



Oct. 30, 2020

# Learning Objectives for Today

- Review of Proposal Presentation
- The Science of Scientific Writing

# Scientific Writing

Purpose:

The goal is to communicate

Accurate information is useless if the reader cannot understand it

# Writing with the Reader in Mind: Expectation and Context

- Readers interpret by themselves
- Example: tracking the temperature of a liquid over a period of time

$t$  (time)=15',  $T$  (temperature)=32°;  $t=0'$ ,  $T=25°$ ;  
 $t=6'$ ,  $T=29°$ ;  $t=3'$ ,  $T=27°$ ;  $t=12'$ ,  $T=32°$ ;  $t=9'$ ,  $T=31°$

<u>time (min)</u>		<u>temperature (°C)</u>	<u>temperature (°C)</u>
Left: Information appears in a pattern (Context)	0	25	25
	3	27	27
	6	29	29
	9	31	31
	12	32	32
	15	32	32
		<u>temperature (°C)</u>	<u>time (min)</u>
		25	0
		27	3
		29	6
		31	9
		32	12
		32	15

# Reader Expectations for the Structure of Prose

The smallest of the URF's (URFA6L), a 207-nucleotide (nt) reading frame overlapping out of phase the NH<sub>2</sub>-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene has been identified as the animal equivalent of the recently discovered yeast H<sup>+</sup>-ATPase subunit 8 gene. The functional significance of the other URF's has been, on the contrary, elusive. Recently, however, immunoprecipitation experiments with antibodies to purified, rotenone-sensitive NADH-ubiquinone oxido-reductase [hereafter referred to as respiratory chain NADH dehydrogenase or Complex I] from bovine heart, as well as enzyme fractionation studies, have indicated that six human URF's (that is, URF1, URF2, URF3, URF4, URF4L, and URF5, hereafter referred to as ND1, ND2, ND3, ND4, ND4L, and ND5) encode subunits of Complex I. This is a large complex that also contains many subunits synthesized in the cytoplasm.\*

# Reader Expectations for the Structure of Prose

The smallest of the URF's (URFA6L), an [A] has been identified as a [B] subunit 8 gene. The functional significance of the other URF's has been, on the contrary, elusive. Recently, however, [C] experiments, as well as [D] studies, have indicated that six human URF's (1-6) encode subunits of Complex I. This is a large complex that also contains many subunits synthesized in the cytoplasm.

# Subject-Verb Separation

The smallest of the URF's (URFA6L), a 207-nucleotide (nt) reading frame overlapping out of phase the NH<sub>2</sub>-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene has been identified as the animal equivalent of the recently discovered yeast H<sup>+</sup>-ATPase subunit 8 gene. The functional significance of the other URF's has been, on the

The smallest of the URF's is URFA6L, a 207-nucleotide (nt) reading frame overlapping out of phase the NH<sub>2</sub>-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene; it has been identified as the animal equivalent of the recently discovered yeast H<sup>+</sup>-ATPase subunit 8 gene.

# The Stress Position

- Stress position: emphasize the material at the end of a sentence
- The relative values of the sentence's contents
- If not:
  - readers get lost
  - or distracted by other less important materials
- The stress position coincides with the moment of syntactic closure

The smallest of the URF's is URFA6L, a 207-nucleotide (nt) reading frame overlapping out of phase the NH<sub>2</sub>-terminal portion of the adenosinetriphosphatase (ATPase) subunit 6 gene; it has been identified as the animal equivalent of the recently discovered yeast H<sup>+</sup>-ATPase subunit 8 gene.



# Practice

- Grammatical subjects -> verbs
- Every unit of discourse -> serve a single function
- Information emphasized -> at points of syntactic closure

Recently, however, immunoprecipitation experiments with antibodies to purified, rotenone-sensitive NADH-ubiquinone oxido-reductase [hereafter referred to as respiratory chain NADH dehydrogenase or Complex I] from bovine heart, as well as enzyme fractionation studies, have indicated that six human URF's (that is, URF1, URF2, URF3, URF4, URF4L, and URF5, hereafter referred to as ND1, ND2, ND3, ND4, ND4L, and ND5) encode subunits of Complex I. This is a large complex that also contains many subunits synthesized in the cytoplasm.

The smallest of the URF's, URFA6L, has been identified as the animal equivalent of the recently discovered yeast  $H^+$ -ATPase subunit 8 gene; but the functional significance of other URF's has been more elusive. Recently, however, several human URF's have been shown to encode subunits of rotenone-sensitive NADH-ubiquinone oxido-reductase. This is a large complex that also contains many subunits synthesized in the cytoplasm; it will be referred to hereafter as respiratory chain NADH dehydrogenase or Complex I. Six subunits of Complex I were shown by enzyme fractionation studies and immunoprecipitation experiments to be encoded by six human URF's (URF1, URF2, URF3, URF4, URF4L, and URF5); these URF's will be referred to subsequently as ND1, ND2, ND3, ND4, ND4L, and ND5.

# The Topic Position

- In the stress position the reader needs and expects closure and fulfillment
- In the topic position the reader needs and expects perspective and context
- What appears at the beginning of sentences matters

“Bees disperse pollen” and “Pollen is dispersed by bees”

- Linkage (looking backward) and context (looking forward)
- Topic position: old information
- Stress position: new information

Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates at which tectonic plates move and accumulate strain at their boundaries are approximately uniform. Therefore, in first approximation, one may expect that large ruptures of the same fault segment will occur at approximately constant time intervals. If subsequent mainshocks have different amounts of slip across the fault, then the recurrence time may vary, and the basic idea of periodic mainshocks must be modified. For great plate boundary ruptures the length and slip often vary by a factor of 2. Along the southern segment of the San Andreas fault the recurrence interval is 145 years with variations of several decades. The smaller the standard deviation of the average recurrence interval, the more specific could be the long term prediction of a future mainshock.

1. The backward-linking old information appears in the topic position.
2. The person, thing or concept whose story it is appears in the topic position.
3. The new, emphasis-worthy information appears in the stress position.

Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates at which tectonic plates move and accumulate strain at their boundaries are roughly uniform. Therefore, nearly constant time intervals (at first approximation) would be expected between large ruptures of the same fault segment. [However?], the recurrence time may vary; the basic idea of periodic mainshocks may need to be modified if subsequent mainshocks have different amounts of slip across the fault. [Indeed?], the length and slip of great plate boundary ruptures often vary by a factor of 2. [For example?], the recurrence interval along the southern segment of the San Andreas fault is 145 years with variations of several decades. The smaller the standard deviation of the average recurrence interval, the more specific could be the long term prediction of a future mainshock.

**Put in the topic position the old information that links backward; put in the stress position the new information you want the reader to emphasize**

# Perceiving Logical Gaps

- When old information does not appear at all in a sentence, whether in the topic position or elsewhere, readers are left to construct the logical linkage by themselves

The enthalpy of hydrogen bond formation between the nucleoside bases 2'deoxyguanosine (dG) and 2'deoxycytidine (dC) has been determined by direct measurement. dG and dC were derivatized at the 5' and 3' hydroxyls with triisopropylsilyl groups to obtain solubility of the nucleosides in non-aqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. From isoperibolic titration measurements, the enthalpy of dC:dG base pair formation is  $-6.65 \pm 0.32$  kcal/mol.



The enthalpy of hydrogen bond formation between the nucleoside bases 2'deoxyguanosine (dG) and 2'deoxycytidine (dC) has been determined by direct measurement. dG and dC were derivatized at the 5' and 3' hydroxyls with triisopropylsilyl groups to obtain solubility of the nucleosides in non-aqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. From isoperibolic titration measurements, the enthalpy of dC:dG base pair formation is  $-6.65 \pm 0.32$  kcal/mol.

We have directly measured the enthalpy of hydrogen bond formation between the nucleoside bases 2'deoxyguanosine (dG) and 2'deoxycytidine (dC). dG and dC were derivatized at the 5' and 3' hydroxyls with triisopropylsilyl groups; these groups serve both to solubilize the nucleosides in non-aqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. Consequently, when the derivatized nucleosides are dissolved in non-aqueous solvents, hydrogen bonds form almost exclusively between the bases. Since the interbase hydrogen bonds are the only bonds to form upon mixing, their enthalpy of formation can be determined directly by measuring the enthalpy of mixing. From our isoperibolic titration measurements, the enthalpy of dC:dG base pair formation is  $-6.65 \pm 0.32$  kcal/mol.

# Locating the Action

Transcription of the 5S RNA genes in the egg extract is TFIIIA-dependent. This is surprising, because the concentration of TFIIIA is the same as in the oocyte nuclear extract. The other transcription factors and RNA polymerase III are presumed to be in excess over available TFIIIA, because tRNA genes are transcribed in the egg extract. The addition of egg extract to the oocyte nuclear extract has two effects on transcription efficiency. First, there is a general inhibition of transcription that can be alleviated in part by supplementation with high concentrations of RNA polymerase III. Second, egg extract destabilizes transcription complexes formed with oocyte but not somatic 5S RNA genes.

is  
is . . . is  
are presumed to be  
are transcribed  
has  
is . . . can be alleviated  
destabilizes

**Readers expect the action of a sentence to be articulated by the verb.**



Transcription of the 5S RNA genes in the egg extract is TFIIIA-dependent. This is surprising, because the concentration of TFIIIA is the same as in the oocyte nuclear extract. The other transcription factors and RNA polymerase III are presumed to be in excess over available TFIIIA, because tRNA genes are transcribed in the egg extract. The addition of egg extract to the oocyte nuclear extract has two effects on transcription efficiency. First, there is a general inhibition of transcription that can be alleviated in part by supplementation with high concentrations of RNA polymerase III. Second, egg extract destabilizes transcription complexes formed with oocyte but not somatic 5S RNA genes.

In the egg extract, the availability of TFIIIA limits transcription of the 5S RNA genes. This is surprising because the same concentration of TFIIIA does not limit transcription in the oocyte nuclear extract. In the egg extract, transcription is not limited by RNA polymerase or other factors because transcription of tRNA genes indicates that these factors are in excess over available TFIIIA. When added to the nuclear extract, the egg extract affected the efficiency of transcription in two ways. First, it inhibited transcription generally; this inhibition could be alleviated in part by supplementing the mixture with high concentrations of RNA polymerase III. Second, the egg extract destabilized transcription complexes formed by oocyte but not by somatic 5S genes.

# Writing and the Scientific Process

1. Follow a grammatical subject as soon as possible with its verb.
2. Place in the stress position the “new information” you want the reader to emphasize.
3. Place the person or thing whose “story” a sentence is telling at the beginning of the sentence, in the topic position.
4. Place appropriate “old information” (material already stated in the discourse) in the topic position for linkage backward and contextualization forward.
5. Articulate the action of every clause or sentence in its verb.
6. In general, provide context for your reader before asking that reader to consider anything new.
7. In general, try to ensure that the relative emphases of the substance coincide with the relative expectations for emphasis raised by the structure.

# Practice

- Each group will present two examples of sentences/short paragraphs that breaks the following principle and the revised paragraphs.  
(search any scientific articles online)

**G1** 1. Follow a grammatical subject as soon as possible with its verb.

**G2 G3** 2. Place in the stress position the “new information” you want the reader to emphasize.

**G4** 3. Place the person or thing whose “story” a sentence is telling at the beginning of the sentence, in the topic position.

**G5 G6 G7** 4. Place appropriate “old information” (material already stated in the discourse) in the topic position for linkage backward and contextualization forward.

**G8** 5. Articulate the action of every clause or sentence in its verb.