

APPENDIX A. LIST OF PRIMARY SOURCES IN *ROLAND*^{HT}

The following is a list of the primary sources encoded in *Roland*^{HT}. Most of them are text-based; works composed in other media are designated as such in the "Title/Name" column (see, for example, *Angoulême Carvings*).

Year	Author	Title/Name	Geo. Origin
~817-836	Einhard	<i>The Life of Charlemagne</i>	France
1095-99	unk.	<i>Song of Roland</i>	France
~1100	Pseudo-Turpin	<i>History of Charles the Great and Orlando</i>	France
1100s	unk.	<i>Angoulême Carvings</i> , stonework	France
~1150	Pfaffe Konrad	<i>Rolandslied</i>	Germany
~1180	Bertrand de Bar-Sur-Aube	<i>The Song of Girart of Vienne</i>	France
XII-XIIIc	unk.	<i>Roland and Oliver</i> , Chartres Cathedral statues	France
XII-XIVc	unk.	<i>Firumbras</i>	England
XII-XIVc	unk.	<i>Reims Triptych</i> , stone carving	France
XIIIc	unk.	<i>Karlamagnús Saga</i>	Norway
XIVc	unk.	<i>The Middle English Song of Roland</i>	England
early XIVc	unk.	<i>Otuel and Roland</i>	England
1300s	unk.	<i>Rouland and Vernagu</i>	England
1308-21	Dante Alighieri	<i>Divine Comedy</i>	Italy
1300-1600	unk.	<i>Cân Rolant</i>	Wales
XVc	unk.	<i>Dubrovnik Roland</i> , statue	Croatia
1471-86	Matteo Maria Boiardo	<i>Orlando Innamorato</i>	Italy

Year	Author	Title/Name	Geo. Origin
1532	Ludovico Ariosto	<i>Orlando Furioso</i>	Italy
1572	unk.	<i>The Tale of Ralph the Collier</i>	Scotland
1605-15	M. de Cervantes Saavedra	<i>Don Quixote</i>	Spain
1810	William Sotheby	<i>Constance de Castille</i>	England
1820	Thomas Campbell	<i>The Brave Roland</i>	England
1831	Laetitia Elizabeth Landon	<i>Roland's Tower. A Legend of the Rhine.</i>	England
1839	Emmeline Stuart-Wortley	<i>The Tower of Roland</i>	England
1849	Henry B. Hirst	<i>The Penance of Roland</i>	USA
1849	William Motherwell	<i>Roland and Rosabelle</i>	Scotland
1855	Robert Browning	<i>Childe Roland to the Dark Tower Came</i>	England
1860	William Caldwell Roscoe	<i>Eliduke, Count of Yveloc</i>	England
1868	Thomas Westwood	<i>An Angler's Dream Under Rolandseck</i>	England
1875	Albert B. Barrows	<i>Roland of Algernon</i>	USA
1901	Robert Williams Buchanan	<i>The Death of Roland</i>	England
1903	John Warren	<i>Roland at Roncesvalles</i>	England
1911	Maurice H. Hewlett	<i>The Birth of Roland</i>	England
1930	Benjamin Low	<i>Roland, A Symphonic Poem</i>	USA
~1941-45	League of Roland	<i>Roland: Country First</i>	England
1942	Adair Forrester	<i>The Children's Story of Roland</i>	England
1949	Peter Racine Fricker	<i>Rollant et Oliver</i>	England
1975	L. Sprague de Camp	<i>The Compleat Enchanter</i>	USA
1978	Warren Zevon	<i>Roland the Headless Thompson Gunner</i>	USA
1982-2004	Stephen King	<i>The Dark Tower I-VII</i>	USA
1994	Gianni Celati	<i>Orlando Innamorato</i>	Italy

Year	Author	Title/Name	Geo. Origin
1995	Greg Roach	<i>The Madness of Roland</i>	USA
1999	Shayne Amaya et al.	<i>Roland: Days of Wrath</i>	USA/Brazil

Table 1. Primary sources encoded in *Roland*^{HT}.

Notably absent from the above list is Shel Silverstein's 1973 "Roland the Roadie and Gertrude the Groupie," written for the rock group Dr. Hook and the Medicine Show. Although parts of the song are encoded, they are commented out because permission to use the work online has been denied by the copyright holder. Encoding data from Silverstein are included in the statistics presented in this dissertation.

APPENDIX B. EXPRESSING SEMANTIC INFORMATION IN HTML

In the first – HTML-based – version of *Roland^{HT}*, semantic information was encoded into the project by way of toggling (using JavaScript) visibility of such information when the user clicks on a given link. For example, the *Song of Roland* excerpt titled "knight, ambassador, coward" has a hyperlink around Roland's words to his stepfather Ganelon: "I love you not a bit":

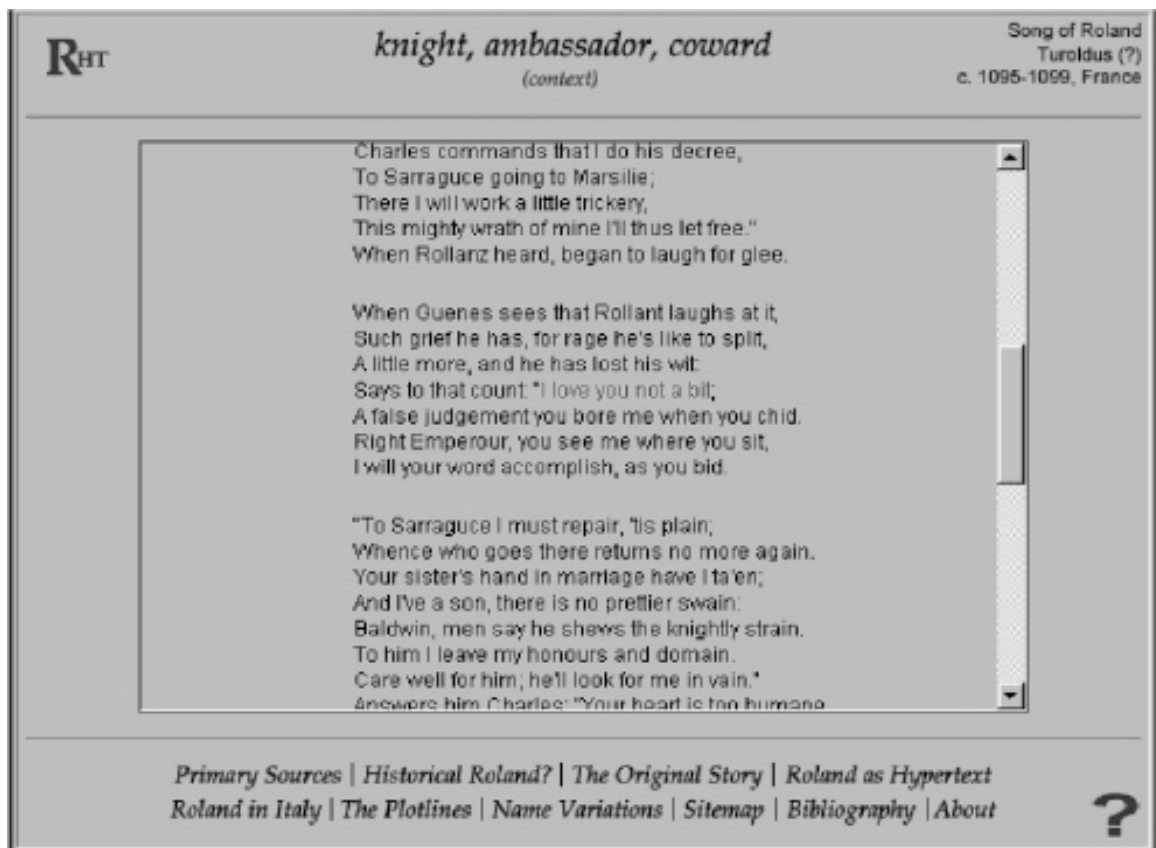


Figure 1. "Knight, ambassador coward" hyperlink.

Love in general, and its lack between Roland and Ganelon in particular, is a recurrent theme in the corpus. An interesting exception to this is the *Karla-*

magnús Saga episode in which the two meet, become kin, and swear fellowship to each other. To highlight this, clicking on the link brings up an overlaid comment, "Ah, but they loved each other once," complete with a link to the excerpt titled "how rollant met guenelun":

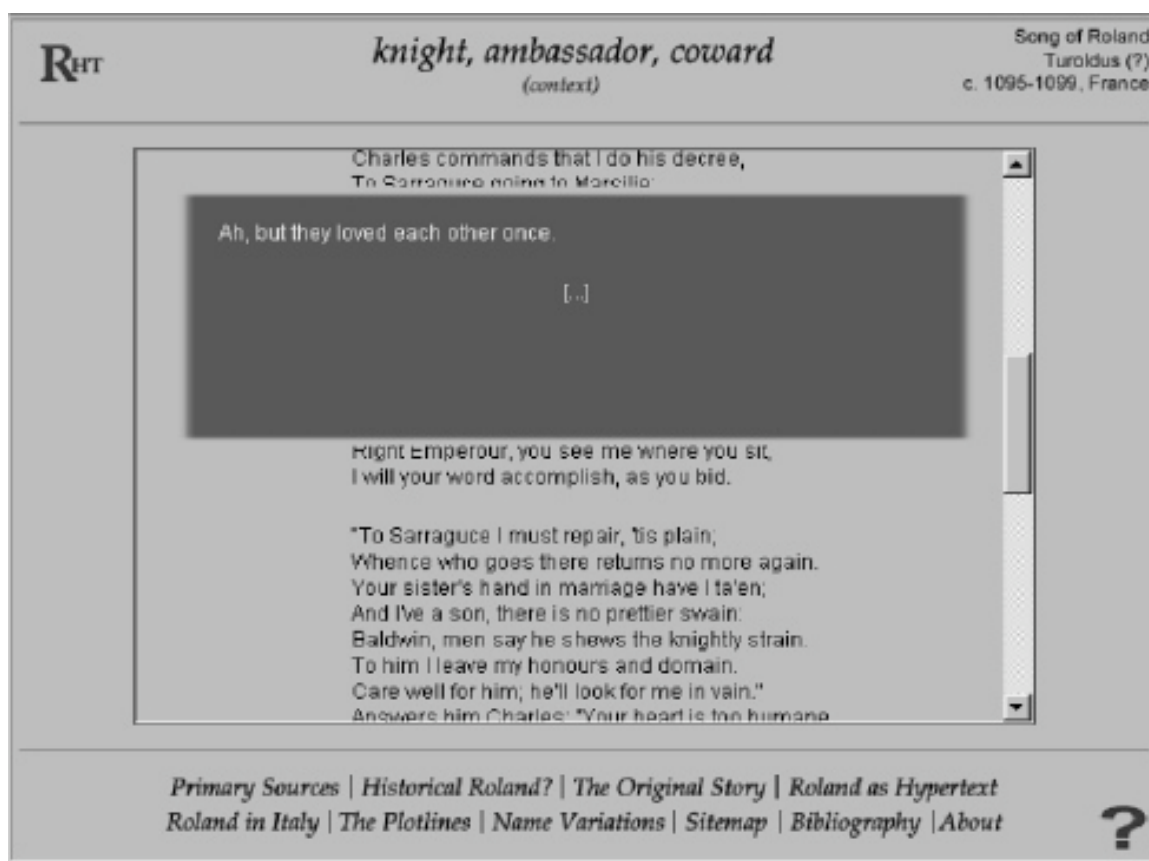


Figure 2. "Knight, ambassador coward": hyperlink clicked.

This approach proved frustrating for most of the twenty or so informal beta testers of the project in 2001: every link required two clicks instead of the customary one, considerably slowing down the reading process. In addition, although the prose blurbs conveyed thematic information, there was no categorized theme set. Consequently there was no way to see, for example, a list of all the excerpts treating the subject of love.

APPENDIX C. CLEANING UP FIRST-PASS XML ENCODING

Having encoded themes and imagery that appeared interesting at the time of encoding, I analyzed the statistics of their occurrence in the corpus. The following themes were deleted (number of occurrences in parentheses):

army (4)	pity (1)
ceremony (3)	prudishness (1)
conspiracy (1)	ritual (2)
disguise (1)	spite (3)
denial (2)	stubbornness (3)
gift (3)	summons (1)
hostility (2)	temper (4)
jealousy (2)	threefold repetition (2)
mastery (1)	writing (2)

Table 2. Deleted themes and imagery.

I also consolidated some rarely occurring theme elements and attributes (marked with an @ symbol before the attribute name) and semantically similar or more appropriate ones:

Original theme/image	Folded into...
arrogance	overconfidence
adultery	kinship
council	counsel
cruelty	violence
crusade	violence
gore	violence
incest	kinship
invulnerability	unconquerable
martyrdom	sacrifice
massacre	violence
prophecy	magic

Original theme/image	Folded into...
prudishness (1)	chastity
prudishness (1)	chivalry
ritual (1)	chivalry
rolands priorities	chivalry
theft	treachery
trust	fellowship
family ties	kinship
@imagined['yes']	@realized['no']

Table 3. Consolidated themes and imagery.

Working on encoded places and character names (which, being recurrent and/or echoes of each other, are one of the elements that tie the corpus together) led to regularization of their names to their presently most widely used forms. Aix and Aix-la-Chapelle became Aachen. Among characters, Basile was changed to Basil, Balsan – to Basin, Blancandrin – to Blancadrin; Geluvis to Geluviz, Gilem to Gille, and Olivieri to Oliver. Problems arose during regularization, some of them of a surprisingly political nature. For example: do I encode a river as located in Taiwan, or in China? (The answer to this was ultimately China, because both of the following conditions are true: the primary source had explicitly placed it there; and the river was in China at the time of the book's writing.)

The other significant challenge was posed by the name of Roland's mother. Her names are sometimes wildly disparate and clearly not variations upon one another. In the corpus, she is called in turn Gille, Gilem and Bertha. Bertha occurs only once, but this is a common name for her in most of the Mediterranean, and so it seems unwise to fold Bertha into Gille. Gilem, however, was folded into Gille.

Places acquired a new attribute – @where – which denotes the country (if working with a city, church, etc.) or the continent (if it is a country, mountain, etc.) within which it is located. This will allow users to search not only for specific places, but for what is located in, for example, Asia.

APPENDIX D. THEME STATISTICS

To arrive at the following statistics, I have broken the corpus works down into four temporal groups: medieval (designated MVL below, works written in or before 1350); Renaissance (REN, 1351-1650), modern (MOD, 1651-1899) and contemporary (CON, 1900-present). For each theme, the number columns represent the number of times that theme is encoded in works from the relevant time period. The percentages were calculated using the total number of <theme> element occurrences in the works from that period. Note that the same theme was, as likely as not, encoded multiple times in the same excerpt, so the numbers below do not represent the number of excerpts in which these themes occur.

Statistics for the top five most often occurring themes for each time period are rendered in boldface.

	MVL #	MVL %	REN. #	REN %	MOD #	MOD %	CON #	CON %
accusation	44	4.64%	4	2.72%	22	4.24%	9	4.10%
anger	30	3.16%	5	3.40%	5	0.96%	3	1.36%
beauty	5	0.53%	0	0.00%	0	0.00%	0	0.00%
betrothal	3	0.32%	0	0.00%	4	0.77%	0	0.00%
chastity	2	0.21%	2	1.36%	3	0.58%	1	0.45%
chivalry	22	2.32%	4	2.72%	19	3.66%	7	3.18%
combat	105	11.06%	17	11.56%	40	7.71%	35	15.91%

	MVL #	MVL %	REN. #	REN %	MOD #	MOD %	CON #	CON %
conquest	6	0.63%	0	0.00%	7	1.35%	2	0.91%
counsel	3	0.32%	1	0.68%	0	0.00%	0	0.00%
courage	30	3.16%	6	4.08%	16	3.08%	1	0.45%
cowardice	9	0.95%	6	4.08%	2	0.39%	1	0.45%
death	65	6.85%	3	2.04%	80	15.41%	12	5.45%
deceit	11	1.16%	7	4.76%	0	0.00%	5	2.27%
defiance	1	0.11%	0	0.00%	2	0.39%	0	0.00%
diplomacy	8	0.84%	0	0.00%	0	0.00%	0	0.00%
dream	6	0.63%	2	1.36%	6	1.16%	1	0.45%
evil	2	0.21%	0	0.00%	4	0.77%	0	0.00%
fear	8	0.84%	1	0.68%	2	0.39%	0	0.00%
fellowship	40	4.21%	0	0.00%	25	4.82%	11	5.00%
glory	4	0.42%	0	0.00%	5	0.96%	0	0.00%
grief	31	3.27%	6	4.08%	25	4.82%	5	2.27%
honesty	1	0.11%	0	0.00%	4	0.77%	0	0.00%
honor	9	0.95%	0	0.00%	9	1.73%	0	0.00%
insult	5	0.53%	1	0.68%	3	0.58%	0	0.00%
journey	3	0.32%	6	4.08%	0	0.00%	0	0.00%
kinship	3	0.32%	0	0.00%	5	0.96%	0	0.00%
knighthood	13	1.37%	0	0.00%	20	3.85%	7	3.18%
lament	9	0.95%	0	0.00%	2	0.39%	0	0.00%
love	29	3.06%	15	10.20%	26	5.01%	16	7.28%
loyalty	19	2.00%	1	0.68%	15	2.89%	7	3.18%
madness	14	1.48%	13	8.84%	2	0.39%	14	6.36%
magic	8	0.84%	1	0.68%	3	0.58%	25	11.36%
marriage	10	1.05%	1	0.68%	10	1.93%	2	0.91%
monjoie	4	0.42%	0	0.00%	0	0.00%	0	0.00%

	MVL #	MVL %	REN. #	REN %	MOD #	MOD %	CON #	CON %
nobility	11	1.16%	0	0.00%	9	1.73%	2	0.91%
omen	20	2.11%	0	0.00%	1	0.19%	0	0.00%
overconfidence	3	0.32%	2	1.36%	0	0.00%	0	0.00%
pain	5	0.53%	1	0.68%	9	1.73%	0	0.00%
piety	25	2.63%	1	0.68%	0	0.00%	8	3.64%
pride	21	2.21%	1	0.68%	8	1.54%	0	0.00%
protection	13	1.37%	1	0.68%	5	0.96%	1	0.45%
quest	5	0.53%	2	1.36%	0	0.00%	0	0.00%
religion	127	13.38%	11	7.48%	43	8.29%	11	5.00%
revenge	9	0.95%	0	0.00%	3	0.58%	1	0.45%
sacrifice	4	0.42%	1	0.68%	2	0.39%	1	0.45%
shame	26	2.74%	2	1.36%	9	1.73%	0	0.00%
storytelling	1	0.11%	0	0.00%	1	0.19%	0	0.00%
strength	24	2.53%	8	5.44%	15	2.89%	4	1.82%
threat	22	2.32%	3	2.04%	8	1.54%	2	0.91%
tower	0	0.00%	0	0.00%	2	0.39%	0	0.00%
treachery	24	2.53%	3	2.04%	11	2.12%	4	1.82%
unconquerable	7	0.74%	3	2.04%	0	0.00%	0	0.00%
violence	30	3.16%	4	2.72%	16	3.08%	11	5.00%
virtue	3	0.32%	1	0.68%	4	0.77%	2	0.91%
weakness	3	0.32%	0	0.00%	7	1.35%	0	0.00%
wisdom	4	0.42%	1	0.68%	0	0.00%	0	0.00%
Total	949	100%	147	100%	519	100%	220	100%

Table 4. Theme statistics.

APPENDIX E. XML

The World Wide Web Consortium (W3C) describes XML documents as "made up of storage units called entities, which contain either parsed or unparsed data. Parsed data is made up of characters, some of which form character data, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. XML provides a mechanism to impose constraints on the storage layout and logical structure." (XML 1.0)

An element, the most widely used kind of XML entity, is enclosed in angle brackets. It consists either of an opening and a closing tag enclosing some content –

```
<element tag="opening">This is an element.</element
tag="closing">
```

– or of a singleton tag "closed" by a forward slash:

```
<singleton_element/>
```

Elements are allowed attributes, which function as sub-categories to the elements' categories; in the above example `tag` is an attribute, whose values are `opening` and `closing`. Elements can also contain other elements; however, contained elements must be wholly contained (beginning and ending tags properly *nested*, like concentric circles). There can be no overlapping of elements:

```
<sentence>This <verb>is</verb> a sentence.</sentence>
```

There are only a few other core rules to XML. Tags must be consistently capitalized (<name> and <Name> are different entities). Finally, XML must be *well-formed*: it must include a prolog (such as an XML declaration, for example `<?xml version = "1.0"?>`) and a root element "no part of which appears in the content of any other element" (XML 1.0) but which can, and usually does, surround all other content.

There may be one or more supplemental documents that define what elements, attributes and other entities may be used in a specific XML project, as well as where and how they may be placed. Such a document is called a DTD (Document Type Definition), or else a schema. If such a document is written and referenced in an XML file's prolog, the additional rules described in it apply to that XML file. Conformance to these rules makes the file *valid*. It is possible for an XML file to be well-formed but not valid if its general syntax is correct but it breaks the rules specified in the DTD or schema.

APPENDIX F. SEMANTIC CODE STRUCTURE AND INTERFACE

Semantic code structure

All excerpt texts and metadata, with the exception of image, video and/or sound files (stored separately), are contained in the file titled *rolandht.xml*. The basic encoding structure of the file (excepting information within an excerpt, for which see next section) follows; for information about XML see Appendix E.

```
<works>
  <work date="" geo="" lang="" name="" type="" timeperi-
od="">
    <header>
      <author id=""></author>
      <title id="">Song of Roland</title>
      <language></language>
      <translator></translator>
      <textnotes caption=""></textnotes>
    </header>

    <excerpt id="" title="">
      <context></context>
      <txt>(see below)</txt>
    </excerpt>
    [more excerpts, if and as needed]
  </work>
  [more works, as needed]
</works>
```

In conventional English, the above can be read as follows. This is a collection of excerpts from different works. Within each work are recorded the date of its creation (if known), its geographic origin, original language, name, type (prose, verse, drama), and time period (necessary for performing the statistical

analysis presented in Appendix D). The author's name (if known), title and translator of the work (if any) are also noted. If general notes on the work are present, they are part of the header section. Each excerpt has a unique ID (necessary for processing for web presentation), a title, a short description of its context within the work, and the text of the excerpt itself. Each work contains one or more excerpts; the *rolandht.xml* document contains several works.

Semantic code structure within an excerpt

Besides the context and any structural encoding (paragraphs, line breaks for verse, etc.), an excerpt may contain the following elements⁶⁰ and/or attributes:

```
<theme name="" who="" accused="" accuser="" charge=""
realized="" simile="" metaphor="" tstart="" tend=""/>
<imagery name="" type="" realized="" magic="" called=""
belongs="" whom="" simile="" metaphor="" tstart=""
tend=""/>
<character name="" collective="" mention="" religion=""
myth="" myth-origin="" historical="" tstart="" tend=""/>
<place name="" type="" where="" myth="" myth-origin=""/>
<speech who="" cont="" internal="" type="" according-
to=""/>
<transl eng=""/>
```

⁶⁰ The elements are presented as singletons here in order to save space, but most of the time they are not. There are no restrictions on the nesting order of the semantic elements within an excerpt.

<note/>

A **theme** may have: a *name* that appears in the middle column of the website; a *who* designation, which attributes a theme (for example fear) to a specific character; a *realized* attribute, which specifies whether the action of the theme is performed or merely discussed. If the theme is accusation, the attributes *accused*, *accuser* and *charge* provide more information about the accusation. There are also attributes to categorize themes as *similes* or *metaphors*; and finally, *tstart* and *tend* refer to the time signatures between which the given theme occurs in a film clip. The four latter attributes may also be present in <imagery>; *tstart* and *tend* may also occur in <character>.

Imagery has a *name* (accessory, sound, nature or animal) and a *type* (weapon, hornblow, water, lion). An accessory may have a name of its own (be called Durendal, for example); *belongs* to someone, but may be in someone else's possession, which is noted in *whom*. An image may have *magic* properties, and may or may not be *realized* (see themes above).

A **character** has a *name* and perhaps a *religion*. The character may be *collective* (the Saxons); a *myth* (Saint Michael) – in which case it will have a *myth-origin* – or *historical*; and finally, the character may be merely *mentioned* but not present in the scene.

A **place** has a *name* (Paris) and a *type* (city); it may be a *myth* (in which case it may also have a *myth-origin*), and it may or may not be geographically contextualized using the *where* attribute.

Speech has a speaker (*who*); may be *internal*, such as a character talking to himself; and may have a *type* (for example lament). The speech of one character may be related by another; such cases are conveyed through the *according-to* attribute. Finally, the *cont* attribute is one of convenience: it permits me to designate a single speech instance spanning more than one structural element (such as a paragraph) without violating the XML hierarchy.

Translations from the Middle English, intended as reading aids, have only one attribute – *eng* for "English" – which contains the translation itself. The `<transl>` element surrounds the word or phrase being translated.

Finally, **notes** are in-line annotations (which appear as quill icons in the interface, see below) that draw attention to particularly interesting semantic connections within the corpus, reference translators' interpretations of obscure geographical locations, and/or provide further contextual information.

Web interface

The web interface consists of four major areas (Fig. 3). At the top (1) are links to bibliography and appendices, and a "reset" link at top right that reloads the page, showing all excerpt titles and marking all of them unread. Under the project title are links to theoretical essays and a help file. The rest of the browser window is divided into three columns: from left to right, the excerpt list with an informational button at top left that explains the ordering and titles of the excerpts (2); the list of characters, imagery and themes present in a displayed excerpt (3); and the text and/or multimedia links of the excerpt itself (4). Clicking on the

(*more...*) link beside any excerpt title reveals contextual information for that excerpt; the link then changes to (*less...*) and can be used to hide the context.

The screenshot shows the main interface of the *RolandHT* tool. At the top, there is a navigation bar with the title "ROLANDHT" and a "reset" button. Below the navigation bar, there are three main panels:

- Panel 1 (Left):** A list of excerpts available for viewing. The list includes titles like "High Council", "Ganelon: Knight, Ambassador, Coward", "Ganelon and Blancandrins", "Ganelon's Treachery", "Charlemagne's Dream", "In the Rearguard", "Courage and Faith", "Obstinate", "Battle", "Nature Weeps", "Guenes, You Traitor", "Valiant Fighter", "The Cost of Pride", "The Cost of Pride - Part II", "Roland Grieves", "Oliver's Death", "To the Last", "Blessing the Dead", "Dying", "Charlemagne Grieves and Avenges", "Poor Aide", "Ganelon Unrepentant", "High Council", "Temper, Temper...", "Gwenlwyd", "Betrayal", "Gwenlwyd And the Paggans", "Charlymaen's Dream", "Gwenlwyd Implements Evil Plan", "Roland And the Rearguard", "Battle Imminent", "Contumacious", "Roland Loves a Battle", and "Nephew Against Nephew". Each title has a "(more...)" link next to it.
- Panel 3 (Center):** A vertical sword image. To its left, there are two sections: "Characters:" listing "Roland", "Marsile", and "Saracens"; and "Themes:" listing "combat", "strength", "anger", and "religion".
- Panel 4 (Right):** A detailed view of an excerpt titled "Roland's Signature Move" by "Otuel and Roland by Unknown, England, early XIV cent.". The text of the excerpt is displayed, with some words highlighted in blue (e.g., "witherlynges").

Figure 3. Main elements of the *RolandHT* interface.

Placing the mouse over a theme (Fig. 4, center) highlights all occurrences of that element in the excerpt (1). Clicking on the theme modifies the excerpts list to show only those excerpts in which the theme occurs (2). The same functions are available for imagery and characters.

This screenshot illustrates the interaction with themes. It shows the same interface as Figure 3, but with specific interactions highlighted:

- Panel 3 (Center):** The "Themes:" list has "strength" highlighted. A mouse cursor is positioned over it. A line labeled "1" points from this theme to the corresponding text in the excerpt on the right.
- Panel 4 (Right):** The excerpt text "Thoruþ the hauberk, the blod was sene, ffor the strokys weren ful kene, That deled were by-twene hem to." is highlighted in blue, corresponding to the "strength" theme.
- Panel 1 (Left):** The list of excerpts has "To the Last (more...)" highlighted in blue. A mouse cursor is positioned over it. A line labeled "2" points from this highlight to the "strength" theme in the central panel.

Figure 4. Working with themes.

The excerpt pane (Fig. 5) may contain any of the following three study aids: information on the work as a whole (1); in-text notes (2); and/or words with available translations (3). Clicking on a quill icon (2) reveals the in-text note (4). Clicking on the information icon (1) reveals a similar closeable window with notes on the work as a whole.

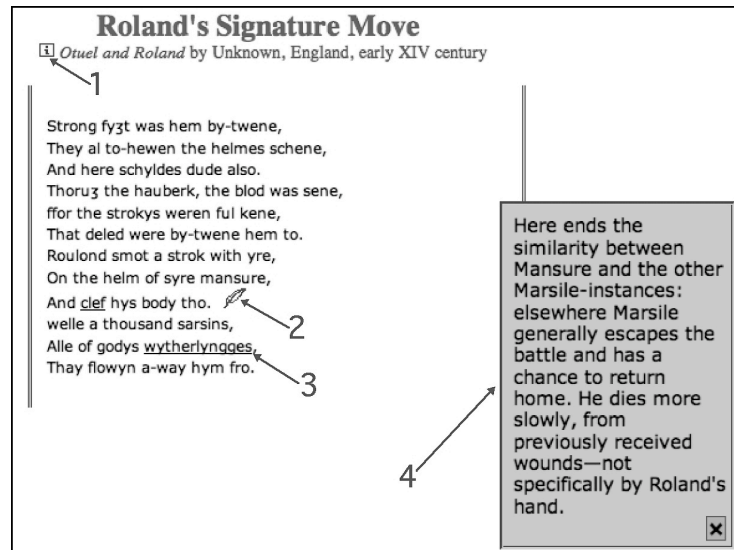


Figure 5. Study aids within the excerpt.

APPENDIX G. SAMPLE TEACHING MODULES

Module 1: Objective

Acquaint the student with critical and paratextual elements of *Roland*^{HT} – information about the source, general text notes, translations, in-text notes and contextual information for each excerpt.

Module 1: Assignment

Load *Roland*^{HT} in your browser. Find the excerpt titled "Forest [Battle at] Runcyvale." Answer the following queries, using the "Help" section if needed:

1. What text is this excerpt from? When was the text written?
2. Which character figures most prominently in the context for this excerpt?
3. Why was the Middle English left untranslated?
4. What do the following words mean: hende; y-slawe; ek; grethed; knyzt?
5. Why is it remarkable that Roncesvalles (Runcyvale) is claimed to be a forest?

Module 2: Objective

Acquaint the student with a multilinear reading process guided by her own interests.

Module 2: Assignment

Load *Roland*^{HT} in your browser. Do the following:

1. Read some excerpts – any from the long list on the left-hand side.

Choose a theme or image that strikes you; click on that theme or image (listed in the middle column) to get a list of excerpts that contain it.

2. Record a path of at least five excerpts connected by the query term you selected; be sure to select excerpts from at least two or three different works.

Note the title, author and date of the work's composition in addition to the excerpt title.

3. Examine closely the words or phrases highlighted in your chosen excerpts when you mouse over the theme or image in the middle column.

4. Write a paragraph or two on what role you think the theme or image plays in the corpus. What might it reveal about Roland? What contributions does this element make to individual plot lines? How are these contributions similar and different among excerpts from different works?

5. What is the most surprising piece of new information you have received in the course of this exercise? Explain in one or two sentences.