PP, we revise our rule to include them in the VP:
\[
\text{VP} \rightarrow \text{V (PP)}
\]

As was the case with the previous example, we observe that the preposition ‘in’ locates the
noun ‘house’, and therefore should be considered as a unit.
\[
\text{PP} \rightarrow \text{P NP}
\]

We might also be able to consider the possibility that a preposition can exist in an NP, but we will refrain from developing that until data suggests otherwise.
<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(35) Perempuan kaya ini minum teh di rumah baru itu.</td>
<td>NP[Woman rich this] VP[drank NP[tea PP[in NP[house new the]]]].</td>
<td>This rich woman drank tea in the new house.</td>
</tr>
<tr>
<td>(36) Orang itu lihat guru dekat jalan besar ini.</td>
<td>NP[Person the] VP[see NP[teacher PP[near NP[road big this]]]].</td>
<td>The person saw a teacher near this big road.</td>
</tr>
</tbody>
</table>

We form the following rule to account for the NP immediately following the verb:

\[
VP \rightarrow V \ ( NP) \ ( PP)
\]

We can also consider the possibility of adding an optional PP inside an NP (and thus allowing for recursion), but since our rules allow for the correct generation of the above phrases, we remain with its simplest form.

| (30) *Perempuan itu orang ini lihat. | NP[Woman the] NP[person this] VP[saw]. | (The woman saw this person.)               |

These are the rules developed so far:

\[
S \rightarrow NP \ VP \\
NP \rightarrow N \ (A) \ (D) \\
VP \rightarrow V \ (NP) \ (PP) \\
PP \rightarrow P \ NP
\]

The ungrammaticality of the above phrase suggests that there cannot be an NP NP VP construction, or that the VP cannot contain a NP in its beginning. A better suggestion is
that a verb must follow the NP it modifies, as we see in phrase (29). None of the rules above allow for the generation of this phrase.

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(27) *Minum perempuan tua ini.</td>
<td>VP[Drink] NP[woman old this].</td>
<td>(This old woman drank.)</td>
</tr>
</tbody>
</table>

The ungrammaticality of the above suggests that the sentence cannot have a VP NP structure. None of the rules above allow for the generation of this phrase.

Updated phrase structure rules II:

(37) Guru itu senang.  
\[NP[Teacher the] happy.\]  
\[The teacher is happy.\]

We note that the phrase is not ‘The happy teacher’, which would be

Guru senang itu.

We can identify the non-verbal predicate expression in the English translation (‘is happy’) that is absent from the Indonesian, and suggest therefore that a sentence need not have a VP:

\[S \rightarrow NP (VP)\]

In order to revise this rule to account for the adjective, and as to where to position the additional NP in the sentence, we look at more data:

<table>
<thead>
<tr>
<th>(38) Orang tua ini senang</th>
<th>NP[Person old this] happy.</th>
<th>This old person is happy</th>
</tr>
</thead>
<tbody>
<tr>
<td>(39) *Orang tua senang ini</td>
<td>Person old happy this.</td>
<td>(This old person is happy)</td>
</tr>
</tbody>
</table>

we note that the NP ‘Orang tua ini’ acts as a NP unit. This is also the case with the NP in the following phrase, as the adjective that modifies the noun ‘teacher’ is encapsulated in the NP construction of N A D.

<table>
<thead>
<tr>
<th>(40) Guru senang itu kaya.</th>
<th>NP[Teacher happy the] rich.</th>
<th>The happy teacher is rich.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(43) Orang kaya itu perempuan.</td>
<td>NP[Person rich the] NP[woman].</td>
<td>The rich person is a woman.</td>
</tr>
<tr>
<td>(47) Orang tua itu perempuan kaya.</td>
<td>NP[Person old the] NP[woman rich].</td>
<td>The old person is a rich woman.</td>
</tr>
</tbody>
</table>

Since both ‘woman’ and ‘the teacher’ are constituents of NPs, we can note that a sentence can consist of two NPs. Using only the data provided we can also arrive at the conclusion that if there are two NPs, the second acts as a predicate that modifies the first. In phrases
(37), (38), and (40) we note that the NP which follows is a **predicate adjective**, and the missing stative verb ‘is’ or (also the **copula**) is assumed. On the other hand, (43) and (47) contain examples of the **predicate nominal**, where the copular verb ‘to be’ links both noun phrases.

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(49) Orang tua itu di rumah besar itu.</td>
<td>NP[Person old this] PP[in NP[house big the]].</td>
<td>The old person is in the big house.</td>
</tr>
</tbody>
</table>

Here we observe an example of a PP that is unaccounted for in our rules. There are two possibilities we can suggest:

1) Rather than having an NP be considered as the predicate of the subject of a sentence, we can consider a PP. However, the data suggests that the preposition is not a common or necessary feature of the language, and so using a PP **without** a preposition to account for previous examples would be difficult to comprehend. We could, however, construct another rule that allows for a NP PP structure in a sentence, though there does not appear to be an immediate need for that rule yet.

2) A preposition can precede a noun in a NP. This would also account for the alternate interpretation of phrase (32), where instead of:

\[
\text{NP[Woman the] VP[sat PP[in NP[house the]]]}
\]

we can develop:

\[
\text{NP[Woman the] VP[sat NP[in house the]]}.\]

We therefore suggest:

\[
\text{NP} \rightarrow (P) N (A) (D)
\]

(54) *Perempuan kaya ini tua minum*

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(54) *Perempuan kaya ini tua minum</td>
<td>NP[Woman rich this] old drank.</td>
<td>(This old rich woman drank.)</td>
</tr>
</tbody>
</table>
If the adjective ‘old’ were to modify ‘woman’ as it does in the English translation, we would have to develop a rule that allows for more than one adjective in a noun phrase, and no other phrase in the data set has a grammatical form including two consecutive adjectives. The ungrammaticality becomes more evident when we compare the above to phrase (26):

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(26) Perempuan tua ini minum</td>
<td>NP[Woman old this] VP[drank].</td>
<td>This old woman drank.</td>
</tr>
</tbody>
</table>

As we examined with phrase (30), we cannot have a NP NP VP construction in a sentence. Furthermore, we can obtain from the data set that the verb must immediately follow the NP is modifies. Here, the person is sitting, so the verb must follow the initial NP. The correct form should be:

(55’) Orang itu duduk dekat restoran.

Additional note : Based on phrases (33) and (34), we suggest suggest the difference between a definite and an indefinite article would define the presence of itu/ini, in that if the phrase were referring to ‘the restaurant’, an *itu* would be added to the end of (55’)

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Literal translation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(34) Orang ini duduk dekat jalan itu.</td>
<td>NP[Person this] VP[sat PP[near NP[road the]].</td>
<td>This person sat near the road.</td>
</tr>
</tbody>
</table>

Our final phrase structure rules are as follows:

\[
\begin{align*}
S & \rightarrow \text{NP}\{(\text{VP})\} \\
\text{NP} & \rightarrow \text{P} \quad \text{N} \quad \text{A} \quad \text{D} \\
\text{VP} & \rightarrow \text{V} \quad \text{NP} \quad \text{PP} \\
\text{PP} & \rightarrow \text{P} \quad \text{NP}
\end{align*}
\]

We note that they are relatively similar to the English phrase structure rules derived in class, as both the rules for VPs and PPs are nearly identical, and the NP is almost exactly the reverse. The difference in NP structure is interesting, since it suggests that the determiner, which allows an English-speaker to describe the quantity or possessor of the noun, is not immediately important in understanding the phrase. This is also true of the adjective, since if the phrase in Indonesian were spoken, we would only understand the precise description of the object *after* being told about the object. If the assumption about the preposition inside an NP is correct, this also suggests the necessity in declaring
The sentence structure is identical, which can be explained by the fact that Indonesian is a Subject-Verb-Object (SVO) language\(^2\), as is English. The most prominent difference is the lack of the verb ‘be’ (as in Russian), which allows the use of a non-verbal predicate, allowing the phrase to ‘assume’ the copula.

It is also interesting to note that, with regards to this data set, the VP must follow immediately after the NP it refers to or modifies. It would be very revealing to know how the following would be translated:

The woman the teacher saw was old. 
The old woman, but not the teacher, saw the old person. 
The student, whose name I cannot remember, should most likely be at the seminar tomorrow.

---