技術英語

Technical English

CUTURI, 高木, 中澤

Source for today's slides: *Science research writing for non-native speakers of English* by Hilary Glasman-Deal, an **excellent book that I recommend.**

Writing a Scientific Paper

Dynamic Programming Algorithm Optimization for Spoken Word Recognition

HIROAKI SAKOE AND SEIBI CHIBA

Abstract—This paper reports on an optimum dynamic programming (DP) based time-normalization algorithm for spoken word recognition. First, a general principle of time-normalization is given using time-warping function. Then, two time-normalized distance definitions, called symmetric and asymmetric forms, are derived from the principle. These two forms are compared with each other through theoretical discussions and experimental studies. The symmetric form algorithm superiority is established. A new technique, called slope constraint, is successfully introduced, in which the warping function slope is restricted so as to improve discrimination between words in different categories. The effective slope constraint characteristic is qualitatively analyzed, and the optimum slope constraint condition is determined through experiments. The optimized algorithm is then extensively subjected to experimental comparison with various DP-algorithms, previously applied to spoken word recognition by different research groups. The experiment shows that the present algorithm gives no more than about two-thirds errors, even compared to the best conventional algorithm.

I. Introduction

It is well known that speaking rate variation causes nonlinear fluctuation in a speech pattern time axis. Elimination of this fluctuation, or time-normalization, has been one of the central problems in spoken word recognition research. At an early stage, some linear normalization techniques were examined, in which timing differences between speech patterns were eliminated by linear transformation of the time axis. Reports on these efforts indicated that any linear transformation is inherently insufficient for dealing with highly complicated fluctuation nonlinearity as well as that time-normalization significantly improves recognition accuracy.

DP-matching, discussed in this paper, is a pattern matching algorithm with a nonlinear time-normalization effect. In this algorithm, the time-axis fluctuation is approximately modeled with a nonlinear warping function of some carefully specified properties. Timing differences between two speech patterns are eliminated by warping the time axis of one so that the maximum coincidence is attained with the other. Then, the time-normalized distance is calculated as the minimized residual distance between them. This minimization process is very efficiently carried out by use of the dynamic programming (DP) technique. The basic idea of DP-matching has been reported in several publications [1]-[3], where it has been shown by preliminary experiment on Japanese digit words that a recognition accuracy as high as 99.8 percent has been achieved, indicating the DP-matching effectiveness.

This paper reports an optimum algorithm for DP-matching through theoretical discussions and experimental studies. Investigations were made, based on the assumption that speech patterns are time-sampled with a common and uniform sampling period, as in most general cases. One of the problems discussed in this paper involves the relative superiority of either a symmetric form of DP-matching or an asymmetric one. In the asymmetric form, time-normalization is achieved by transforming the time axis of a speech pattern onto that of the other. In the symmetric form, on the other hand, both time axes are transformed onto a temporarily defined common axis. Theoretical and experimental comparisons show that the symmetric form gives better recognition than the asymmetric one. Another problem discussed concerns slope constraint technique. Since too much of the warping function flexibility sometimes results in poor discrimination between words in different categories, a constraint is newly introduced on the warping function slope. Detailed slope constraint condition is optimized through experimental studies. As a further investigation, the optimized algorithm is experimentally compared with several varieties of the DP-algorithm, which have been applied to spoken word recognition by some research groups [3]-[6]. The optimized algorithm superiority is established, indicating the validity of this investigation.

Manuscript received February 17, 1977; revised September 7, 1977.
The authors are with Central Research Laboratories, Nippon Electric Company, Limited, Kawasaki, Japan.

II. DP-Matching Principle

A. General Time-Normalized Distance Definition

Speech can be expressed by appropriate feature extraction as a sequence of feature vectors.

$$A = a_1, a_2, \cdots\cdots, a_i, \cdots\cdots, a_I$$
$$B = b_1, b_2, \cdots\cdots, b_j, \cdots\cdots, b_J. \quad (1)$$

Consider the problem of eliminating timing differences between these two speech patterns. In order to clarify the nature of time-axis fluctuation or timing differences, let us consider an i-j plane, shown in Fig. 1, where patterns A and B are developed along the i-axis and j-axis, respectively. Where these speech patterns are of the same category, the timing differences between them can be depicted by a sequence of points $c = (i, j)$:

$$F = c(1), c(2), \cdots\cdots, c(k), \cdots\cdots, c(K), \quad (2)$$

where

$$c(k) = (i(k), j(k)).$$

This sequence can be considered to represent a function which approximately realizes a mapping from the time axis of pattern A onto that of pattern B. Hereafter, it is called a warping function. When there is no timing difference between these

Writing a Scientific Paper

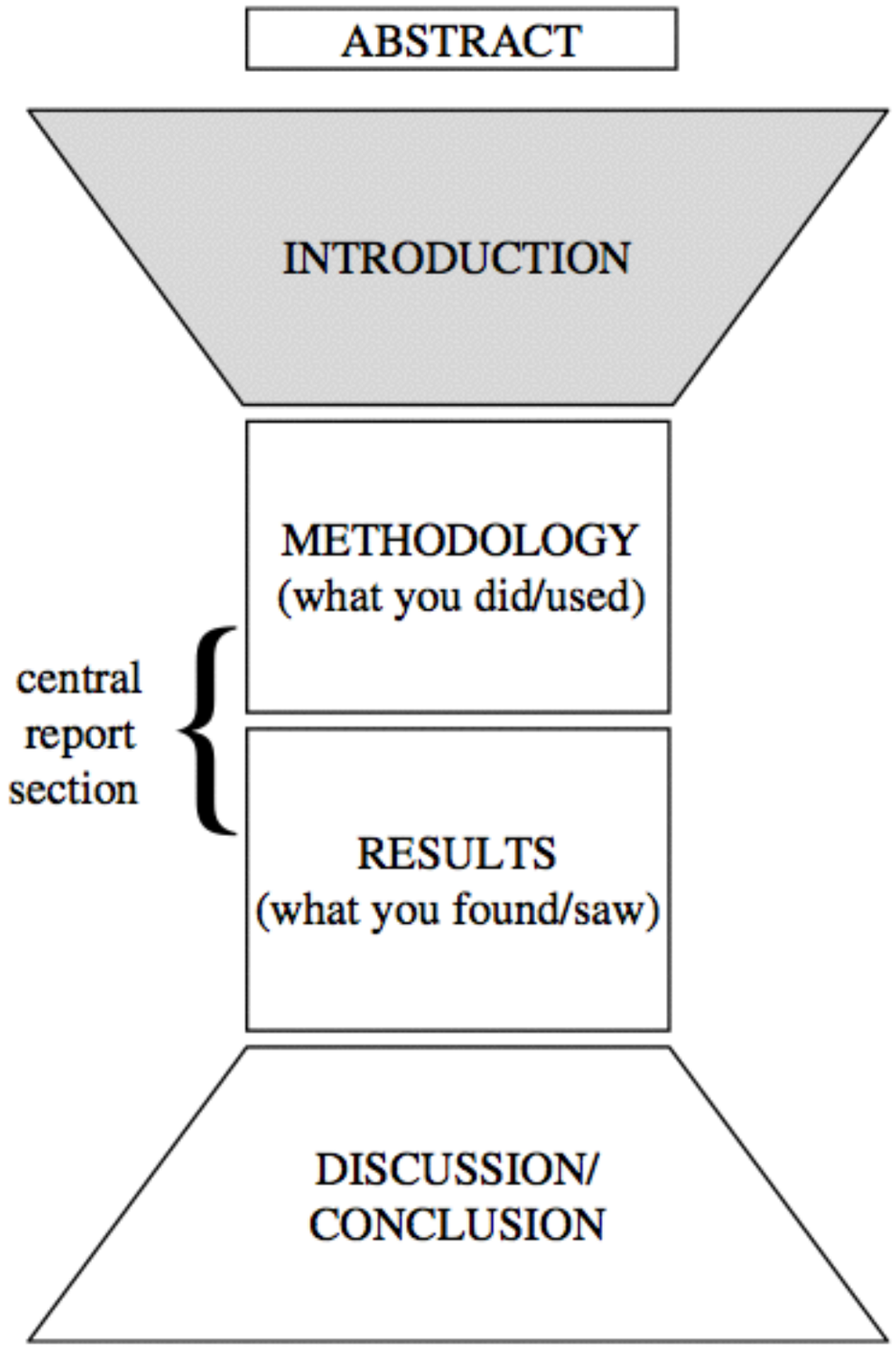


Fig. 1. The shape of a research article or thesis.

Abstract

Table 1. Characteristics of a well-written abstract.

Stands on its own without need to read the paper
States the hypothesis, question, or objective of the study
Completes the story by answering the hypothesis, question, or objective
Contains the same key words and terms as the title and the introduction
Follows the correct style and format
Follows the order of the main text (e.g., IMRAD)
Stays within the allowed word count
Does not contain information absent in the paper
Does not make conclusions unsupported by the data
Limits the use of abbreviations
Does not include references
Does not cite tables or figures

IMRAD format (Introduction, Methods, Results, and Discussion)

The Abstract and the Elevator Talk: A Tale of Two Summaries Thomas M. Annesley

Abstract

- Former US President Woodrow Wilson once said,
  - *If I am to speak ten minutes, I need a week for preparation;*
  - *if fifteen minutes, three days;*
  - *if half an hour, two days;*
  - *if an hour, I am ready now.*
- **Entire pages of information can be written in a short time**
- **The process of condensing that information with well-chosen words takes much more time.**

Introduction

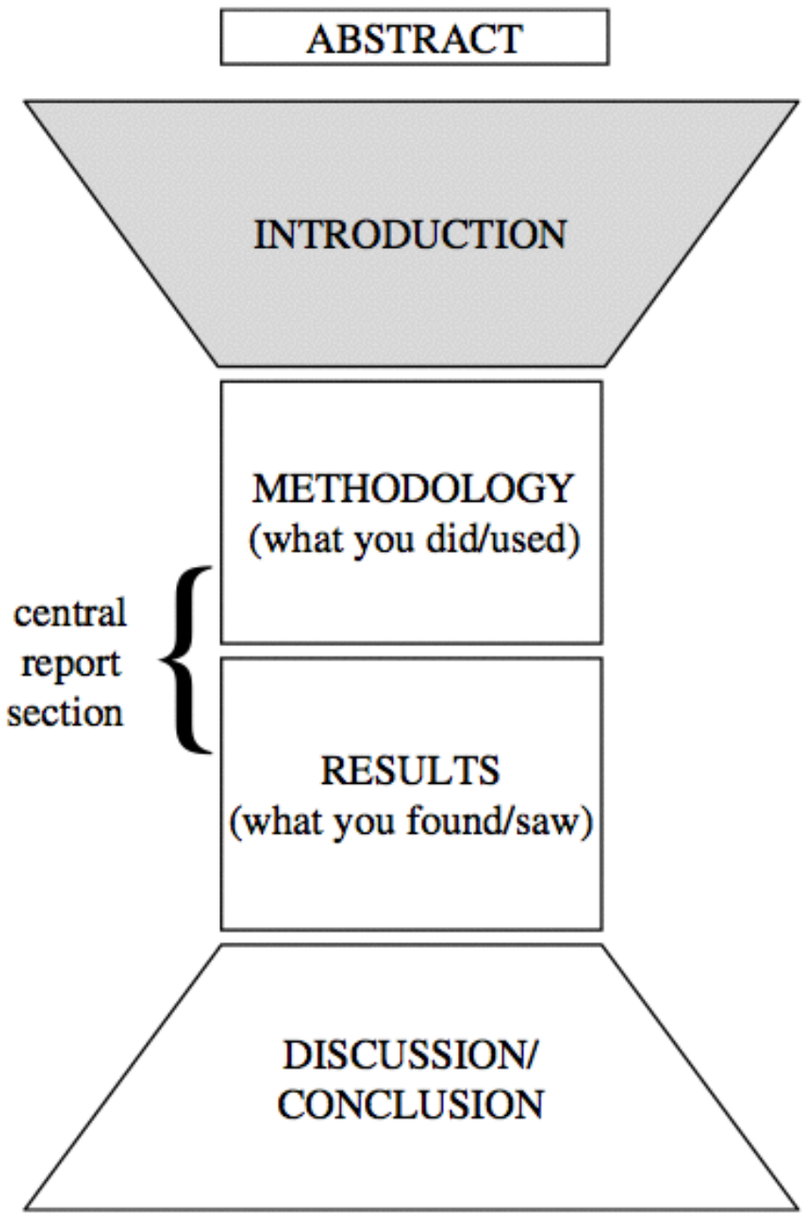


Fig. 1. The shape of a research article or thesis.

Science research writing for non-native speakers of English by Hilary Glasman-Deal.

Write an Introduction

- Overall Structure
- Grammar and important language skills
- Build a Model

Structure

- *How do writers usually start an introduction?*
- *What types of information should be in my introduction?*
- *How do writers usually end an introduction?*

Structure

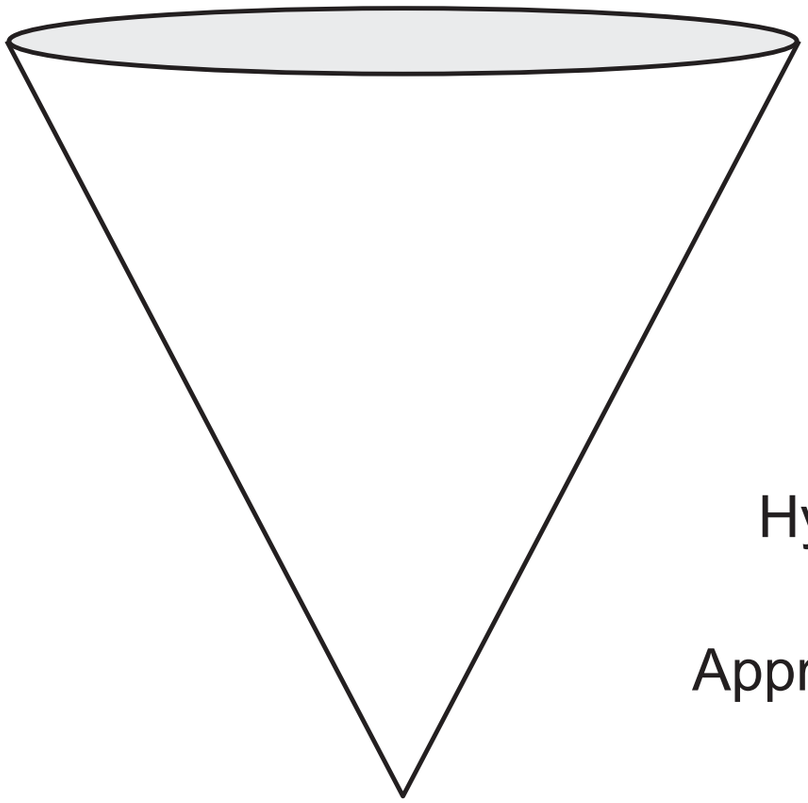
- Introduction & Conclusion are asymmetric (非対称)

- Start with general topics, converge to your contribution or point.

- Highlight the importance of the problem you discuss and explain as simply as possible your contribution

- Close by laying out the structure

Structure



Background, known information

Knowledge gap, unknown information

Hypothesis, question, purpose statement

Approach, plan of attack, proposed solution

Fig. 1. The introduction: a cone or funnel.

"It was a cold and rainy night":
Set the Scene with a Good Introduction Thomas M. Annesley

Introduction: Example

The synthesis of flexible polymer blends from polylactide and rubber

Introduction

1 *Polylactide (PLA) has received much attention in recent years due to its biodegradable properties, which offer important economic benefits.* **2** *PLA is a polymer obtained from corn and is produced by the polymerisation of lactide.* **3** *It has many possible uses in the biomedical field*[1] *and has also been investigated as a potential engineering material.*[2,3] **4** *However, it has been found to be too weak under impact to be used commercially.*[4]

5 *One way to toughen polymers is to incorporate a layer of rubber particles*[5] *and there has been extensive research regarding the rubber modification of PLA.* **6** *For example, Penney et al. showed that PLA composites could be prepared using blending techniques*[6] *and more recently, Hillier established the toughness of such* composites.[7] **7** *However, although the effect of the rubber particles on the mechanical properties of copolymer systems was demonstrated over two years ago,*[8] *little attention has been paid to the selection of an appropriate rubber component.*

8 *The present paper presents a set of criteria for selecting such a component.* **9** *On the basis of these criteria it then describes the preparation of a set of polymer blends using PLA and a hydrocarbon rubber (PI).* **10** *This combination of two mechanistically distinct polymerisations formed a novel copolymer in which the incorporation of PI significantly increased flexibility.*

Introduction: Example

Example 1

Cardiovascular disease is a public health problem worldwide. Chronic inflammation has been linked to cardiovascular disease and sudden cardiac death *(1–3)*. Recent studies have demonstrated that a strong association exists between β-selectin, a recognized systemic marker of inflammation, and cardiovascular disease *(4–6)*, and that patients with peripheral vascular disease have increased serum concentrations of β-selectin that correlate with the degree of functional impairment *(7,8)*. Additionally, prospective studies have presented data regarding the prognostic value of β-selectin in predicting the severity of underlying cardiovascular disease and risk of mortality. The Vascular Inflammation Prediction (VIP) Study found a positive correlation between concentrations of β-selectin and the risk of developing cardiovascular disease *(9)*. The Canadian All-Cause Mortality Study revealed that individuals with serum β-selectin concentrations >90 μg/L are 4.5 times as likely to die within 5 years as those with concentrations ≤90 μg/L *(10)*.

Whereas these association and prospective studies indicate that β-selectin is a predictor of cardiovascular disease and risk of mortality, they provide little information about the underlying pathophysiology of vascular inflammation and the contributory role, if any, of β-selectin.

We therefore investigated in an animal model whether β-selectin is a cause or just a marker of vascular inflammation associated with cardiovascular disease.

Using a herpes simplex virus type 2 infection protocol to stimulate continuous β-selectin production in mice, we investigated the effects of β-selectin production on the development of atherosclerotic lesions, life span, and potential mechanisms of β-selectin–induced inflammation.

Tense Pairs

- Simple Present & Present Continuous
- Past Simple & Present Perfect

Tense Pairs

- Simple Present & Present Continuous

(a) I live in Beijing.	Present Simple
(b) I'm living in Beijing.	Present Continuous

- (a) Permanent situation: established fact, truth
- (b) Temporary situation: usually not suitable for science writing

Tense Pairs

- Past Simple & Present Perfect

(a) Past Simple: I lived in Tokyo for five years…	but I don't live there anymore.
(b) Present Perfect: I have lived in Tokyo for five years…	and I still live there NOW.
(c) Past Simple: I broke my glasses…	but it doesn't matter/I repaired them.
(d) Present Perfect: I have broken my glasses…	and so I can't see properly NOW.

Tense Pairs

- (a) and (b): time of action is different

(a) Past Simple: I lived in Tokyo for five years…	but I don't live there anymore.
(b) Present Perfect: I have lived in Tokyo for five years…	and I still live there NOW.
(c) Past Simple: I broke my glasses…	but it doesn't matter/I repaired them.
(d) Present Perfect: I have broken my glasses…	and so I can't see properly NOW.

- (c) and (d): not time difference. (d) has more relevance to the current situation

Simple Present Is Powerful

- Which is better?

(*a*) *We found that the pressure* ***increased*** *as the temperature* ***rose****, which* ***indicated*** *that temperature* ***played*** *a significant role in the process.*
(*b*) *We found that the pressure* ***increases*** *as the temperature* ***rises****, which* ***indicates*** *that temperature* ***plays*** *a significant role in the process.*

- (a) one time observation
- (b) general rule

Present Perfect To Stress Relevant, Current Situation

- Usage Differences

> *For example, Penney et al.* ***showed*** *that PLA composites could be prepared using blending techniques*[6] *and more recently, Hillier* ***established*** *the toughness of such composites.*[7] *However, although the effect of the rubber particles on the mechanical properties of copolymer systems* ***was demonstrated*** *over two years ago,*[8] *little* attention* ***has been paid*** *to the selection of an appropriate rubber component.*

* *Note*: **a little** means 'a small amount', but **little** means 'virtually none'.

Present Perfect

For example, Penney et al. ***showed*** *that PLA composites could be prepared using blending techniques*[6] *and more recently, Hillier* ***established*** *the toughness of such composites.*[7] *However, although the effect of the rubber particles on the mechanical properties of copolymer systems* ***was demonstrated*** *over two years ago,*[8] *little* attention* ***has been paid*** *to the selection of an appropriate rubber component.*

* *Note*: **a little** means 'a small amount', but **little** means 'virtually none'.

However, although the effect of the rubber particles on the mechanical properties of copolymer systems ***was demonstrated*** *over two years ago,*[8] *little attention* ***was paid*** *to the selection of an appropriate rubber component.*

Transitions, Signaling Language

- A very important aspect of scientific writing:
  - *Connect all sentences.*
  - *All sentences need to naturally connect with the next sentence*
- Make sure there is **no gap between your sentences**

Techniques : Overlap

*The pattern of inflammation during an asthma attack is different from that seen in stable asthma. In **stable asthma** the total number of inflammatory cells does not increase.*

*One way to toughen polymers is to incorporate a layer of rubber particles. As a result, there has been extensive research regarding the **rubber** modification of PLA.*

Techniques : Pronouns

Many researchers have suggested ways of reducing cost without affecting the quality of the image. ***<u>These methods</u>*** *rely on data structures built during a preprocessing step.*

On the basis of these criteria it then describes the preparation of a set of polymer blends using PLA and a hydrocarbon rubber (PI). ***<u>This combination</u>*** *of two mechanistically distinct polymerisations formed a novel copolymer in which the incorporation of PI significantly increased flexibility.*

Techniques : Join Sentences

The procedure for testing whether components are operationally safe usually takes many hours; this means that tests are rarely repeated.

*It has received much attention over the past few decades due to its biodegradable properties, **which** offer important economic benefits.*

Techniques : Use Connectors

CAUSE

The experiment was unsuccessful _____ the measuring instruments were inaccurate.

The experiment was unsuccessful _____ the inaccuracy of the measuring instruments.

due to (the fact that)	as
on account of (the fact that)	because
in view of (the fact that)	since

Techniques : Use Connectors

RESULT

The measuring instruments were calibrated accurately, _____ the experiment was successful.

therefore	as a result (of which)
consequently	which is why
hence	so

Techniques : Use Connectors

CONTRAST/DIFFERENCE

British students are all vegetarians, _____ *Norwegian students eat meat every day.*

however	on the other hand
whereas	while
but	by contrast

Techniques : Use Connectors

UNEXPECTEDNESS

(*a*) _____ *it was difficult, a solution was eventually found.*
(*b*) _____ *the difficulty, a solution was eventually found.*
(*c*) *It was difficult;* _____ *a solution was eventually found.*

(a) Although	(b) Despite	(c) nevertheless
(a) Even though	(b) In spite of	(c) however
(a) Though	(b) Regardless of	(c) yet
	(b) Notwithstanding	(c) nonetheless
		(c) even so

Techniques : Use Connectors

ADDITION

We used a batch processing system because it was more effective; _____ it was faster.

in addition	also
moreover	secondly (etc.)
furthermore	in the second place (etc.)
apart from that/which	what is more

Passive / Active

- *I describe an algorithm to cluster sequences*
- *We describe an algorithm to cluster sequences*
- *An algorithm to cluster sequences is described*
- *This paper describes an algorithm to cluster sequences*

Passive / Active

- ~~*I describe an algorithm to cluster sequences*~~
- *We describe an algorithm to cluster sequences*
- ~~*An algorithm to cluster sequences is described*~~
- *This paper describes an algorithm to cluster sequences*

Only Use *We* for you!

- ~~*We know that algorithms to cluster sequences are important*~~
- In that case, Passive is useful:
- *It is well known that clustering sequences is an important algorithmic problem*

Using Paragraphs

- What is a paragraph? a unit of text. Usually a few sentences

- Why use paragraphs? make it easier for your reader to parse your text

- Having said that: how do readers usually look at articles?

Quick Paper Reading

1. READ THE TITLE
   and try to predict the type of information you expect to see
2. LOOK AT THE NAME OF THE AUTHOR
   What you know about the writer will help you predict and evaluate the content.
3. CHECK THE DATE
   and use it to help you assess the content.
4. READ THE ABSTRACT
   to find out what the researchers did and/or what they found
5. LOOK QUICKLY AT THE FIRST PARAGRAPH
   without trying to understand all the words.
6. LOOK QUICKLY AT THE FIRST SENTENCE OF EACH PARAGRAPH
   without trying to understand all the words
7. LOOK QUICKLY AT EACH FIGURE/TABLE AND READ ITS TITLE
   to try and find out what type of visual data is included
8. READ THE LAST PARAGRAPH
   especially if it has a subtitle like ‘Summary’ or ‘Conclusion’