Guide to Presenting Rules in the Linguistics Olympiad

This guide introduces ways to show your linguistic rules when writing your solutions for the Linguistics Olympiad. Before you start reading, please keep these three things in mind: (1) Following this guide will not necessarily give you more points. You will get points based on how complete and accurate your answer is. (2) This guide does not limit how you write your answers. You are free to use any other logical way to show the rules—for example, using plain text instead of the symbols introduced in this guide. (3) The content provided in this guide is only a set of general principles. If a problem has its own special requirements, you should follow those instructions to make sure your answer meets the requirements.

1.   Describe your rules as clearly as possible.

Make sure to write your rules with certainty. For example, write "A is B" instead of "A might be B" or "A can be B." Phrases like "might be" or "can be" imply that it also "might not be" or "does not have to be."

If you want to say that A has multiple possibilities, such as being either B or C, and there are reasons why A becomes B or C, it is better to say: "A is B under one condition, and A is C under another condition." It is not as good to say: "A might be B or C, depending on the conditions." For more on how to write conditions, please see sections 7 to 10.

2.   Describe all the rules you find, even if they seem very basic.

Basic rules are often the easiest to overlook. Be sure to write them all down so you don't miss out on any points. For example, in problems about Morphology or Syntax, you should describe the structure of words or sentences (i.e., word order). In problems about Writing Systems, you should describe the writing direction and what a single symbol represents (such as a syllable or a phone).

3.   Use linguistic terms and generalize into hypernyms as much as possible, instead of just listing individual words.

If you do not use linguistic terms or fail to generalize into hypernyms, your answer may be considered incomplete, and you might not get points. For example, suppose that in APLO 2025 Problem 4, the marking scheme requires identifying that for noun phrase structures, female animals use **N ʦuhu**. If your answer only lists that "male chiefs and male workers use **N-õ**," "female chiefs and daughters use **N-uhu**," "female sloths and female monkeys use **N ʦuhu**," and "male sloths and male monkeys use **N**," without generalizing into hypernyms such as human, animal, female, or male, your answer might be seen as insufficient. As a result, you might not receive the points for that part of the problem.

For another example, in English, the second person does not distinguish between singular and plural. If you write *you* in your rules instead of *second-person singular*, you might not meet the requirements of the marking scheme because you failed to exclude the "second-person plural" (or vice versa). To be more detailed, if the marking scheme requires identifying that "the second-person

singular is expressed as...", you might get points because *you* already includes the singular form. However, if the marking scheme requires identifying that "only the second-person singular is expressed as...", you cannot get points because *you* does not exclude the second-person plural.

Similarly, in English, the first person does not distinguish between dual and plural. If you write *we* in your rules instead of *first-person dual*, you might not meet the requirements of the marking scheme because you failed to exclude the "first-person plural" (or vice versa).

4. You do not need to write down your reasoning process or a word bank.

When answering, you should describe the "rules" in detail, not the "process of finding them." There is no need to write "we know this from..." for each rule. The reasoning process is not part of the marking scheme.

When answering, you should describe the "rule" in detail, not "one-to-one word translations." There is no need to compile a dictionary or a vocabulary list. While making a word bank helps you find the answers, it is not necessary. If you have already finished writing the answers and rules, you do not need to go back and list every single word just for the sake of doing it. However, please note that if certain affixes have specific functions—for example, different shapes of objects use different classifier affixes—you should still list all those affixes and what they mean, as seen in APLO 2025 Problem 4.

Some official solutions from APLO and IOL show a word bank, such as in APLO 2024 Problem 4 or APLO 2025 Problem 2. These are merely there to help you fully understand the problem after the competition; they are not instructions for you to write a word bank during the competition.

Please remember that these are only general principles. If a problem clearly asks you to write down your reasoning or a word bank, you should still follow those instructions to meet the requirements.

5. You can use linguistic symbols or commonly used symbols to present or simplify the rules you find.

To save space in your answers, you may use linguistic symbols or commonly used symbols. Appendix 1 explains the meanings of symbols such as "C, V, σ, H, L, Ø, #, [ ], / /, | |, < >, $, ." which are generally used at the phonological level, while Appendix 2 explains the meanings of symbols such as "S, O, V, Vi, Vt, A, P, adj, N, Ø, I, II, III, *sg.*, *du.*, *pl.*, *m.*, *f.*, ♂, ♀" which are generally used at the syntactic level. You do not need to provide extra explanations when using these symbols in your answers. If you choose to use other self-defined symbols, you must provide their definitions in your answers.

When using the symbols mentioned above, please note the following:

(1) Since H and L can stand for *High/Low* or *Heavy/Light*, you should generally explain their meanings in your answers. Otherwise, it may be unclear whether σH and σL refer to high/low tone syllables or heavy/light syllables. If the meaning is clear from the problem—for example, if the problem states that the language has high and low tones—you may skip the

explanation.

(2) Although V can stand for both *Vowel* and *Verb* and therefore has multiple meanings, the former is usually paired with consonants or other phonetic symbols, while the latter is usually paired with symbols like subjects and objects. Thus, they are unlikely to be confused, and an explanation is generally not needed.

(3) You can use subscript numbers to distinguish symbols, such as C_1 and C_2 for the first and second consonants.

(4) You can define sub-concepts of symbols by adding subscript text. For example, use V_{+w} or $V_{rounded}$ for rounded vowels, or V_i and V_f for the initial and final vowels, respectively. In these cases, you should explain your definitions.

(5) If there is no need to distinguish between [ ] and / /, you may choose either one. If there is no need to distinguish between sounds and writing systems, these symbols can be omitted.

(6) There are many ways to mark stress; you should follow the marking style described in the problem.

(7) You may use other linguistic symbols not mentioned in Appendix 1 or Appendix 2. However, to avoid losing points due to incomplete definitions, we recommend that you fully explain the meaning of any such symbols.

6. You can use hyphens "-" to indicate the relative position of a symbol.

For example, **b-** can represent **b** at the beginning of a syllable or word, or the prefix **b-**; **-b-** can represent **-b-** in the middle of a syllable or word, or the infix **-b-**; and **-b** can represent b at the end of a syllable or word, or the suffix **-b**.

7. You can use the equals sign "=" to mean "is/be," and use arrows "→" or the greater-than sign ">" to mean "changes into."

For example, the solution for Problem 1 of IOL 2025 describes System A:

$\beta \times 20(+\alpha)$＝**ke** β (**da** α)

$\beta \times 2 + 10 \rightarrow$ **ke pɟe-da** $\boxed{\beta+1}$

$\beta \times 20 + 15 \rightarrow$ **ke ko-da** $\boxed{\beta+1}$

This means "$\beta \times 20(+\alpha)$ is **ke** β (**da** α), where $\beta \times 20 + 10$ changes into **ke pɟe-da** $\beta+1$, and $\beta \times 20 + 15$ changes into **ke ko-da** $\beta+1$."

For example, the solution for Problem 3 of IOL 2025 describes verbs (excerpt):

-n-V**-** > **-naa-**, else**-**V**-** = preceding vowel; **nβ** > **mb**; **nr** > **nd**

This means "**-**V**-** changes into **-naa-** when preceded by **n-**, else **-**V**-** is the preceding vowel; **nβ** changes into **mb**; and **nr** changes into **nd**."

Additionally, you can use the inequality sign "≠" to represent the negation of "is/be," which means "is not."

8. You can use a list format or a multi-line left curly bracket "{" to represent changes under different conditions, and write the respective conditions on the right as

needed.

For example, the solution for Problem 5 of IOL 2022 describes the penultimate syllable:

*ʔV > V
*lV > **li**
*bV(h)L > **piL**
otherwise CV(V) > C**a**
-h > ∅

This means "***ʔV** changes into V, ***lV** changes into **li**, ***bV(h)L** changes into **piL**, otherwise CV(V) changes into C**a**, and **-h** changes into ∅."

For example, the solution for Problem 2 of APLO 2025 describes sentence structures:

$$\begin{cases} \text{S (O) V—\textbf{zu}} & \text{V = active} \\ \text{S V—\textbf{βɛ}} & \text{V = stative} \end{cases}$$

This means "when the verb is an active verb, the sentence structure is S (O) V—**zu**; when the verb is a stative verb, the sentence structure is S V—**βɛ**." The same content can also be rewritten in a list format using a colon ":", for example:

V = active: S (O) V—**zu**
V = stative: S V—**βɛ**

Please note: Although a colon can also be replaced by an arrow "→" to represent the so-called "if... then..." in the field of logic, arrows in the official APLO and IOL solutions are generally only used to mean "changes into" as introduced in section 7. To develop consistent answering habits, we recommend that you use arrows only to indicate the direction of change and not as a logical symbol.

To give another example, the solution for Problem 5 of APLO 2024 describes the structure of numerals (excerpt):

$$10X + Y = \begin{cases} \boxed{10X}\text{-ri Y} & 0 < Y \leq 4 \\ \boxed{10\,(X+1)}\ \textbf{maben Y} & 4 < Y \leq 9 \end{cases}$$

This means "regarding 10X + Y, when $0 < Y \leq 4$, the numeral structure is $\boxed{10X}$-ri Y; when $4 < Y \leq 9$, the numeral structure is $\boxed{10\,(X+1)}$ **maben Y**."

9. When expressing "sound changes," you have multiple options for writing environmental conditions.

When describing the change of a sound or an affix under a specific environmental condition, as mentioned in Section 8, you can describe the environmental condition on the left (optionally connected by a colon ":"). For example, the solution for Problem 5 of IOL 2017 describes:

before another prefix **le-ven-** > **len-**
before **p**   **-n** > **m**
before **k**   **-n** > **ng**

This means "l**e-ven-** changes into **len-** when followed by another prefix; **-n** changes into m before **p**, and into m before **k**."

On the other hand, as mentioned in Section 8, you can also use a multi-line left curly bracket and

write the environmental conditions on the right. When writing conditions on the right, you can generally make the rules more concise by utilizing the symbols introduced in Section 5. For example, the solution for Problem 4 of APLO 2023 describes the notes on person (excerpt):

$$\mathbf{R} = \begin{cases} \mathbf{n} & \#\_ \\ \mathbf{r} & \text{otherwise} \end{cases}$$

This means "**R** is **n** word-initially; otherwise, it is **r**."

In addition to the above, there is a professional way to write environmental conditions: write the condition on the right, use an underscore to represent the position, and separate it with a forward slash. This notation is equivalent to one of the formal ways to write phonological rules. For example, the solution for Problem 4 of APLO 2023 describes the notes on verb structures (excerpt):

me > **mo** / _o

This means "**me** changes into **mo** before **o**."

Please note: Due to formatting reasons, the official IOL solutions may use a vertical bar instead of a forward slash; however, this is not a standard linguistic practice. In actual answers, a vertical bar can easily be confused with the letter "l" or the number "1" if your handwriting is messy. Therefore, we still recommend using a forward slash. For example, the solution for Problem 3 of IOL 2016 describes person (excerpt):

tense: past **-s**, present/future **-r** (**r** > **d** | **n**_)

This means "use the suffix -s to express the past tense, and the suffix -r to express the present/future tense; r changes into d before n." In this notation, the vertical bar " | " can be replaced by a forward slash " / ."

For example, the solution for Problem 3 of IOL 2023 describes:

kb > **mb**; **e o** > **i u** | _C*VC*VC*#

This means "**kb** changes into **mb**; **e** and **o** change into **i** and **u**, respectively, before C*VC*VC*#." Please note: in this case, the same number of elements appear on both sides of the greater-than sign ">" and are separated by spaces to mean "applying rules respectively." This is a lazy and not recommended way to write rules. If the number of elements on each side is different—for example, if you mistakenly write "**a e o** > **i u** | _C*VC*VC*#" or "**e o** > **i ɨ u** | _C*VC*VC*#"—you may lose points because it is impossible to determine which vowel changes into which. A better way to write this would be "before C*VC*VC*#, **e** → **i** and **o** → **u**," though you are not limited to this format.

10. You can use tables to present or simplify the rules you find.

You can use a table to show the content that applies when specific row and column conditions are met. For example, the solution for Problem 2 of IOL 2025 describes possessive pronouns (excerpt):

	inalienable possession		alienable possession	
	kinship term	body part	singular	plural
singular	**V́***	**V**	**V́V̀n**	**V́nV̀g**
plural	**V̄g**		**V̄yV̀n**	**V̄yV̀g**
person marking	**V**: 1/2/3 → **a/ɔ/ɛ**		**V**: 1/2/3 → **ə/u/i** 3 → add a suffix **-ī**	

The notations for changes described in Sections 7–9 can be presented either inside or outside of a table.

Tables can be used independently or for partial descriptions of a structure. If it does not affect understanding, the borders of a table can be omitted. For example, the solution for Problem 4 of APLO 2025 describes noun phrase structures when animate:

	♂	♀				♂	♀
1	**bakʷõ**	**bakʷuhu**	+		human	N-**õ**	N-**uhu**
2	**dũhũtaha**				animal	N	N **tsuhu**
3	**wãmedukʷa**						

There is no specific order for table rows and columns. However, when reading horizontally, it is common practice to follow the row labels first and then the column labels. For example, in the case above, one would typically read, "it is **bakʷõ** for the first person ♂." Swapping rows and columns does not affect your answer; for instance, the example above can also be rewritten as:

	1	2	3			human	animal
♂	**bakʷõ**	**dũhũtaha**	**wãmedukʷa**	+	♂	N-**õ**	N
♀	**bakʷuhu**				♀	N-**uhu**	N **tsuhu**

11. You can use other mathematical symbols or logical symbols to present or simplify the rules you find, but it is not required.

If you are familiar with other mathematical symbols, you can apply them to your rules. For instance, symbols such as "+, −, ∈, ∉, ∋, ∌, { }" can be used to simplify the rules you wish to express. Appendix 3 explains the meanings of these mathematical symbols. Please note that using mathematical symbols is not mandatory; you should consider your own level of familiarity with these symbols and choose an appropriate way to answer.

For example, the solution for Problem 2 of IOL 2022 describes:

l → C | C_ (C∈{**b, f, ł, n**})
b → **m** | _C (C denotes any consonant)

This means "after **b**, **f**, **ł**, or **n**, **l** changes into the same sound; before any consonant, **b** changes into **m**." In this notation, the scope of C in the conditions is expressed by listing the elements of a set. The vertical bar " | " can be replaced by a forward slash " / " (see Section 9).

In addition to mathematical symbols, logical symbols sometimes appear in APLO and IOL solutions. However, logical symbols are difficult to use and are not encouraged. Appendix 4 explains the meanings of logical symbols such as "∃, ∄, ∧, ∨, ∀, ¬, ⇔, ∵, ∴" that have appeared in previous

solutions. For example, the solution for Problem 2 of APLO 2025 describes verbs (The original text of this problem uses the Andika font, so **a** is presented as **ɑ**. Regarding the distinction between **a** and **ɑ**, please see Section 12):

V.	active		stative
	∃O	∄O	—
dɑ	lay down	lie down	—
tɛ	lift	climb	—
bɑlɑ	grow (it)	grow	be big
gutu	shorten		be short

Among these, "∃O" means an object exists, and "∄O" means an object does not exist. From this, one can read that "when the verb **dɑ** expresses an action with an object, it means *to lay something down*; when it expresses an action without an object, it means *to lie down*." V. is a self-defined symbol.

12. Other Notes: Avoid writing errors that cause ambiguity.

The following describes common issue caused by poor writing habits or messy handwriting among contestants.

(1) Mistaking the letter **i** for **ì**, **í**, or **ī**

Due to poor habits or excessive pressure when dotting the letter, i may be written as **ì**, **í**, or **ī**. While this may be acceptable if it does not cause confusion, it can lead to misunderstandings in problems where strict distinctions are required (e.g., if the problem uses **ì** or **í** to represent stress, or **ī** to represent a long vowel). In such cases, this may result in a loss of points.

(2) Confusing the letters **a** and **ɑ**

Although a and ɑ are generally just the difference between printed and handwritten styles, they must not be confused in the International Phonetic Alphabet (IPA) or in certain orthographies using the Latin alphabet. Additionally, some contestants have a habit of writing **a** in a writing style similar to **2** or **α**; this can be confused with **ə** and may lead to a loss of points.

(Note: In Cameroonian languages, the uppercase version of **ɑ**, which is **Ɑ**, is also used.)

(3) Confusing the letter **h** and **n**

If the length of the vertical stroke is unclear, it may lead to confusion and a loss of points.

(4) Confusing the letter **ɑ** and **u**, or **h** and **b**

The loop of **ɑ** and the bottom of **b** must be fully closed. Otherwise, they may be confused with **u** or **h**, leading to a loss of points.

(5) Other confusions

Certain symbols in the International Phonetic Alphabet (IPA) are often written incorrectly. For example, **ɯ** is wrongly written as **ω** or **w**, **ʃ** as **f** or **f**, **ɔ** as **c**, and **ɛ** as **ȝ**. It is best to follow the style of the symbols provided in the problem exactly. Do not omit strokes or add serifs

on your own to avoid confusion with other symbols, which may lead to a loss of points.

Appendix 1: Symbols for Phonological Levels and Their Meanings

C	consonant	[ ]	phonetic or relatively narrow transcription
V	vowel	/ /	phonemic or relatively broad transcription
σ	syllable	| |	glyph markers (often omitted)
H	high or heavy	< >	grapheme markers
L	low or light	$	syllable boundary (for phonological rules)
∅	zero, null	.	syllable boundary (IPA)
#	word boundary		

Appendix 2: Symbols for Syntactic Levels and Their Meanings

S	subject	I	first person
O	object	II	second person
V	verb	III	third person
Vi	intransitive verb	sg.	singular
Vt	transitive verb	du.	dual
A	agent	pl.	plural
P	patient	m.	masculine
adj	adjective	f.	feminine
N	noun	♂	male
∅	zero, null	♀	female

Arabic numerals are sometimes used instead, provided that there is no risk of confusion.

Appendix 3: Symbols for Mathematical Operations and Their Meanings

+	presence or combination (conventional)	∋	contains sth.
−	absence or deduction (conventional)	∌	does not contain sth.
∈	belongs to sth.	{ }	listing elements of a set
∉	does not belongs to sth.		

Appendix 4: Symbols for Logicals and Their Meanings

∃	exists	¬	not
∄	does not exists	⇔	equivalence
∧	and	∵	because
∨	or	∴	therefore
∀	for all		

Hsiang-Yu Lei
Ver 1.1, 2026.3.24