**TX5 Reading**

Subclass of: [S](#_3cqmetx)4 Observation

Superclass of:

Scope Note: Subclass of the [S4](#_1rvwp1q) Observation, referring to the semiotic procedure of decoding (and therefore understanding) a written text. This procedure can be carried out for scientific purposes, in order to analyse and study the text according to different disciplinary perspectives. The reading activity, thus, is intended as a specific observation (S4) in which the decoding of the signs is performed, i.e. the linguistic value is recognised and the message is understood. Cases in which decoding does not happen (e.g., the observer is able to describe the signs but not to assign a specific linguist value to them), the S4 class could be used as it is. For study purposes, the reading procedure requires a scientific autoptic examination of the text as preparatory action for the study. An autoptic examination consists of an accurate analysis of the surface and the signs and prescribes the use of specific tools and procedures, for establishing as faithfully as possible the exact value of each sign drawn on the physical feature.

Examples:

* The autoptic investigation of the South inscription ([TX1](#_4bvk7pj)) on the Arch of Constantine ([E22](#_2r0uhxc)) made by Rodolfo Lanciani between 1893 and 1901.
* The reading of the Greek text present on the Derveni papyrus ([E22](#_2r0uhxc)).

In First Order Logic:

 [TX5](#_1664s55)(x) ⇒ [S](#_3cqmetx)4(x)

Properties:

[TXP10](#_3q5sasy) read (was read by): [TX1](#_4bvk7pj) Written Text

**NEW**

**TX5 Text Recognition**

Subclass of: [S](#_3cqmetx)4 Observation, E65 Creation

Superclass of:

Scope Note: This class comprises activities of recognizing physical features on some surface, often an instance of TX4 Writing Field, as an arrangement of series of identifiable glyphs of some known script, deciphered or not, in an order characteristic for a text.

For study purposes, the text recognition procedure requires a scientific autoptic examination of the text. An autoptic examination consists of an accurate analysis of the surface and the signs and prescribes the use of specific tools and procedures, for establishing as faithfully as possible the exact value of each sign drawn on the physical feature. Deterioration of the original medium or “sloppy” writing may render parts of the original text as unrecognizable or ambiguous, which may be annotated in the transcript following epigraphic standards. A text recognition typically results in a record of an equivalent sequence of graphemes on another persistent medium in a scholarly established form of representation of the respective graphemes, often called a “transcript”.

An instance of TX5 Text Recognition may in particular apply even to a single glyph, typically forming part of an instance of TX5 Text Recognition applying to a larger sequence of glyphs containing the former glyph.

The recognition process may be assisted by a mechanical means, imaging technology, or a traditional squeeze for incised glyphs. In case the recognition process is solely based on the latter, the observation concerns only the representations on the latter as present to the researcher in some physical form or projection and should unambiguously be documented as such.

In case the recognized text has not been documented in a transcript, text recognition may constitute implicit part of an overarching reading process, instance of TXXX Reading, which has resulted in other noteworthy propositions related to the content of the recognized text. On the other side, recognition of single glyphs or contracted parts of texts, as they are characteristic for the use of ligatured scripts, may quite well be implicitly supported by the reader’s comprehension of the text and the creator of the transcript may have chosen not to annotate parts that she regarded as unambiguous. Since these cases can often hardly be separated from the shape recognition of the glyphs in isolation, documenting such implicit comprehension as a separate process may not be relevant. It is however regarded as good practice to document explicitly the reading process and associated interpretative reasoning for any non-trivial resolution of ambiguity or gaps in the recognized text that has a bearing on the transcript or further completion of the transcript.

Examples:

* The autoptic investigation of the South inscription ([TX1](#_4bvk7pj)) on the Arch of Constantine ([E22](#_2r0uhxc)) made by Rodolfo Lanciani between 1893 and 1901.

In First Order Logic:

 [TX5](#_1664s55)(x) ⇒ [S](#_3cqmetx)4(x)

 TX5(x) ⇒ E65(x)

Properties:

[TXP10](#_3q5sasy) deciphered (was deciphered by): E25 Human-Made Feature

TXPxx1 deciphered via the representation (was representation used for deciphering) : E36 Visual Item

TXPxx2 used copy or representation of (was deciphered via copy or representation): TX1 Written Text

TXPxx3 recorded transcript (was recorded by): Grapheme Sequence

**TXP10 deciphered text (was deciphered by)**

Domain: [TX5](#_1664s55)Text Recognition

Range: E25 Human-Made Feature

Subproperty of [O8](#_25b2l0r) observed (was observed by)

Quantification: one to one (0,1:0,n)

Scope note: This property associates an instance of TX5 Text Recognition with an instance of E25 Human-Made Feature carrying a sequence of glyphs that was recognized in the respective activity of text recognition. Typically, the associated instance of E25 Human-Made Feature is more specifically an instance of TX1 Written Text, however, a text may also be recognized from a mechanical copy, a photography, squeeze or other form of material copy of a written original, which would not by itself constitute an instance of TX1 Written Text.

 If the text was actually recognized only from a digital representation, this property may not be used, but the property *TXPxx1 deciphered via the representation* should be used instead.

Examples:

* The autoptic investigation ([TX5](#_1664s55)) carried out by Rodolfo Lanciani between 1893 and 1901, *deciphered* the South inscription ([TX1](#_4bvk7pj)) on the Arch of Constantine.

In First Order Logic:

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX5](#_1664s55)(x)

 [TXP10](#_kgcv8k)(x,y) ⇒ [E25](#_34g0dwd)(y)

[TXP10](#_kgcv8k)(x,y) ⇒ [O8](#_25b2l0r)(x,y)

**TXPxx1 deciphered via the** **representation (was representation used for deciphering)**

Domain: [TX5](#_1664s55)Text Recognition

Range: E36 Visual Item (Digital Object?)

Subproperty of

Quantification: one to one (0,1:0,n)

Scope note: This property associates an instance of TX5 Text Recognition with an instance of E36 Visual Item (digital object?), capturing the optical impression of an instance of TX1 Written Text by some mechanical method, that was used for recognizing the text without access to the original text and without an explicitly documented material copy or electronic display device that was used for the process.

If the text was actually recognized from an autoptic recognition or from a material reproduction, this property may not be used but the property “TXP10 deciphered text (was deciphered by)” should be used instead.

This property should also not be used, if the recognition of the text was actually carried out from the original text or a material copy of it together with an auxiliary instance of E36 Visual Item (digital object?). In this case, the use of the auxiliary material should be documented with the more general property *P16 used specific object.*

Examples:

* Reading Antikythera glyphs with BTI imaging.

In First Order Logic:

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX5](#_1664s55)(x)

 [TXP10](#_kgcv8k)(x,y) ⇒ [E36](#_34g0dwd)(y)

**TXPxx2 used** **copy or representation of (****was deciphered via copy or representation)**

Domain: [TX5](#_1664s55)Text Recognition

Range: [TX1](#_4bvk7pj) Written Text

Quantification: one to one (0,1:0,n)

Scope note: This property associates an instance of TX5 Text Recognition carried out only via copies of a text with the original instance of TX1 Written Text that was represented on the used copies or digital surrogates.

 This property is to be used only for non-autoptic recognition. If the recognition of the text was actually carried out from the original text, the property *TXP10 deciphered text* should be used for associating the instance of TX5 Text Recognition with the original instance of TX1 Written Text.

Examples:

In First Order Logic:

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX5](#_1664s55)(x)

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX1](#_4bvk7pj)(y)

**TXPxx3 recorded transcript (was recorded by)**

Domain: [TX5](#_1664s55)Text Recognition

Range: Grapheme Sequence

Subproperty of [P94](#_1jlao46) has created (was created by)

Quantification: one to one (0,1:1,1)

Scope note: This property associates an instance of TX5 Text Recognition with an instance of Grapheme Sequence that was created by this activity of text recognition for recording and representing as faithfully as possible the exact value of each sign drawn on the physical material of the regognized instance of TX1 Written Text.

Examples:

* The autoptic investigation ([TX5](#_1664s55)) carried out by Rodolfo Lanciani between 1893 and 1901, *read* the South inscription ([TX1](#_4bvk7pj)) on the Arch of Constantine.

In First Order Logic:

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX5](#_1664s55)(x)

 [TXP10](#_kgcv8k)(x,y) ⇒ [TX1](#_4bvk7pj)(y)

[TXP10](#_kgcv8k)(x,y) ⇒ P94(x,y)

**TX6 Transliteration**

Subclass of: E65 Creation

Superclass of:

Scope Note: This class comprises activities of exactly re-writing (i.e., re-encoding) an instance of Grapheme Sequence, i.e., the characters of a text, a contiguous part or a single character of it, by using a writing system ([TX3](#_43ky6rz)) different from that of the original text, without changing the order of characters or words, by using standard correspondences.

This operation may apply a 1:1 relation between the signs of the two writing systems, a “transliteration” in the narrower sense (e.g., the ALA-LC Romanization of Greek to Latin). It may also apply an approximation of the sounds of a language, as defined by the source writing system, by that of the target writing system, normally called a “transcription” (e.g., the “rōmaji” Romanization of Japanese), or a mixture of both (e.g. the ELOT 743 Type 2 – transcription of Greek to Latin letters). In a broader sense, the term “transcription” also applies to the activity of re-encoding a text using the same writing system (see example 1). The [P16](#_2iq8gzs) *used specific object (was used for)* property can be used to specify the applied method of correspondence.

Examples:

* Transcription, in Latin letters, of the Latin inscription(s) ([TX1](#_4bvk7pj)) on the Arch of Constantine ([E22](#_2r0uhxc)) reported in *Corpus Inscriptionum Latinarum* (CIL VI 1139).
* The transliteration and the transcription of the ancient Persian name of king Darius I, written in Persian cuneiform , into Latin script as respectively ‘da-a-ra-ya-va-u-ša’ and ‘Dârayavauš’.

In First Order Logic:

 [TX6](#_xvir7l)(x) ⊃ E65(x)

Properties:

 TXP3 renders (is rendered by): TX5 Text Recognition

TXP11 transcribed (was transcribed by): TX8 Grapheme

**TX8 Grapheme**

Subclass of: [E55](#_3hv69ve) Type

Superclass of:

Scope Note: This class comprises symbols used as kinds of atomic units with distinctive value in a given writing system in order to represent linguistic units of some level to encode elements of a message. According to the typology of the writing system, the represented linguistic units can be phonemes (as in Latin), syllables (as in Mycenaean Linear B), up to complete words (as in Chinese and Sumerian scripts).

A writing system also provides the conventions determining how the graphemes are to be used to write a language (orthographic rules).

In some writing systems, graphemes may also be used as auxiliary signs, for instance, for disambiguating senses of homonyms, as in Japanese writing system, or to mark the semantic categories of the words, as in the ancient Egyptian determinatives.

Examples:

* The abstract unit “S” of the Latin alphabet, used to represent the /s/ sound
* The abstract unit  of the ancient Persian syllabary, used to represent the /da/ syllable.
* The abstract unit “安” of the Han script, used to represent the meaning “peace”.

In First Order Logic:

 [TX8](#_1x0gk37)(x) ⊃E55(x)

**TXxx1 Grapheme Occurrence**

Subclass of: [E](#_3hv69ve)90 Symbolic Object, Txx2 Grapheme Sequence

Superclass of:

Scope Note: This class comprises single occurrences of a Grapheme used as an atomic unit at a particular position in the abstract form of a given particular piece of text.

Examples:

* The ideal letter “S” of the Latin alphabet, used to represent the /s/ sound, rendered by the specific S-shaped feature engraved on the second line of the South inscription on the attic of the Arch of Constantine (see section 1.3.1)
* The ideal ‘da’ syllabogram of the ancient Persian syllabary, used to represent the /da/ syllable rendered by the cuneiform sign  engraved on the first line of Darius I’s inscription ([TX1](#_4bvk7pj)) in Bagistan.

In First Order Logic:

 xx1TX8(x) ⊃E90(x)

**TXxx2 Grapheme Sequence**

Subclass of: [E](#_3hv69ve)90 Symbolic Object

Superclass of:

Scope Note: This class comprises particular sequences of Graphemes used for representing the abstract written form of a section of a given particular text. .

Examples:

In First Order Logic:

 TXxx2(x) ⊃E90(x)

**TXxx2 Script**

Subclass of: [E](#_3hv69ve)89 Propositional Object

Superclass of:

Scope Note: This class comprises functionally complete sets of mutually different graphemes employed by one or more languages, regardless linguistic specific use rules. On the other side, a writing system refers also to the set of relations between symbols and linguistic units they represent. The same language may be written using different scripts.

Examples:

* The Latin script used by the Italian and English writing systems.
* The Latin and the Greek scripts used for the encoding of the Oscan language, creating the Oscan-Greek and Oscan-Latin writing systems.



Figure 1: Text recognition, glyph sequences recognition and grapheme association.



Figure 2: Text recognition, deciphering and recording transcripts, and the new Reading conceptualisation