Modelling Digital Editing

Elena Pierazzo

Université Stendhal Grenoble 3

Editing

- To prepare someone else's text for publication
- Texts in the wild do not look like the text in your printed paperback
- Language, writing habits and conventions have changed over time
- Texts exists in multiple versions, all different
- Which one is the "correct" one?

For example

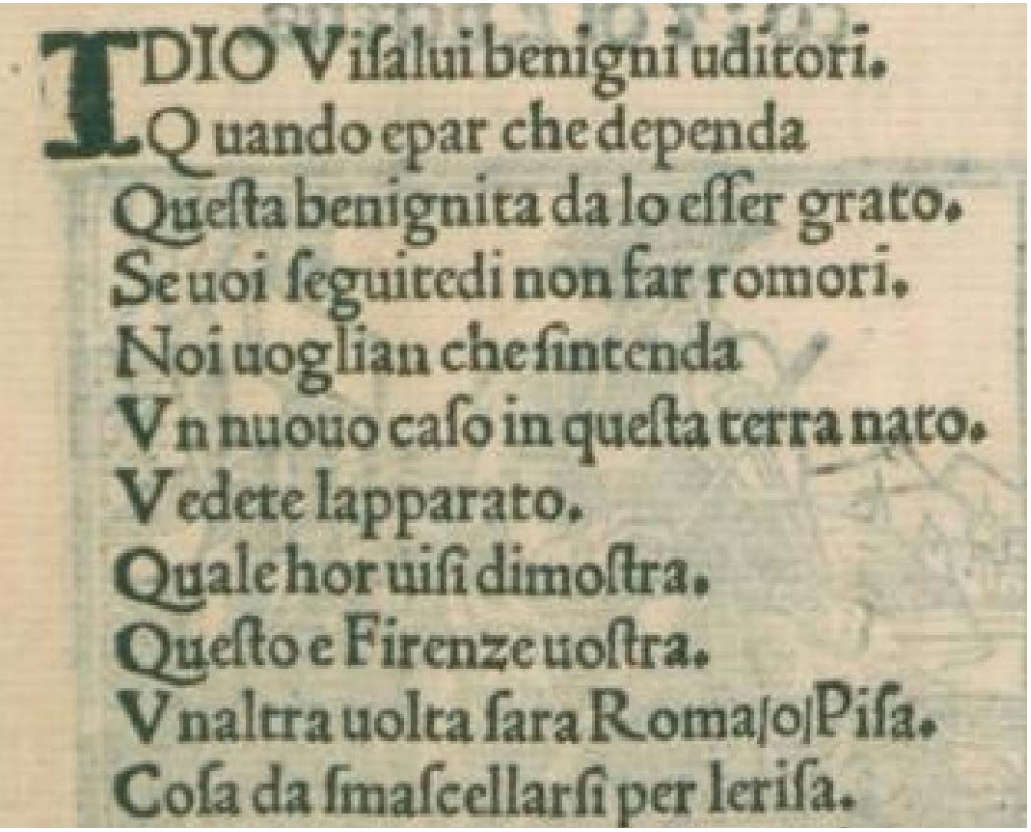
- For Italian:
  - Accents and apostrophes introduced in 1501; generalized in 1530s; regularized in 1870s
  - Spacing of words: gradually introduced from the XI century; generalized in 1530s; regularized in 17th century
  - Abbreviations: came from Latin writing habits; disappeared in early 20th century

Crisis of the alphabet

Did not exist in Latin:

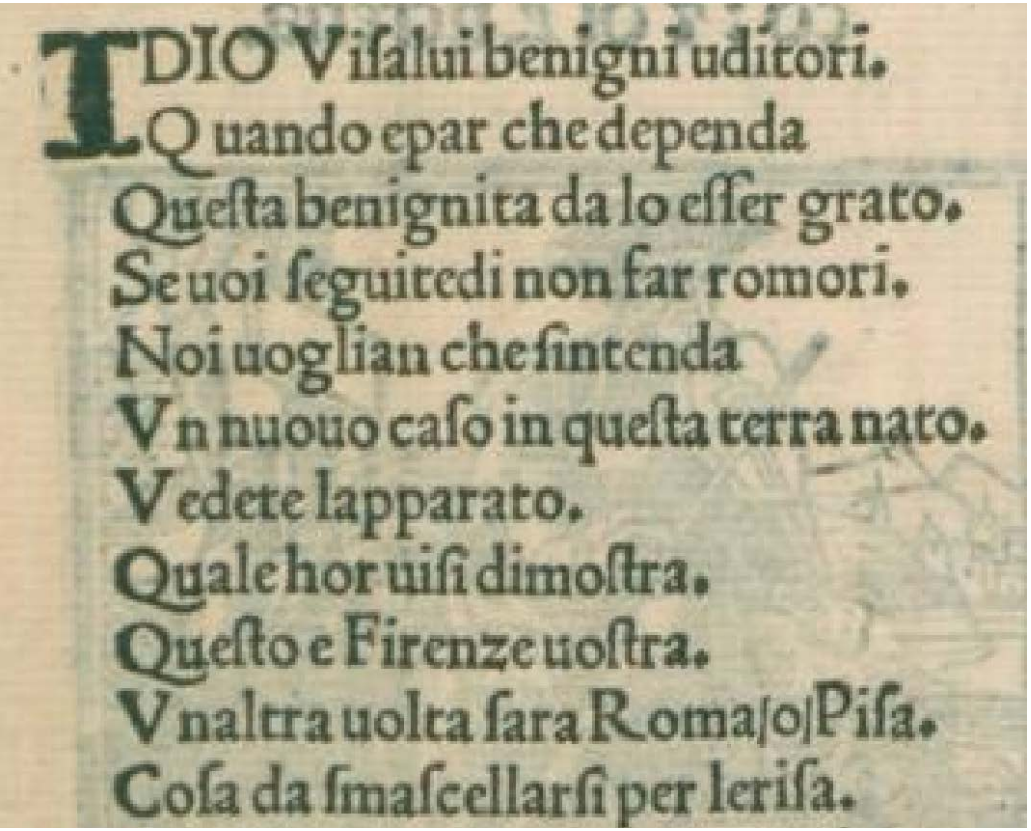
- sc: sh, ch, scs, sch
- j: dh, gj, gh, hgh
- ci: çi, ci
- ch: ki, ch, ki, hc
- gn: ngn, ñ, nj
- gl: lgl, ł, łł, lj
- z: ti, cti, // dj, dz
- v/U

Machiavelli, *La Mandragola*


IDIO Viſalui benigni uditori.
Quando epar che dependa
Queſta benignita da lo eſſer grato.
Se uoi ſeguitedi non far romori.
Noi uoglian cheſintenda
Vn nuouo caſo in queſta terra nato.
Vedete lapparato.
Quale hor uiſi dimoſtra.
Queſto e Firenze uoſtra.
Vnaltra uolta ſara Roma|o|Piſa.
Coſa da ſmaſcellarſi per leriſa.

IDIO Viſalui benigni uditori.
Qunado epar che dependa
Questa benignita da lo esser grato.
Seuoi seguite di non far romori
Noi uoglian cheſintenda
Vn nuouo caſo in questa terra nato.
Vedete lapparato.
Quale hor uiſi dimostra.
Questo e Firenze uostra.
Vnaltra uolta ſara Roma|o|Pisa.
Coſa da ſmaſcellarſi per leriſa.

Machiavelli, *La Mandragola*


IDIO Vi salui benigni uditori.
Quando e par che dependa
Questa benignita da lo esser grato.
Se uoi seguite di non far romori.
Noi uoglian che sintenda
Vn nuouo caso in questa terra nato.
Vedete lapparato.
Quale hor uisi dimostra.
Questo e Firenze uostra.
Vn altra uolta sara Roma o Pisa.
Cosa da smascellarsi per le risa.

Idio vi salvi benigni uditori,
Quando e' par che dependa
Questa benignità da lo esser grato;
Se voi seguite di non far romori,
Noi voglian che s'intenda
Un nuovo caso in questa terra nato.
Vedete l'apparato
Quale or vi si dimostra:
Questo è Firenze vostra,
Un'altra volta sarà Roma o Pisa,
Cosa da smascellarsi per le risa.

Machiavelli, *La Mandragola*

IDIO Viſalui benigni uditori.
Quando epar che dependa
Questa benignita da lo esser grato.
Seuoi seguite di non far romori
Noi uoglian cheſintenda
Vn nuouo caſo in questa terra nato.
Vedete lapparato.
Quale hor uiſi dimostra.
Questo e Firenze uostra.
Vnaltra uolta ſara Roma|o|Pisa.
Coſa da ſmaſcellarſi per leriſa.

Idio vi salvi benigni uditori,
Quando e' par che dependa
Questa benignità da lo esser grato;
Se voi seguite di non far romori,
Noi voglian che s'intenda
Un nuovo caso in questa terra nato.
Vedete l'apparato
Quale or vi si dimostra:
Questo è Firenze vostra,
Un'altra volta sarà Roma o Pisa,
Cosa da smascellarsi per le risa.

TOTAL: 33 intervention on 57 words

Digital Editing

- New heuristics, new models
- Facsimile, transcriptions of witnesses, automatic collation, cladistic method
- Digital archives: Homer Multitext, Whitman archives
- Genetic Archives: Becket, Melville etc.
- We need new theories of editing!

Modelling

Epistemological activity: we take a domain, we analyse it, we select the features we are interested in, we study their relationships, we classify them and we use these to create a model of such domain.

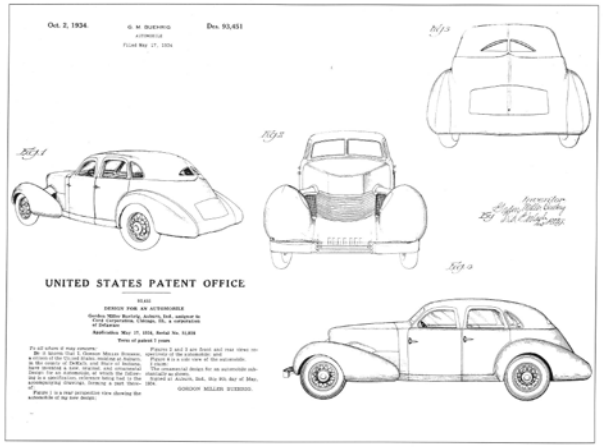
To model is to **simplify**.

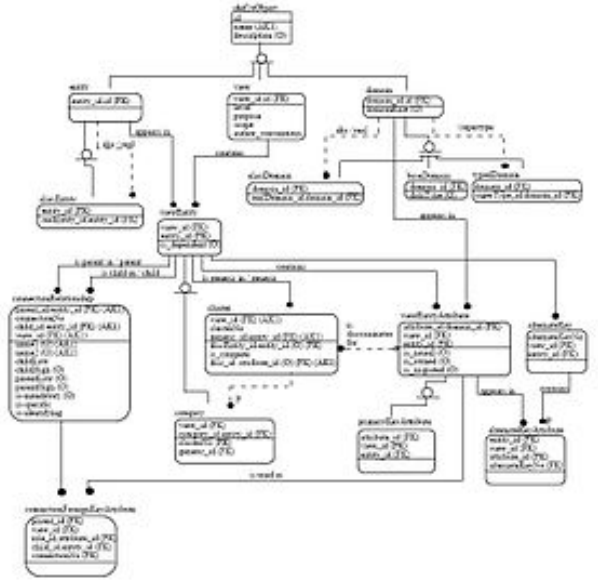
Model-of: to model a domain in order to study it, to learn from it

Model-for: to model a domain to build something new

Models (McCarty)

- “a representation of something for purposes of study, or a design for realizing something new”
- “smaller than the object”
- “the fundamental dependence of any computing system on an explicit, delimited conception of the world or ‘model’ of it”

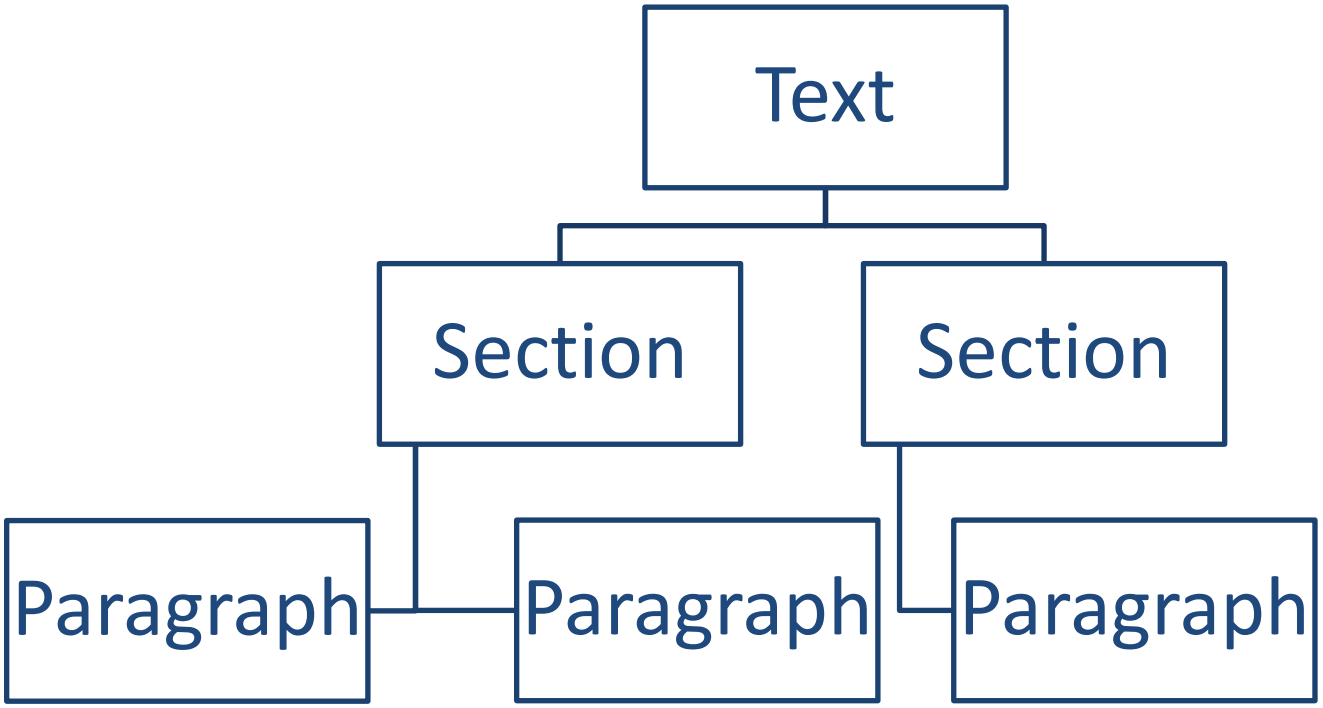




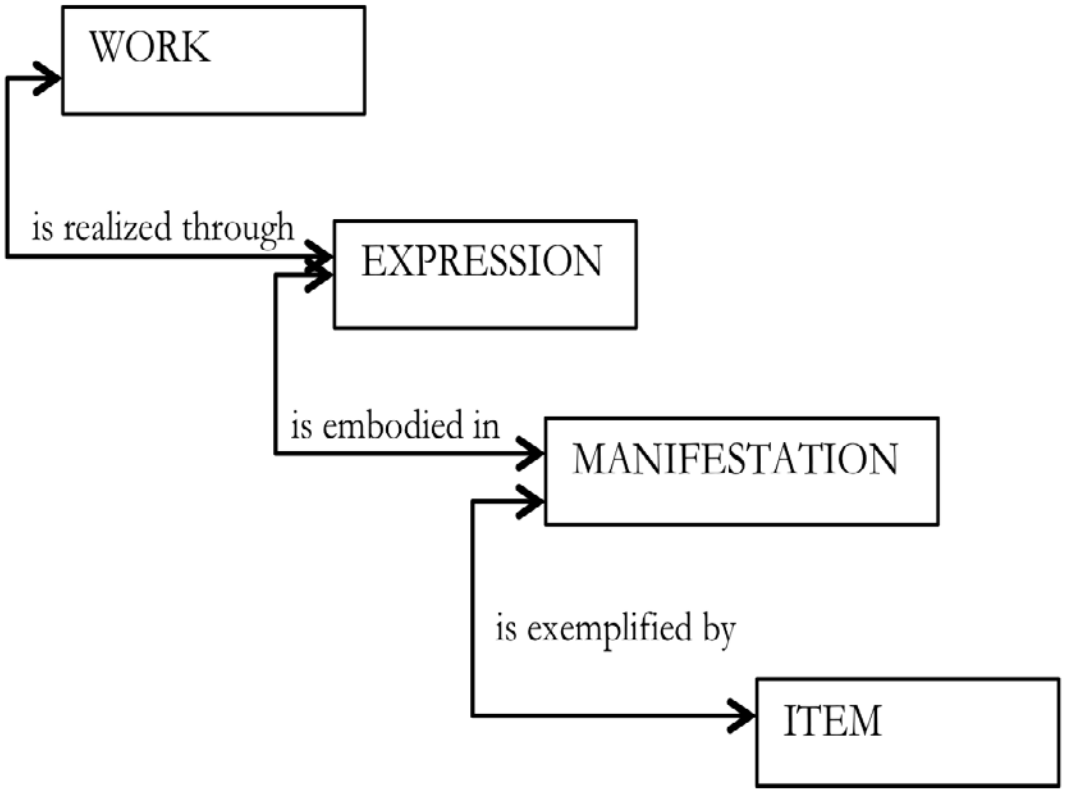
Models and texts

- Texts are models : they present an organised selection of some of the infinite facts that can be derived from documents (Sperberg-McQueen)
- How are texts-as-models created? Can we model the modelling process necessary to create texts as models?
- Existing Models: OHCO, FRBR

Describing the content: OHCO



Describing relations among bibliographical entities: FRBR



A model of: D.E.M.

- Digital Editing Model
- Modelling the documents/texts/works from an editorial point of view
- Modelling the editorial act

Entities

- Document
- Text
- Iterpretants
- Readers
- Author
- Editors
- Editions
- …

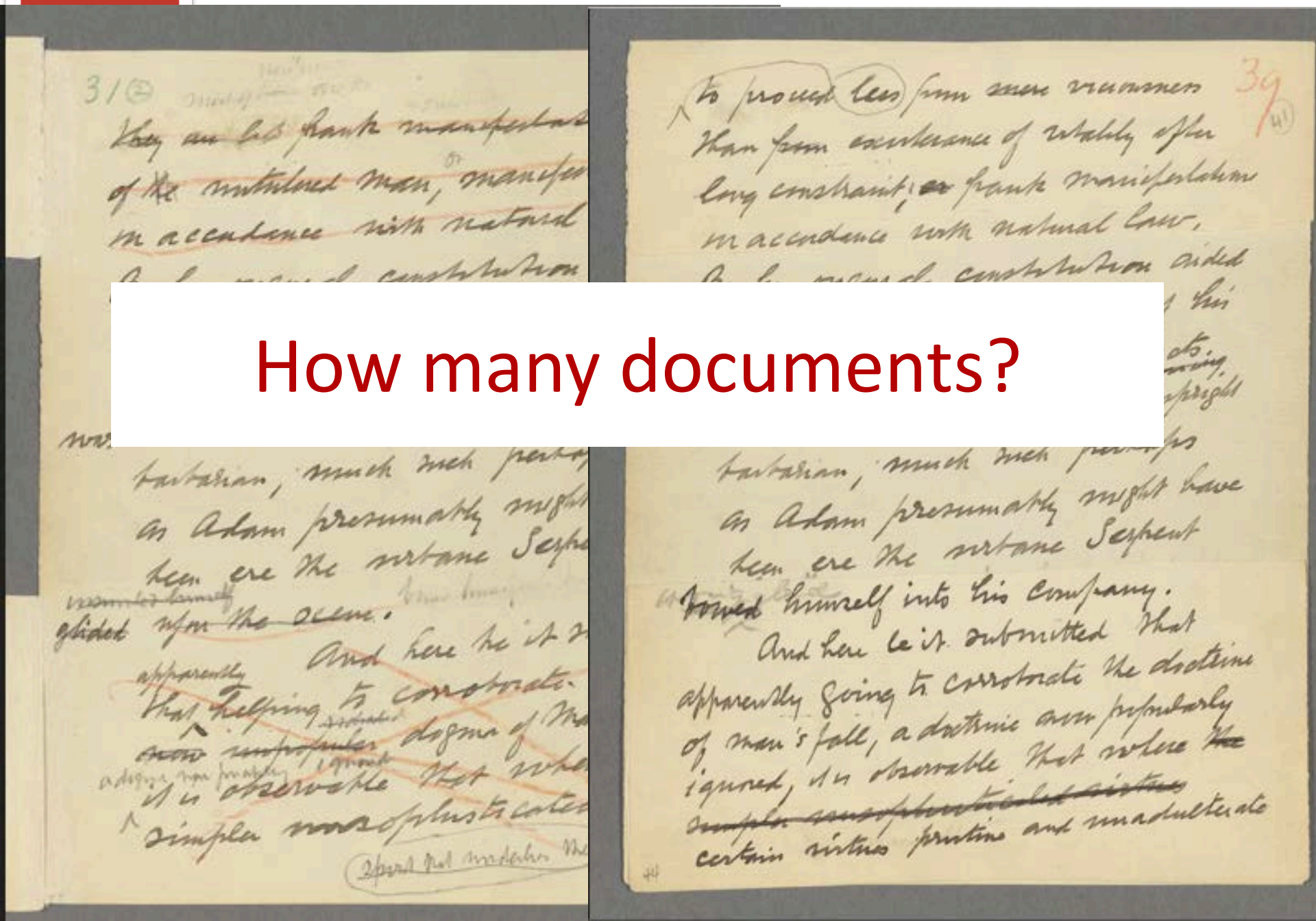
Documents

A document is defined by three sets of characteristics:

- It has, or had in the past, some physical dimensions which can be measured (length, weight, number of leaves, so on).
- It can, or could in the past, be found somewhere.
- It has, or had in the past, some signs on it that can be recognised by a competent reader as words written in a language.

Documents

- Have a physical substance
- Are also abstractions and editorial decisions



How many documents?

Verbal Text Bearing Objects

A document may contain loads of stuff: words, jots, stains, music, illuminations, dead flies…

However it HAS (for this model only!) to contain words, the rest is optional

Facts and dimensions

- Documents are made of facts: dimension, decoration, material, marks, signs, colours, words, punctuation, strokes, stiches, gatherings, folios, pages, weight, drop capitals…
- Facts are grouped according to any organizing principles
- Each reader selects and groups such facts in different ways
- We call these groups **dimensions**

Some dimensions

- Linguistic
- Semantic
- Graphematic
- Literary
- Genetic
- Iconic
- Artistic
- Historical
- Codicological
- Cultural

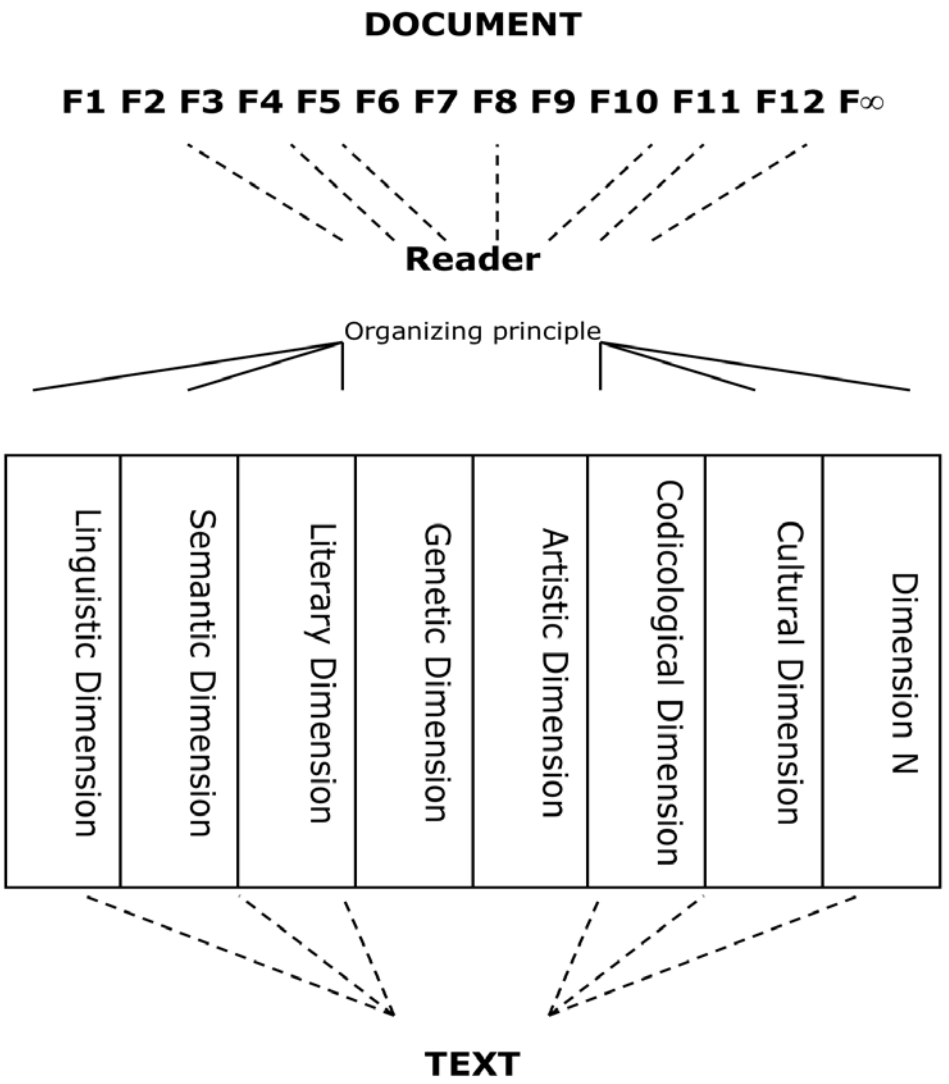
Any particular point of view / purpose / organising principle will select different sets of facts grouping them in different sets of dimensions

Texts and model

- Texts are made up by Readers by selecting some facts from the infinite set, organising them according to their needs/preferences/purposes, making editorial assertions
- A Text is a model that includes the verbal dimension (TM)
- Models that do not include the verbal dimension are not text*, they are other types of models (“document’ model” or DM)

*Remember! Our point of view is the point of view of an editor

First step: texts and documents

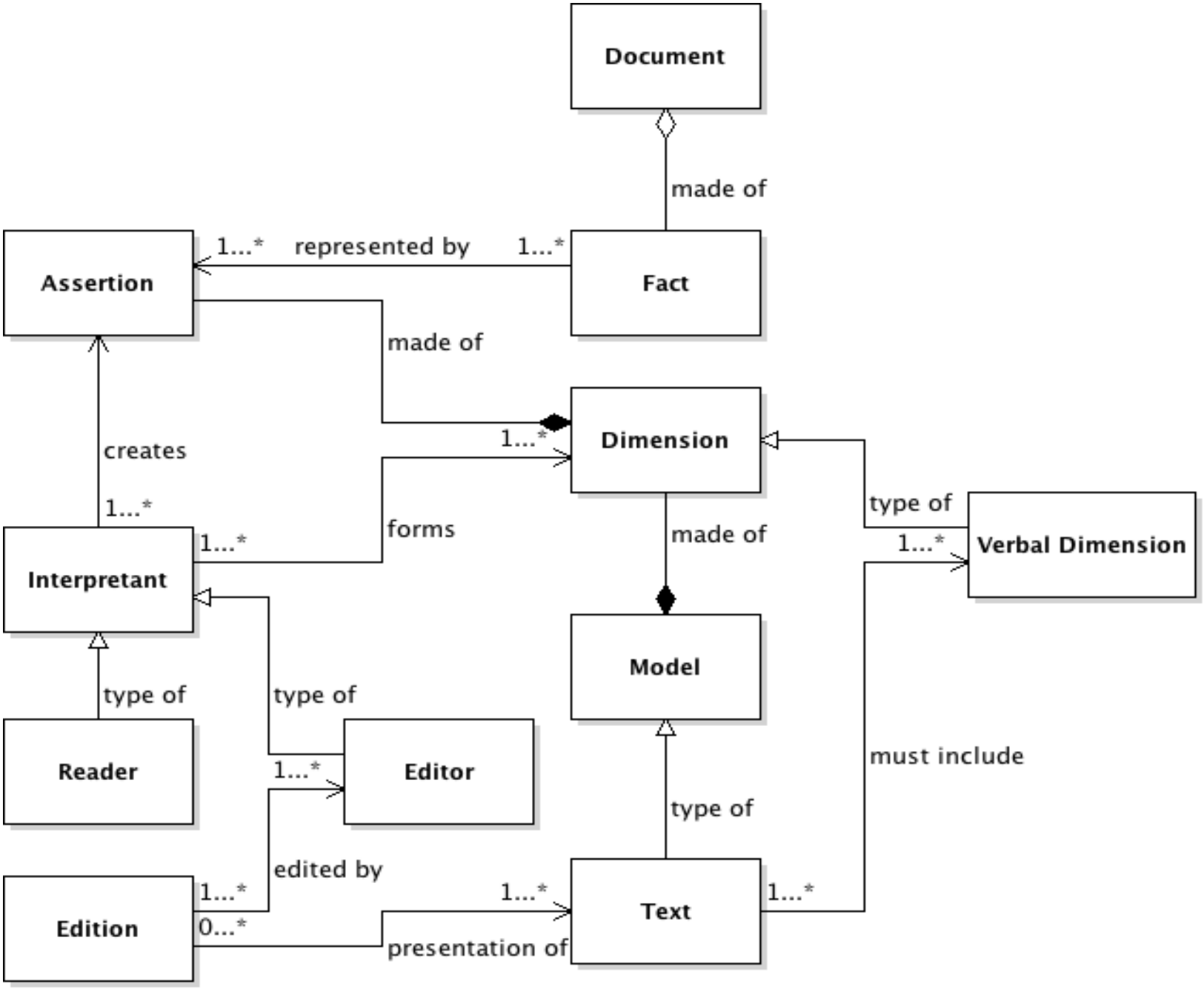


Generalising and specialising

- Editors are specialised readers
- …and so are paleographers, art historians, codicologists, etc.
- We call them **interpretants**
- Not all ‘interpretants’ actually read the words
- Not all ‘interpretants’ read the words only
- Interpretants can be interested in different facts, in any combination

The Formalism: UML

- Unified Modeling Language
- Purpose of representing in a graphic:
  - Any activities (jobs)
  - Individual components of the system
    - And how they can interact with other software components.
  - How entities interact with others (components and interfaces)
  - External user interface


Document
made of
Assertion
1...*
represented by
1...*
Fact
made of
creates
1...*
1...*
Dimension
1...*
forms
1...*
Interpretant
made of
type of
1...*
Verbal Dimension
Model
type of
type of
Reader
1...*
Editor
must include
type of
edited by
1...*
Edition
0...*
1...*
Text
1...*
presentation of

Documents(s)/Text(s)/Version(s)

- Many documents can transmit more or less* the same** text

* More or less/**same: how many variants are allowed before we can speak of a different text?

*** how many variants before we can speak of a different version?

A new entity: the work

When we (=editors) state that

- many documents present *the same text* even with variants;
- and/or a text exists in many versions;

We recognise the existence of a **super-entity** with respect to the one of the documents and of the text: the **work**

The work

- An editorial statement: “a regulative principle” (Eggert)
- The abstract “idea” that determines if many documents present more or less the same* text
- The sum of all documents that present the same* text
- The sum of all texts derived from the documents that present the same* text
- The sum of all texts resulting from the combination of texts derived from multiple documents that present the same* text

*same: editorial statement; in spite of some “minor” differences

Model-text and model-critical text

- Text Model (TM): textual model derived from one document (edition given one witness)
- Critical Text Model (CTM): textual model produced by combining facts derived from different documents (edition given multiple witnesses)

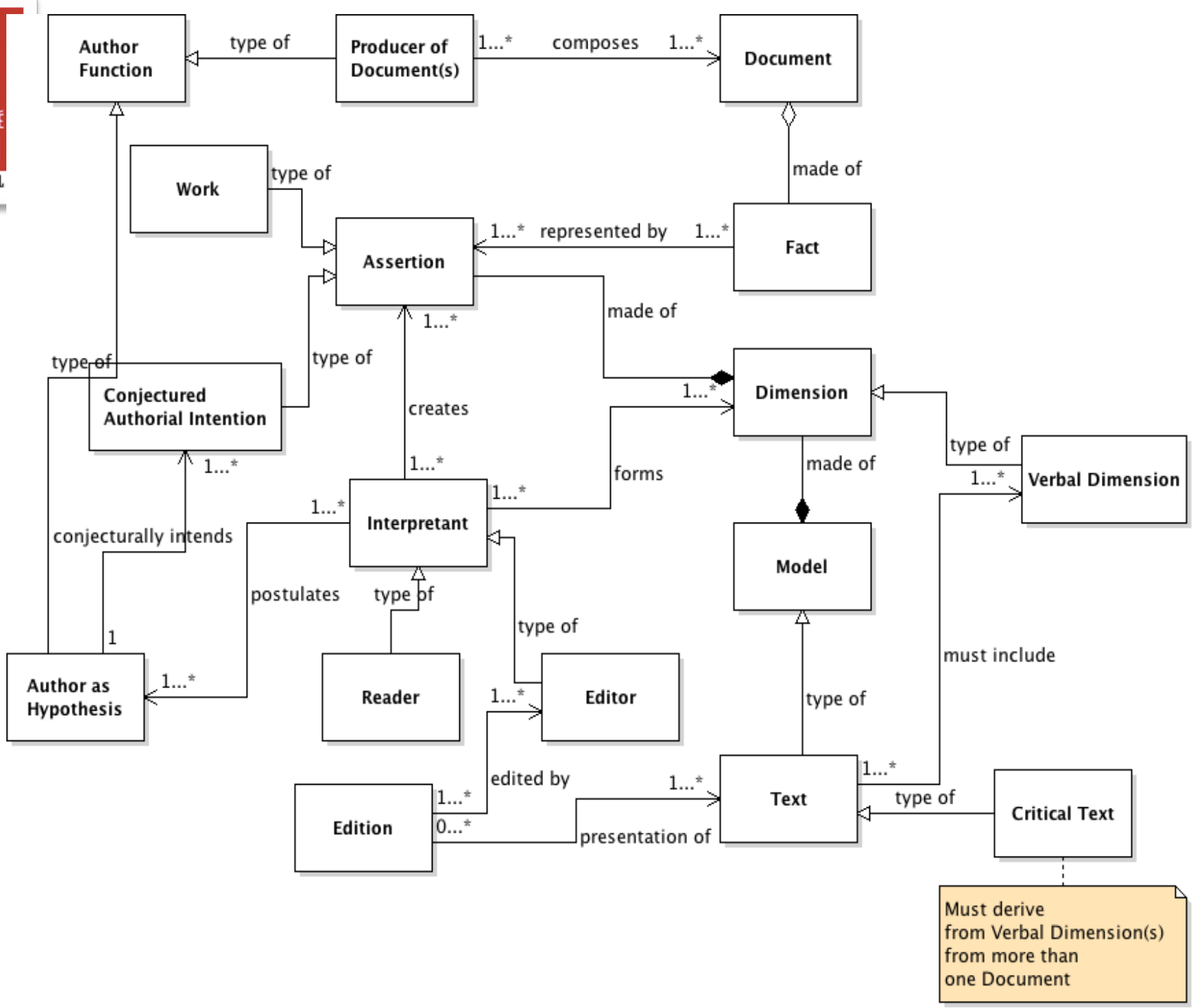
Where's the author?

Good question…

- Do we need one? Yes and no
- The Author (function) is necessary to produce the first text (document?)
- The Author intention, the one that produced the work as an idea, is necessary only if you want it

Versions

- Works exist in versions, at least one (Shillingsburg)
- A version represents a stage of composition of the work
- Such stage is considered more or less conclusive by the author at some point
- A version is an authorial/editorial statement
- A version is preserved in document(s), which are made of facts, etc. and behaves as a sub-work


Author Function
type of
Producer of Document(s)
1...*
composes
1...*
Document
made of
Work
type of
Assertion
1...*
represented by
1...*
Fact
made of
1...*
type of
type of
Conjectured Authorial Intention
creates
1...*
Dimension
type of
Verbal Dimension
1...*
1...*
1...*
forms
made of
1...*
1...*
Interpretant
conjecturally intends
Model
postulates
type of
type of
1
must include
Author as Hypothesis
1...*
Reader
1...*
Editor
type of
edited by
1...*
1...*
Text
type of
Critical Text
Edition
1...*
0...*
presentation of
Must derive
from Verbal Dimension(s)
from more than
one Document

PART 2: MODELLING TEXTUAL TRANSMISSION

Textual transmission determines variations

Scholars at the Library of Alexandria discovered that if you compare (‘collate’) any two copies of any given work, these were different. Very.

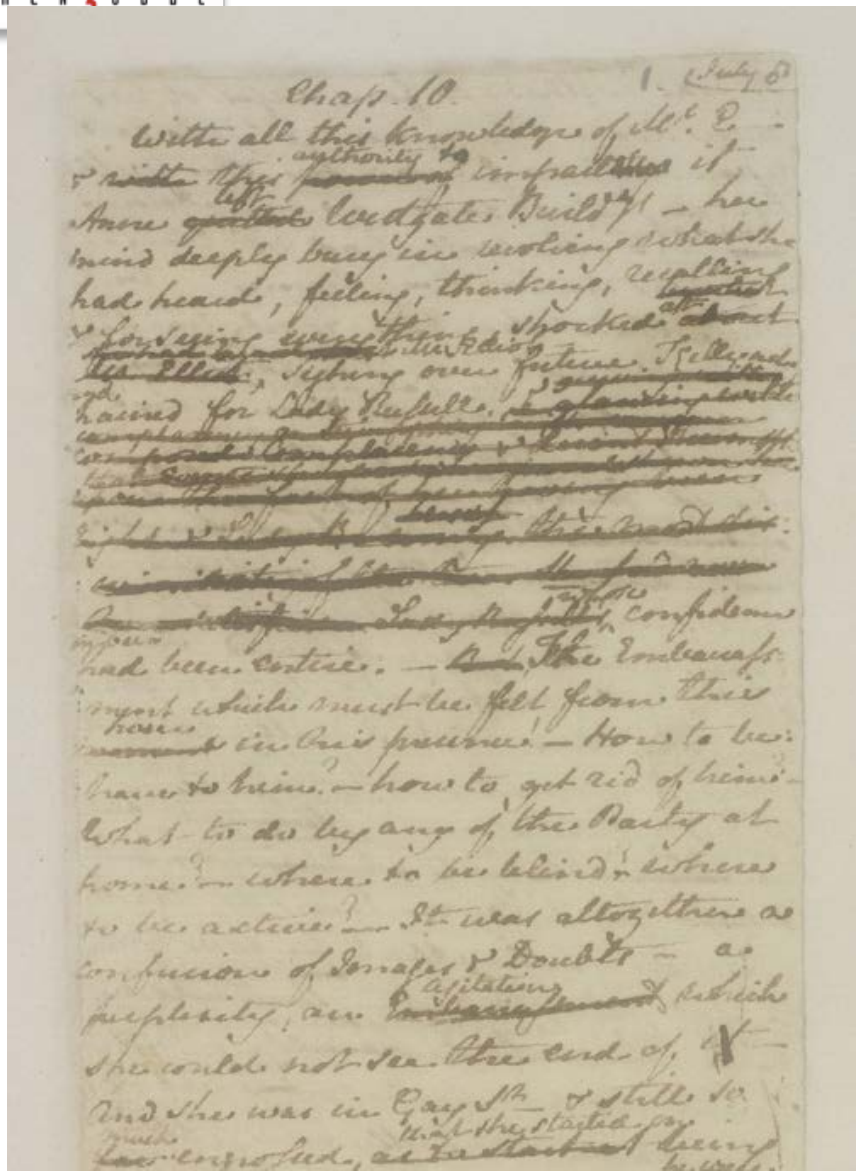
Editorial theories and methods

- Provide different ways on how to make sense of the variation
- Provide different definition of what “correct” means
- Defined by the role they assign to the physical support of texts, i.e. the text-bearing objects, i.e. documents

Variation as a fact

- Variation is not an accident of textual transmission
- Variation is a fact: you cannot transmit text without introducing variants
- (Cerquiglini, *Eloge de la Variante*)
- Can we model variation? Which are the factors that introduce variation? Are all variation “errors”?

Are these the same “thing”?



Chapter 10

With all this knowledge of Mr Elliot and with this authority to impart it Anne left Westgate Buildings, her mind deeply busy in revolving what she had heard, feeling, thinking, recalling and foreseeing everything, shocked at Mr Elliot, sighing over future Kellynch and pained for Lady Russell whose confidence in him had been entire. The embarrassment which must be felt from this hour in his presence! How to be have to him? How to get rid of him? What to do by any of the party at home? Where to be blind? Where to be active? It was altogether a confusion of images and doubts

At a certain level of abstraction they are the same thing

The “thing” is what we call the **text**

Can we separate text from its support?

Our entire culture is based on this separation.

Text transmission as a communication act

CODE

Source → SENDER → Channel → RECEIVER → Destination

Message *Noise* Message

Shannon Weaver 1948: variation is due to the **noise** of the channel

What is a Text?

A text is a **linguistic architecture** that conveys a **meaning** which is potentially **understandable** to at least one group of **receivers** which have the capabilities to decipher the **code** in which the message is encoded.

CODE

Source → SENDER → Channel → RECEIVER → Destination

Message *Noise* Message'

- In theory one can re-encode the text into another code
- In theory one can use another channel
- In theory all of this can be done **without loss of information**

CAN WE USE THIS MODEL TO DISCUSS TEXTS FOUND IN MANUSCRIPTS?

The source(s)

- The source or the sources: the Shannon Weaver model imply one single source
- Textual transmission otherwise…

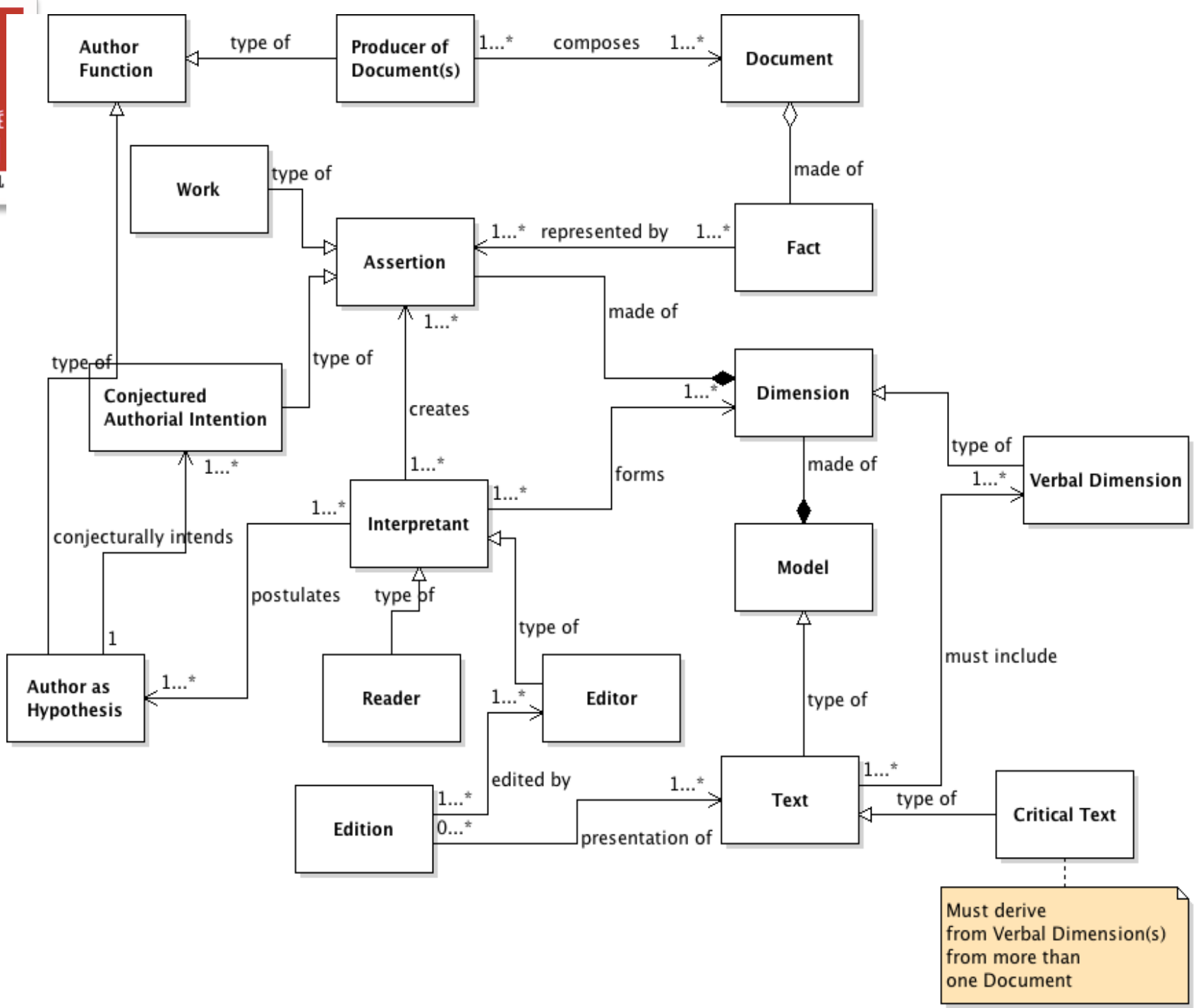
CODE

Sources → SENDER → Channel → RECEIVER → Destination

Message *Noise* Message'

The message (text) is a model

- A model is the result of selection of features from a given domain
- A text is a model because it is composed of a selection of facts from one or many documents
- A model for texts


Author Function
type of
Producer of Document(s)
1...*
composes
1...*
Document
made of
Work
type of
Assertion
1...*
represented by
1...*
Fact
made of
1...*
type of
type of
Conjectured Authorial Intention
creates
1...*
Dimension
type of
Verbal Dimension
1...*
1...*
forms
made of
1...*
1...*
Interpretant
conjecturally intends
Model
postulates
type of
type of
1
must include
Author as Hypothesis
1...*
Reader
1...*
Editor
type of
edited by
1...*
Text
1...*
type of
Critical Text
Edition
1...*
0...*
presentation of
Must derive
from Verbal Dimension(s)
from more than
one Document

The code

- Language
- Grammar
- Syntax
- Rhetoric
- Orthography
- Writing system, conventions…

Without loss?

CODE′ CODE CODE′′

Source → SENDER → Channel → RECEIVER → Destination

Message ***Noise*** **Message'**

Without loss?

PAROLE' LANGUE PAROLE''

Source → SENDER → Channel → RECEIVER → Destination

Message ***Noise*** **Message'**

Saussure, Cours de Linguistique Générale, 1916

Without loss?

CODE’ CODE CODE’’

Source → SENDER → Channel → RECEIVER → Destination

Message ***Noise*** **Message’**

Without loss?

CODE′ CODE CODE″

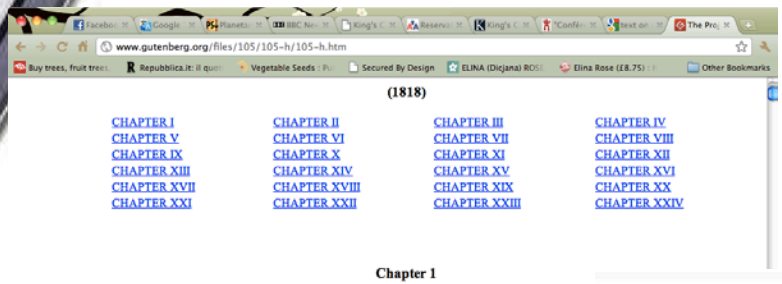
Source → SENDER → Channel → RECEIVER → Destination

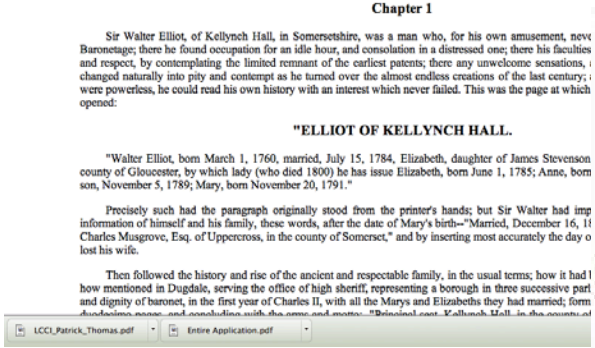
Message ***Noise*** **Message'**

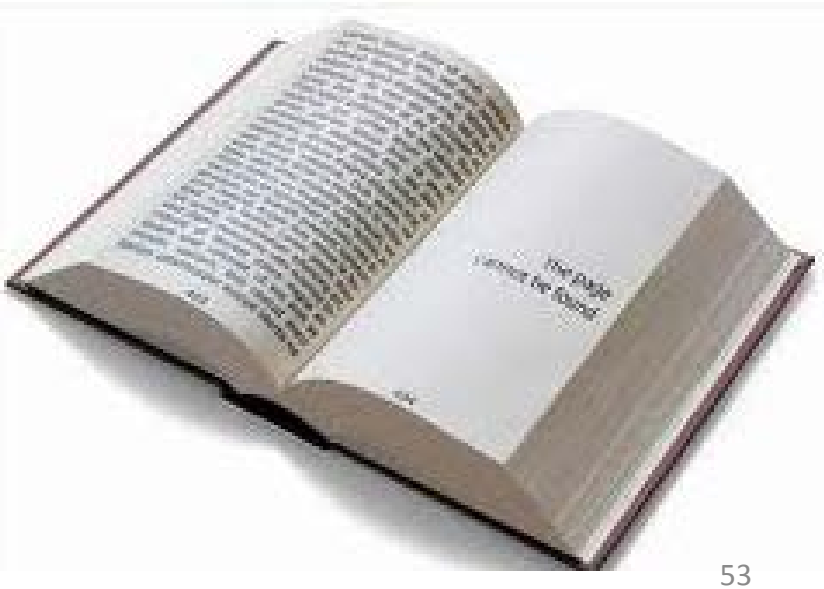
The Channel

- Scroll
- Codex (Manuscript)
- Printed book
- The screen of a computer
- The screen of a mobile phone
- Audio
- The eyes/brain (perceptive network)

Are these the same “thing”?



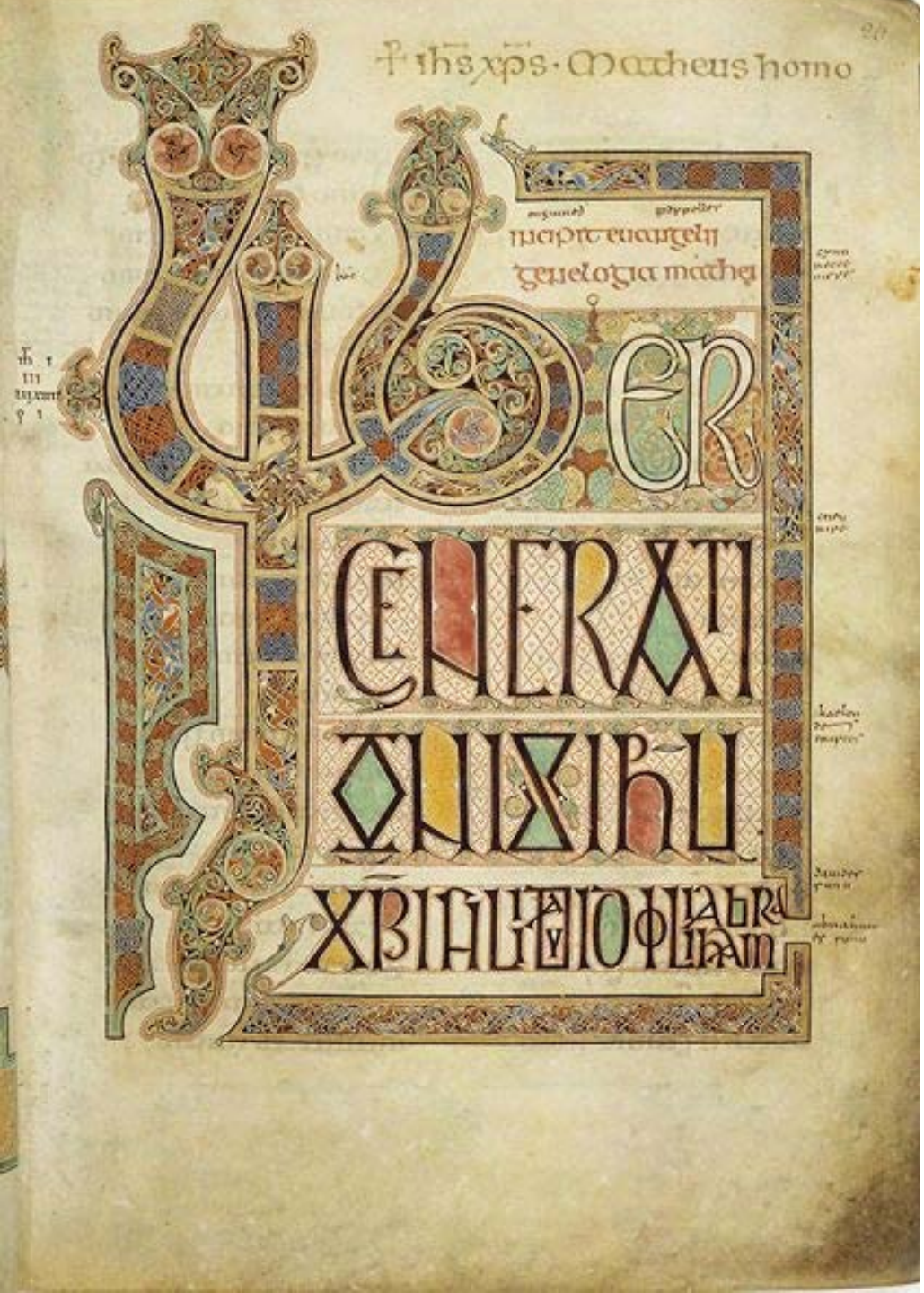




McLuhan 1962

The Noise of the Channel

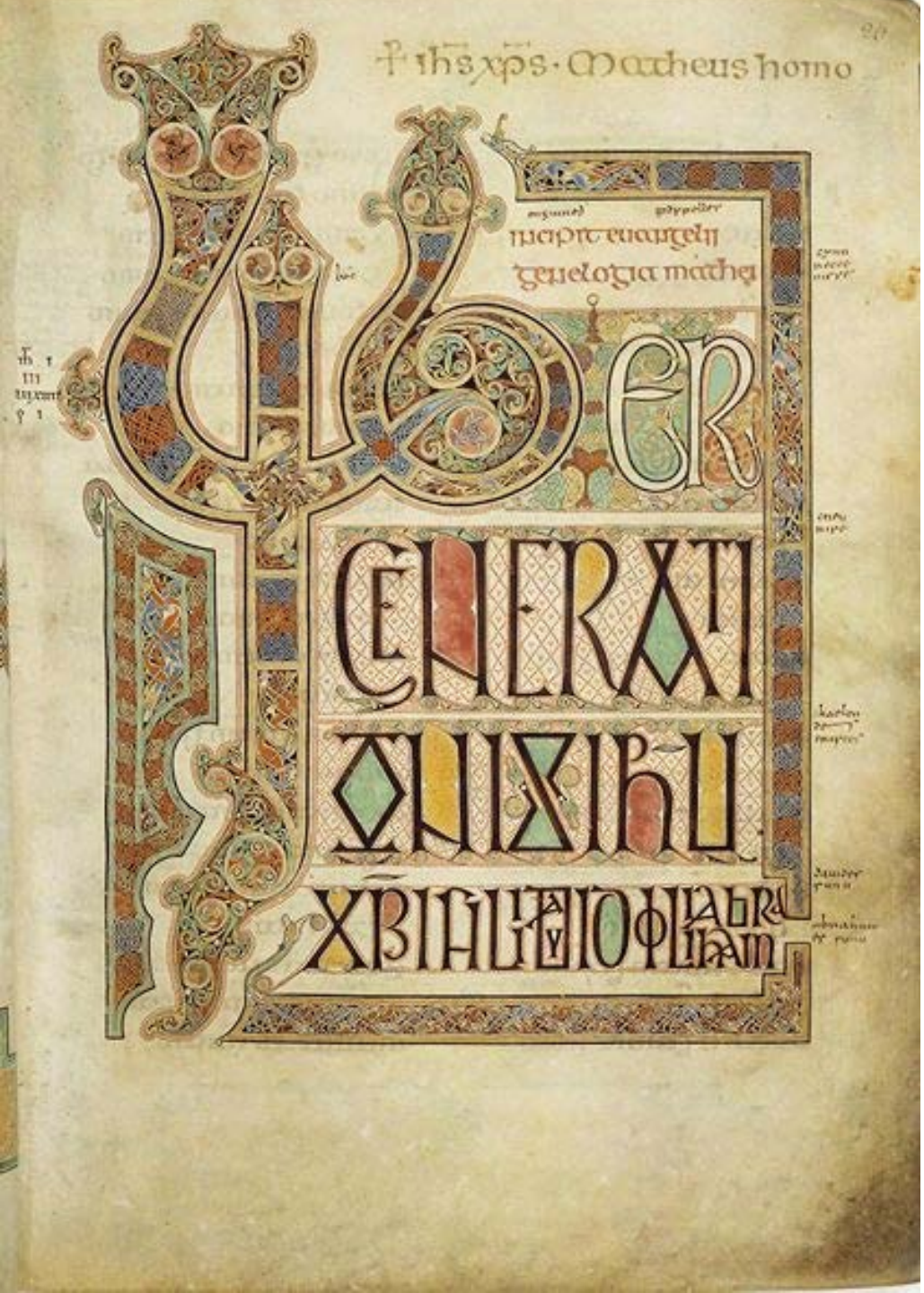
- In the writing system
- In the writing conventions
- In the style of writing
- In the support
- In the layout
- In the screen colors
- In the pronunciation



Are these the same "thing"?

Incipit Evangelii
Genelogia Mathei

Liber Generationis
Iesu Christi filii
David filii Abraham

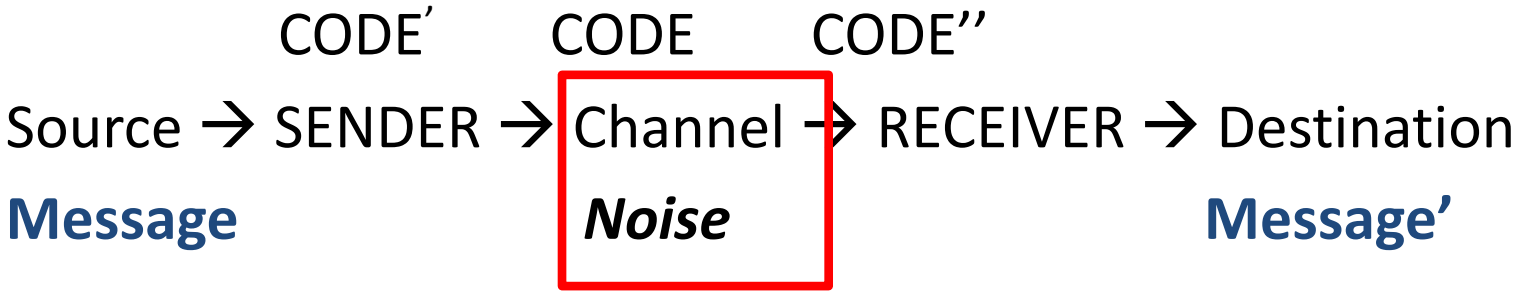


Text without noise?

Incipit Evangelii
Genelogia Mathei

Liber Generationis
Iesu Christi filii David
filii Abraham

The noise for someone is someone else's scholarship

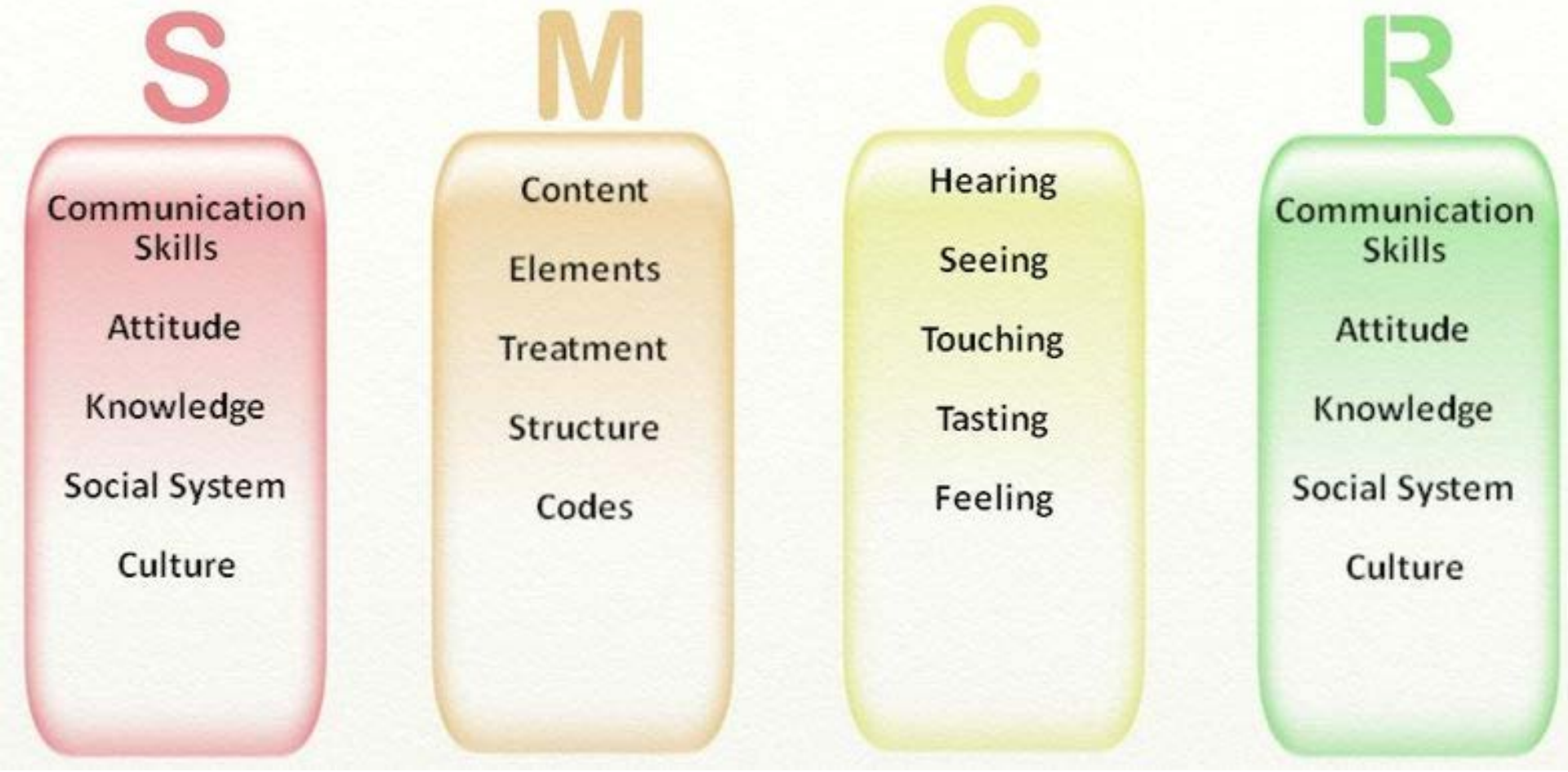


The destination

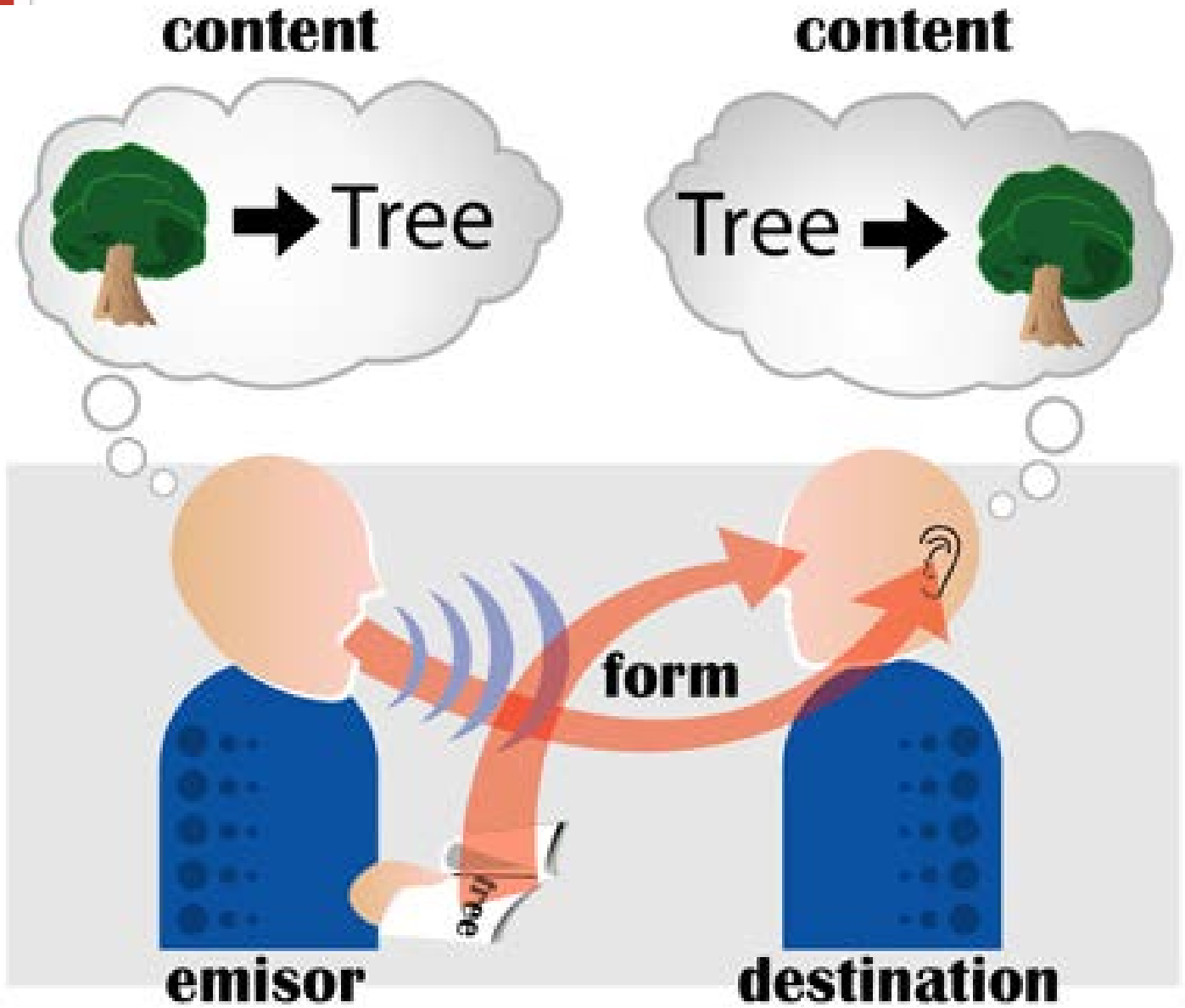
- Who is the target of the message?

- The scholar? A generic one, or a textual editor?
- A reader? A competent reader or a lay reader?

Another Model: Berlo's SMCR (1960)



Schramm: a social interaction



The environment

Jakobson's Function Languages Model

CONTEXT

SENDER → MESSAGE → RECEIVER

CHANNEL

CODE

Documentary-aware Editing

- Documents do not only contains words
- Where we find texts matters
- This is not a new concept

A social and cognitive approach?

- The message, the way we transmit and receive it depends also from our understanding, mood, social constraints, culture, etc.
- Communications follows semiotic rules:
  - Syntactic
  - Pragmatic
  - Semantic

A new definition of text

A text is a **multidimensional** message that conveys a set of meanings transmitted by various **codes** which are potentially understandable to at least one group of **Interpretants** with the capability and interest to decipher at least one of such codes

Textual transmission as communication

Textual transmission is a **selective and lossy act of communication** where the Interpretants-senders and Interpretants-receivers willingly or unwillingly makes the selection of which code to decipher from the number potentially available from the source that contains them.

Textual transmission as communications

- Textual transmission is a complex event
- Variation and distortion of the message happens at any level
- Variation is inevitable: what about making the most of it?
- What about valorising variation?

A new model is needed: next episode!

Elena Pierazzo & Geoffroy Noël

With many thanks to Peter Stokes

Department of Digital Humanities

King's College London

Université Stendhal Grenoble 3

UFR Langues