

Introduction to Linguistics:

Interdisciplinary aspects: Text Technology

Dafydd Gibbon

B.A. British And American Studies
Basic Module 2

Winter Semester 2006/2007

[Class Website](#)

LINGUISTICS
AND
OTHER DISCIPLINES

THE BACKGROUND TO
YOUR
WEB PORTFOLIO

Web portfolio

- Motivation:
 - easier access and interaction than via paper / email
 - means of becoming familiar with everyday use of electronic media
 - a form of “Applied Text Linguistics”
 - a source of materials / tasks for the class
- So:
 - if you do not have a free web site yet, get one!
 - hint: you can start with a blog (web log) website

Creating a portfolio web site

- My own teaching web site contains a kind of portfolio of my teaching goals, methods and content:
- [Dafydd Gibbon Teaching Winter 2006/2007](#)
- Your own websites can be
 - structured according to your own preferences
 - constructed with web editing software
 - made by hand
 - but they must be professionally formatted, look good, and be easy to use...

Several ways to make a website

- Run your own web server
 - for example on a DSL line, with the Apache server
 - save your HTML files
 - Use the university website
 - and upload your HTML files
 - Use another web service provider
 - and upload your HTML files
 - Use blogging software
 - and make a weblog (blog)
- <My laptop http server>

What is a website, actually?

- A hypertext document with
 - embedded document objects
 - linked document objects
 - and therefore a text...

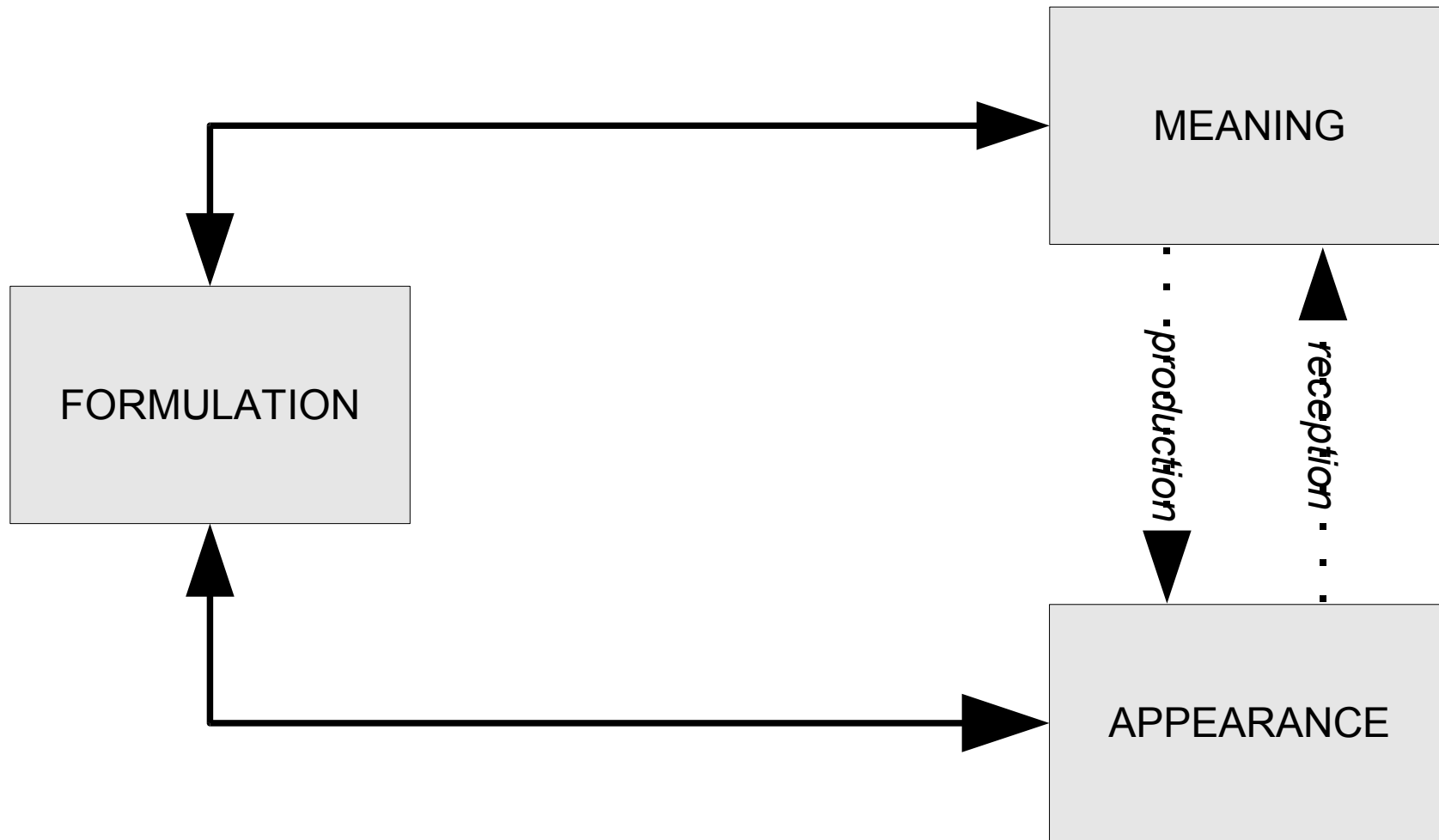
What is a hypertext, actually?

- A hypertext document is a *text*
 - either with conventional hierarchical parts
 - or as a complex network of parts
- For example:
 - Any document on the World Wide Web
 - electronic dictionary
 - blog
 - e-commerce site
 - [Google](#) (and of course this slide, since it is linked...)
 - A help document for a computer application

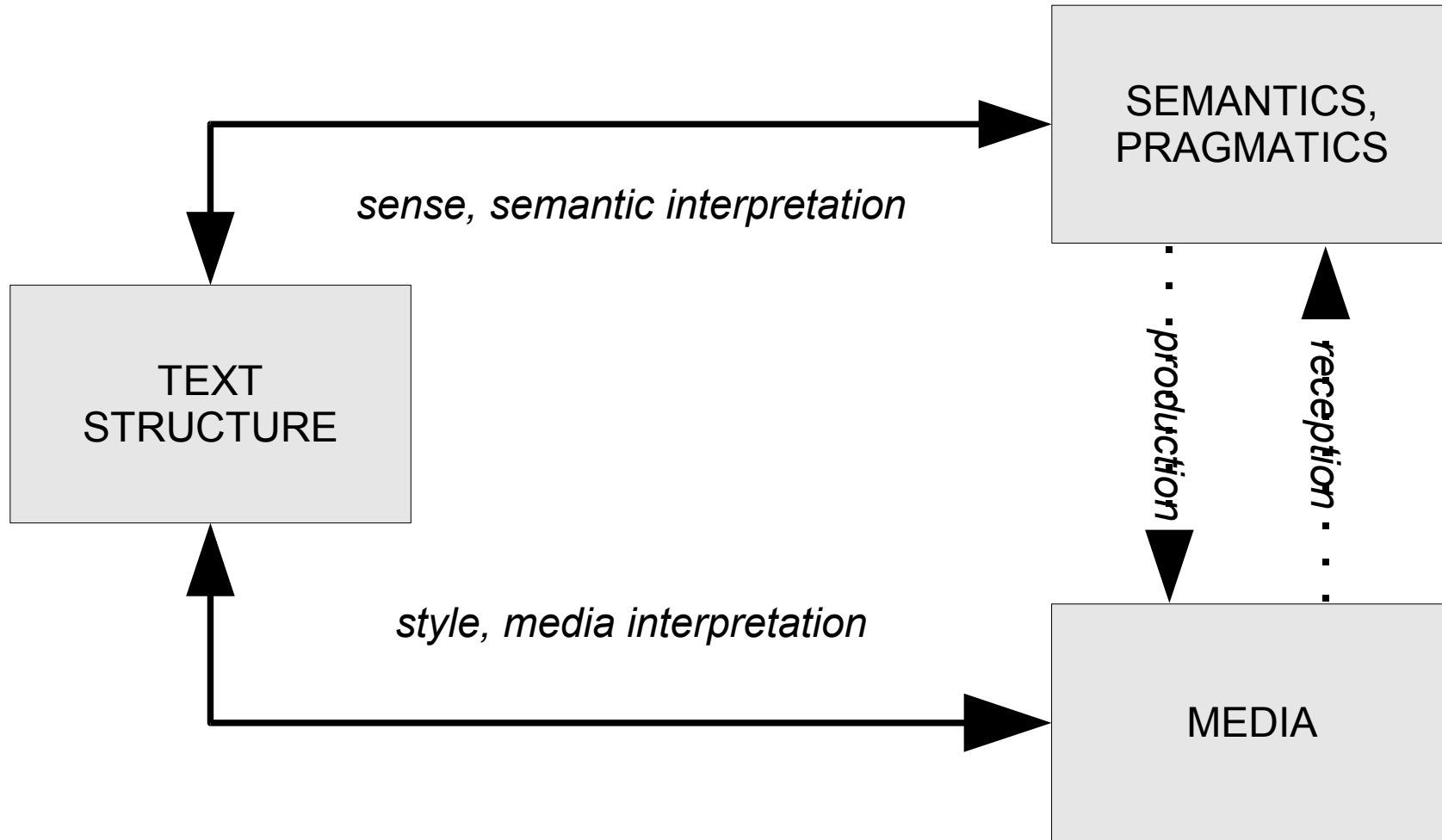
What is a text, actually?

- Discussion:
 - Examples:
 - ...
 - Which properties does a text have?
 - appearance...
 - meaning...
 - structure...
- Consequences for a linguistic theory of text?

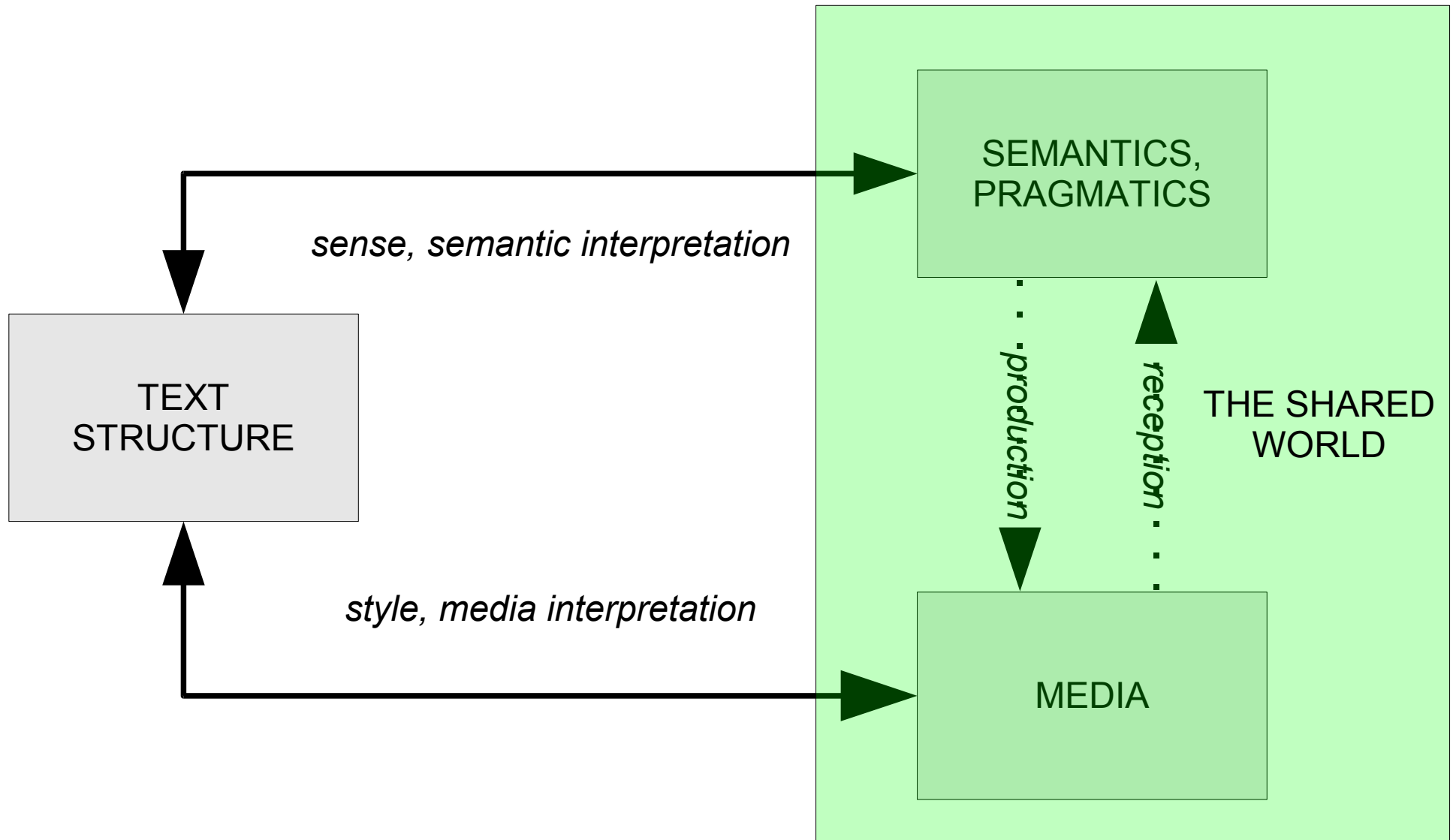
Properties of a text



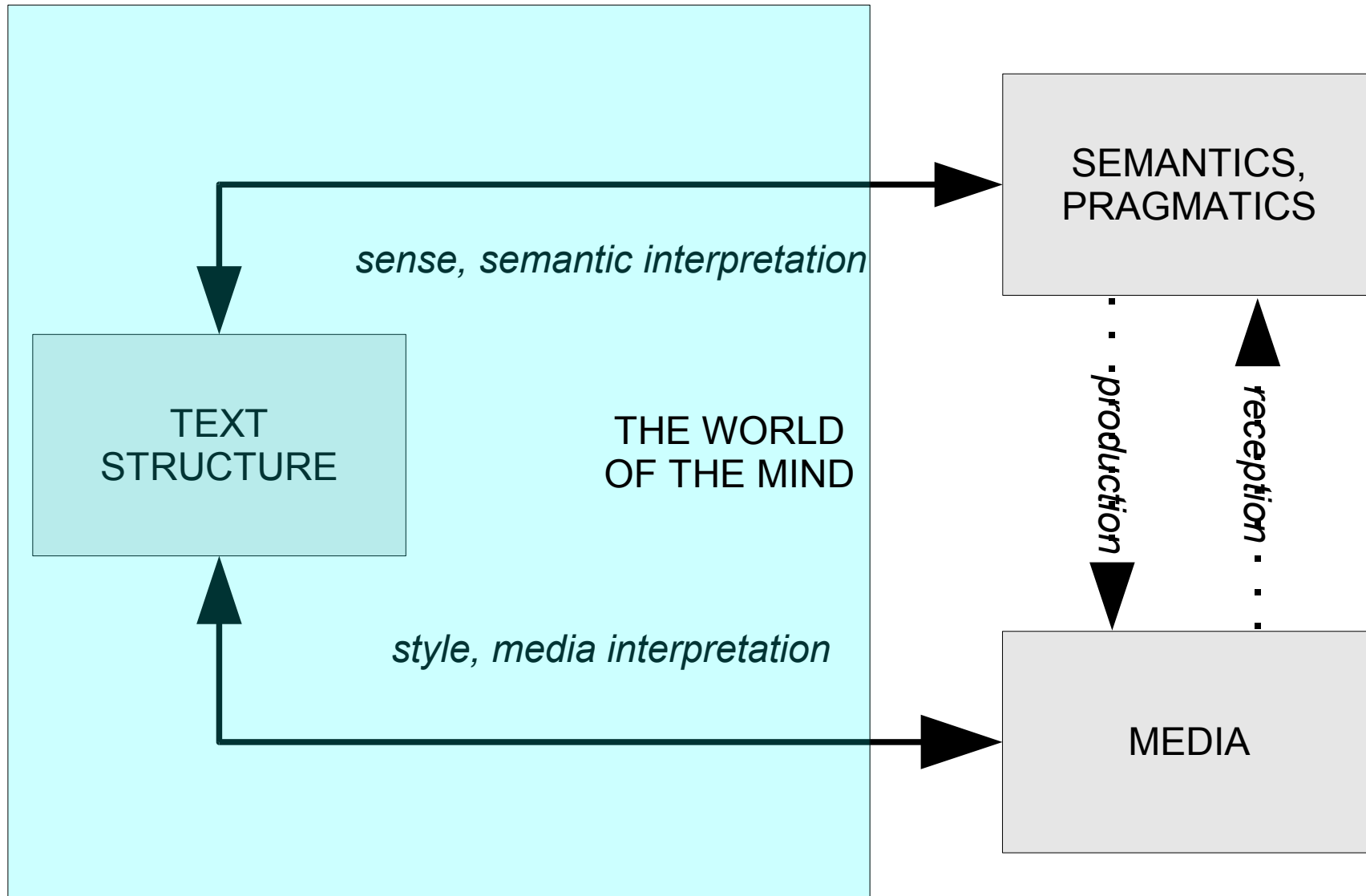
Properties of a text



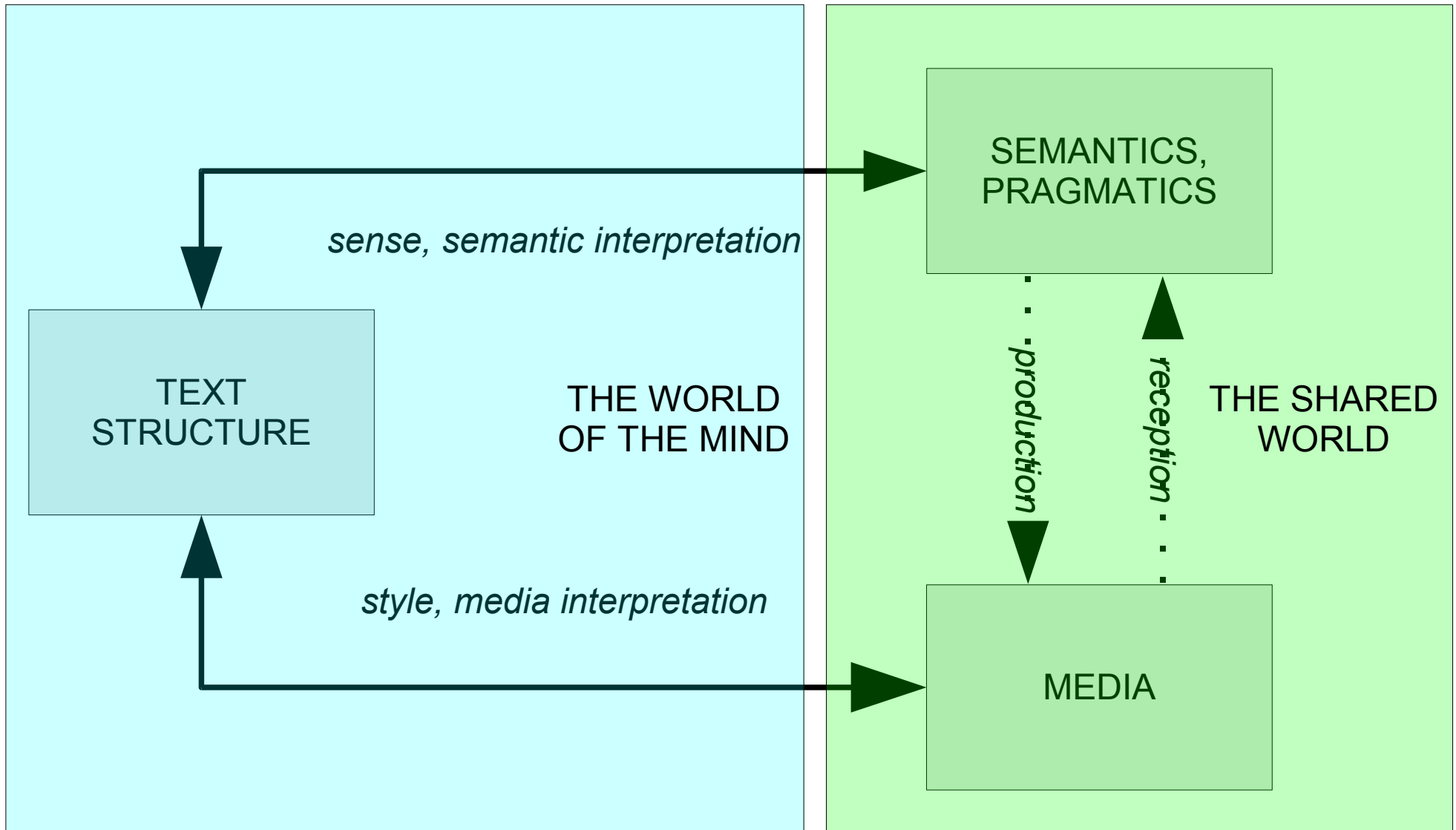
Properties of a text



Properties of a text



Properties of a text



APPLYING TEXT THEORY PROFESSIONALLY

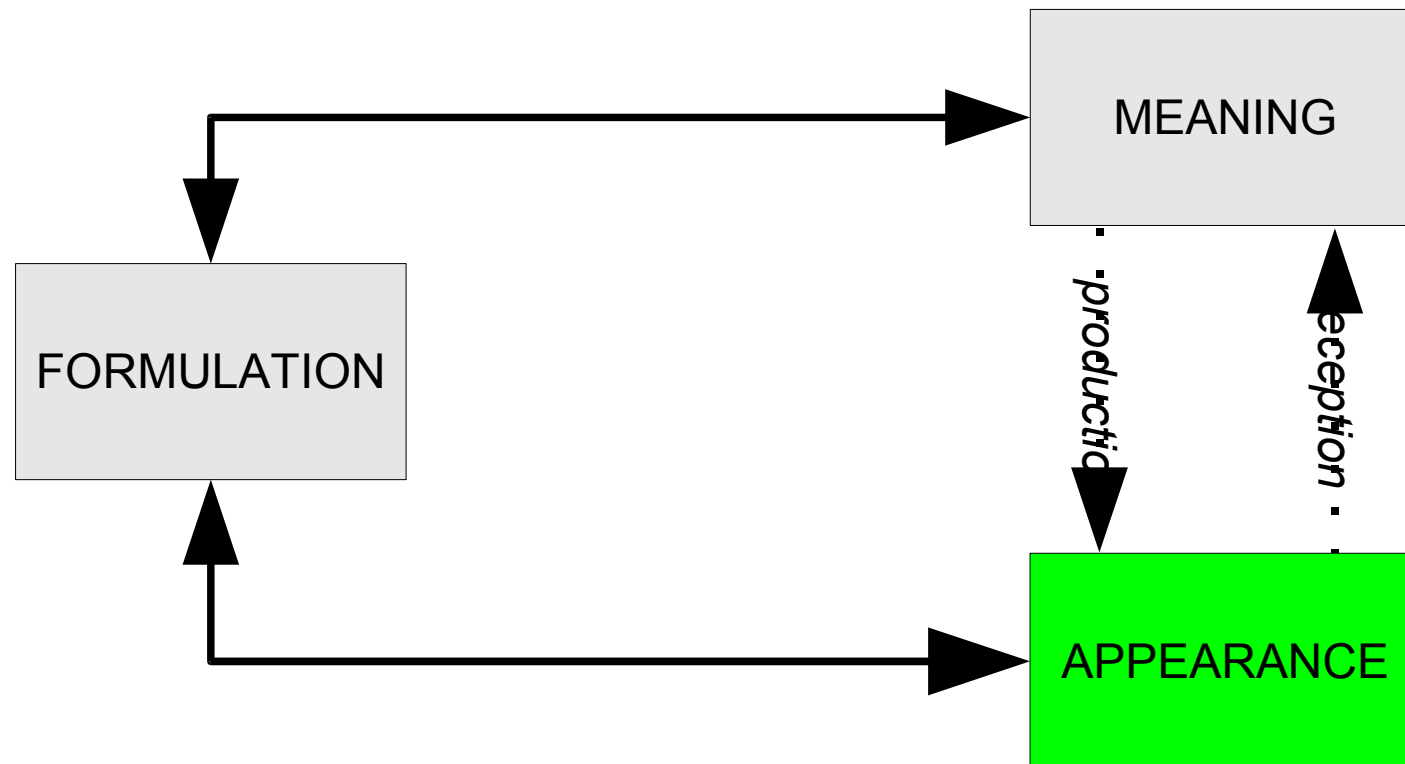
"What a Linguist Needs to Know about Word Processing"

<http://wwwhomes.uni-bielefeld.de/~gibbon/Docs/LinguisticsAndWordProcessing03.pdf>

- Because linguists know about
 - text construction
 - spelling (cf. spell checkers)
 - correct inflection (cf. grammar checkers)
 - thesaurus as a writer's help
 - word prediction/completion (cf. also mobile phones)
 - capitalisation
 - use of correct quotation marks
 - translation of terms for localisation to other languages

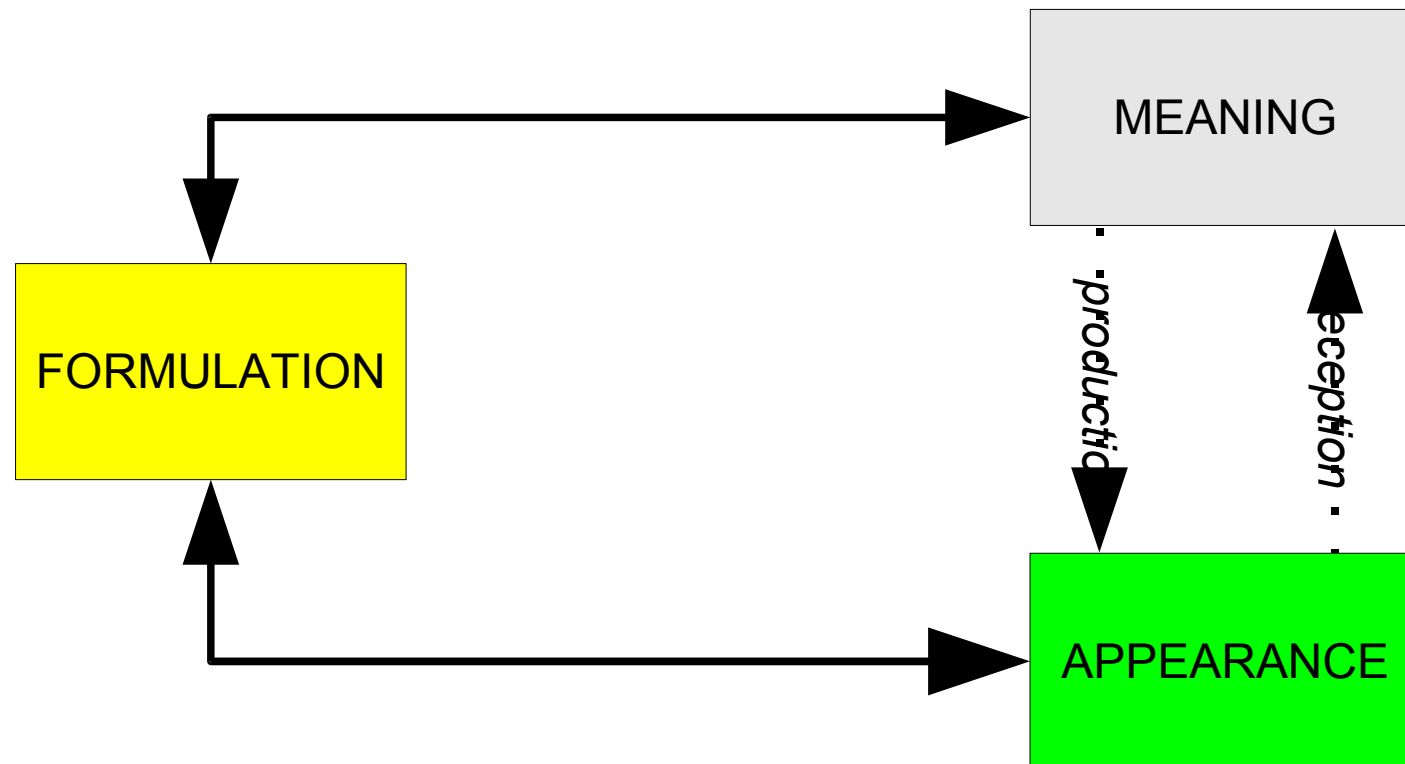
Word Processing

- Word processing, with OpenOffice or MS-Word is an area of *Applied Linguistics*.
- Most people do not know how to do it: they just concentrate on APPEARANCE:



Word Processing

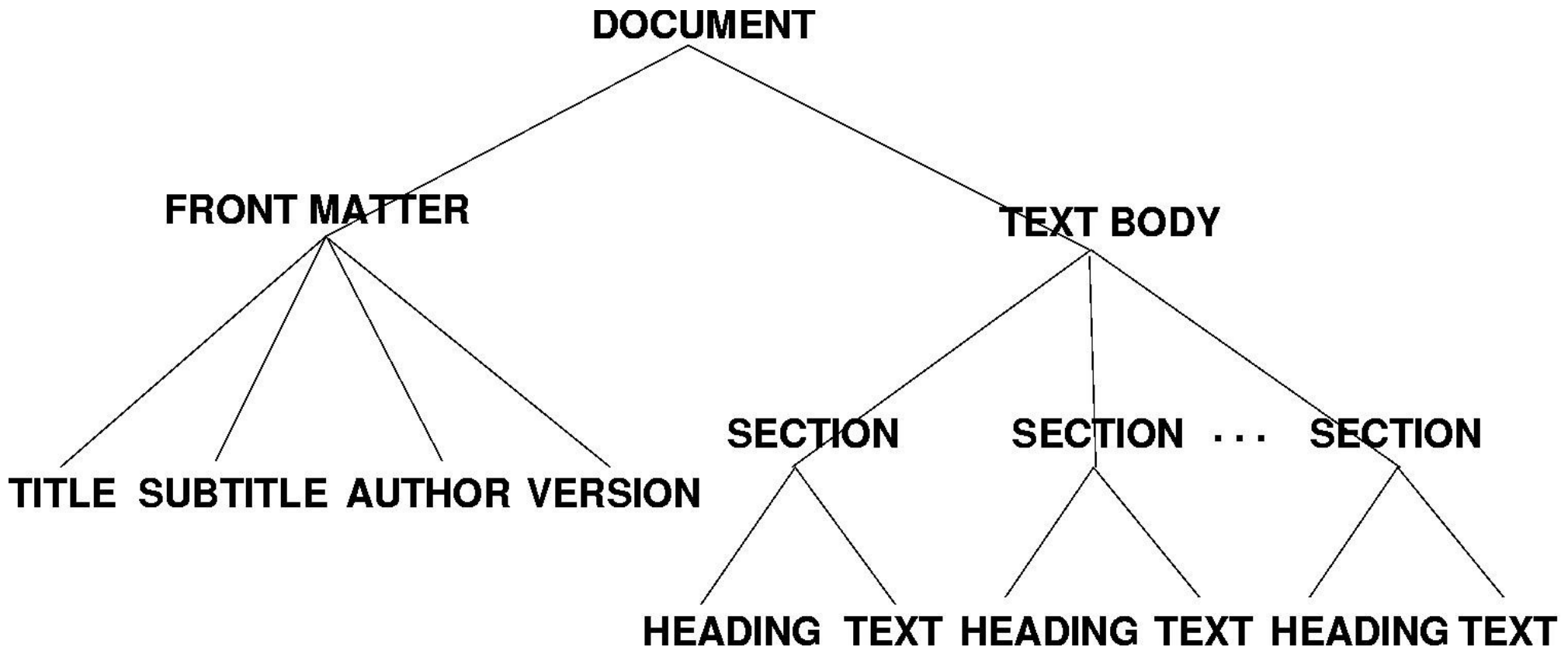
- Word processing, with OpenOffice or MS-Word is an area of *Applied Linguistics*.
- However, much more important is FORMULATION in terms of TEXT OBJECTS:



TEXT OBJECTS
AND
DOCUMENT OBJECTS

TEXT OBJECTS

Formulation: text structure



Word processor text object hierarchy

- Character:
 - Properties: font, size, highlights, ...
- Paragraph:
 - Properties: upper, lower, left, right margins
 - Types:
 - Default (general definitions - don't actually use this!)
 - Text body
 - Headings (different levels of sub-headings)
 - Lists
 - numbered lists (ordered lists)
 - unordered lists (bullet lists)
 - Tables
- Figure

Text objects and their properties

Character	Font	Arial, Helvetica, Times Roman, Courier, ...
	Size	10pt, 12pt, ...
	Bold	bold, non-bold
	Italic	italit, non-italic
	Underline	underline, non-underline
	Colour	red, orange, ... white, black

Text objects and their properties

Paragraph	Alignment	left, right, centred, justified
	Numbering	enumerated, bulleted
	Tabulator	horizontal tab settings
	Indentation	left & right margin, first line indentation
	Spacing	gap above and below
	Line	1, 1.5, 2, ...
	Frame	sub-attributes: line type, thickness, colour, ...

More text objects: Lists

- Lists:
 - properties:
 - line spacing
 - indenting
 - spacing between list marker and text
 - ...
 - types:
 - ordered lists (numbered lists)
 - properties: leading number
 - unordered lists (bullet lists)
 - properties: leading dot / dash / arrow / ...

1. apples
 2. Oranges
 3. Pears
 4. Quinces
- apples
 - Oranges
 - Pears
 - Quinces

More text objects: Tables

<i>Family name</i>	<i>First name</i>	<i>Date of birth</i>	<i>Place of birth</i>	<i>Telephone</i>	<i>Email</i>
Albertson	James	1983-07-23	Norwich, UK	+4942345332	ja@none.co.uk

- Tables
 - parts:
 - header rows
 - properties: spacing, fonts, ...
 - content rows
 - properties: spacing, fonts, ...
 - columns:
 - properties: spacing, ...
 - cells
 - properties: padding, fonts, ...
 - properties: borders, ...

More text objects: Figures

- Figures:
 - parts:
 - picture
 - caption
 - cross-reference
- Advantages:
 - Numbering is automatic
 - Moving the figure changes the numbering

Figure 1 shows the Golden Horn of Gallehus.



Figure 1: The Golden Horn of Gallehus.

Figure 2 shows the ear.



Figure 2: The ear.

PARAGRAPH STYLES

Paragraphs: the unprofessional method

- The unprofessional method:
 - paragraph objects can be “hacked” - i.e. their properties can simply be defined by using the “bold”, “centred” etc. properties, one paragraph at a time
 - but this
 - is a waste of time and energy (and intelligence): changes have to be made separately to each paragraph
 - inflexible
 - leads to inconsistency
 - creates problems with converting into other media (e.g. hypertext for the internet)

Paragraphs: the professional method

- The professional method:
 - *Types* of paragraph objects should be defined as *types*, by means of *styles* (German: Formatvorlagen)
 - and this is
 - time-saving: changes apply instantly to the whole document
 - versatile
 - permits consistent formatting
 - permits easy conversion into other media

Typical paragraph styles

- Default/Standard (predefined - the granddad of all styles)
- Pre-Title (user-defined)
- Title (predefined but modified),
- Subtitle (predefined but modified),
- Author (user-defined),
- Version (user-defined),
- Heading 1 (i.e. level 1, not the first heading; predefined but modified),
- Heading 2 (i.e. subheading of level 2, not the second heading; predefined but modified),
- Text body (predefined but modified),

Style families - property inheritance

- Similar styles may “inherit” properties from more basic or general “parent” styles, which they are “linked” to
- Default - don't use this - its the granddaddy of all styles!!!
 - Title
 - Heading
 - Heading 1
 - Heading 2
 - ...
 - Text body
 - Text body indented
 - LongQuotation
 - Bibliography
 - ...

DOCUMENT OBJECTS

Superimposed Document Objects

- Document
 - Filename
 - ...
- Page:
 - Orientation: portrait, landscape
 - Margins: top, bottom, left, right
- Running titles:
 - Headers
 - Footers
- Fields for insertion into running titles etc.
 - Page number
 - Total number of pages
 - Date
 - ...

Document objects and their properties

Document	Name	relevant metadata
	Author	relevant metadata
	Version	relevant metadata
Page	Size	sub-attributes: size, top, bottom, left, right margins, header, footer, frame, ...
	Margin	sub-attributes: top, bottom, left, right margin
	Header	special page-linked paragraph type
	Footer	special page-linked paragraph type

Homework

- What are text objects?
 - Name two, and give typical properties.
- What are document objects?
 - Name one and give typical properties.
- What are paragraph styles?
 - Name two types of paragraph
 - For each type, give their typical properties