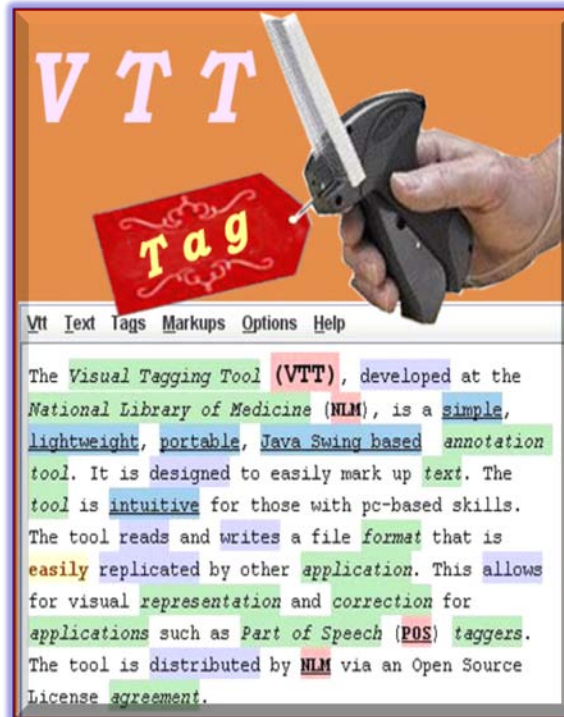


Visual Tagging Tool (VTT)



Dr. Chris J. Lu

The Lexical Systems Group

NLM. LHNCBC. CGSB

<http://SPECIALIST.nlm.nih.gov/vtt>

12, 2010

Table of Contents

- Introduction
- Manual tagging
- VTT file format (integration)
- Questions

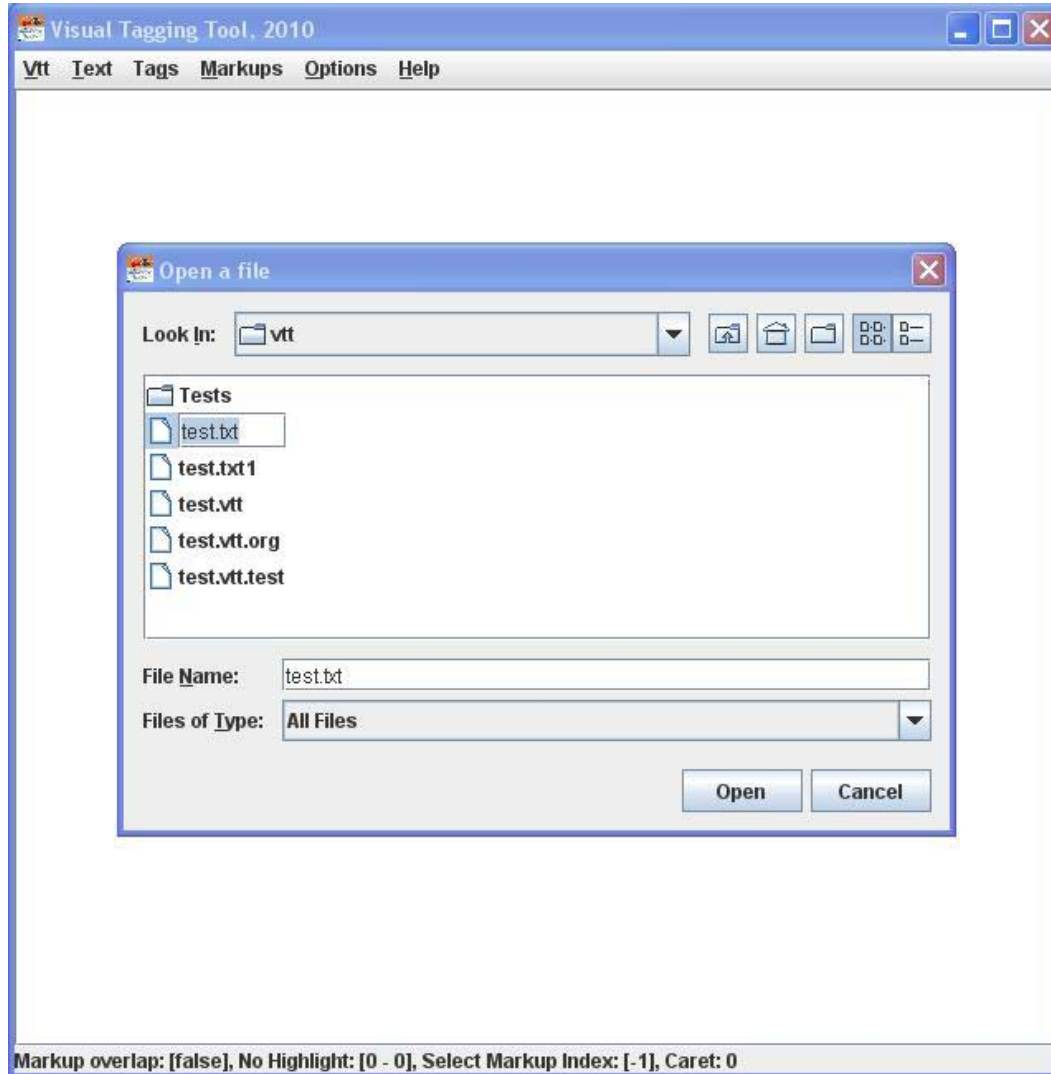
Introduction

- A simple, easy, lightweight, portable, Java Swing based annotation tool
- Shows tagged text in different visual effects: **color**, **font**, size, **bold**, *italic*, underline, etc.
- Developed to ease the human tagging process (markup text)
- Can be integrated with other NLP programs
- Full documents & supports
- Free distributed with open source code (since 2009)

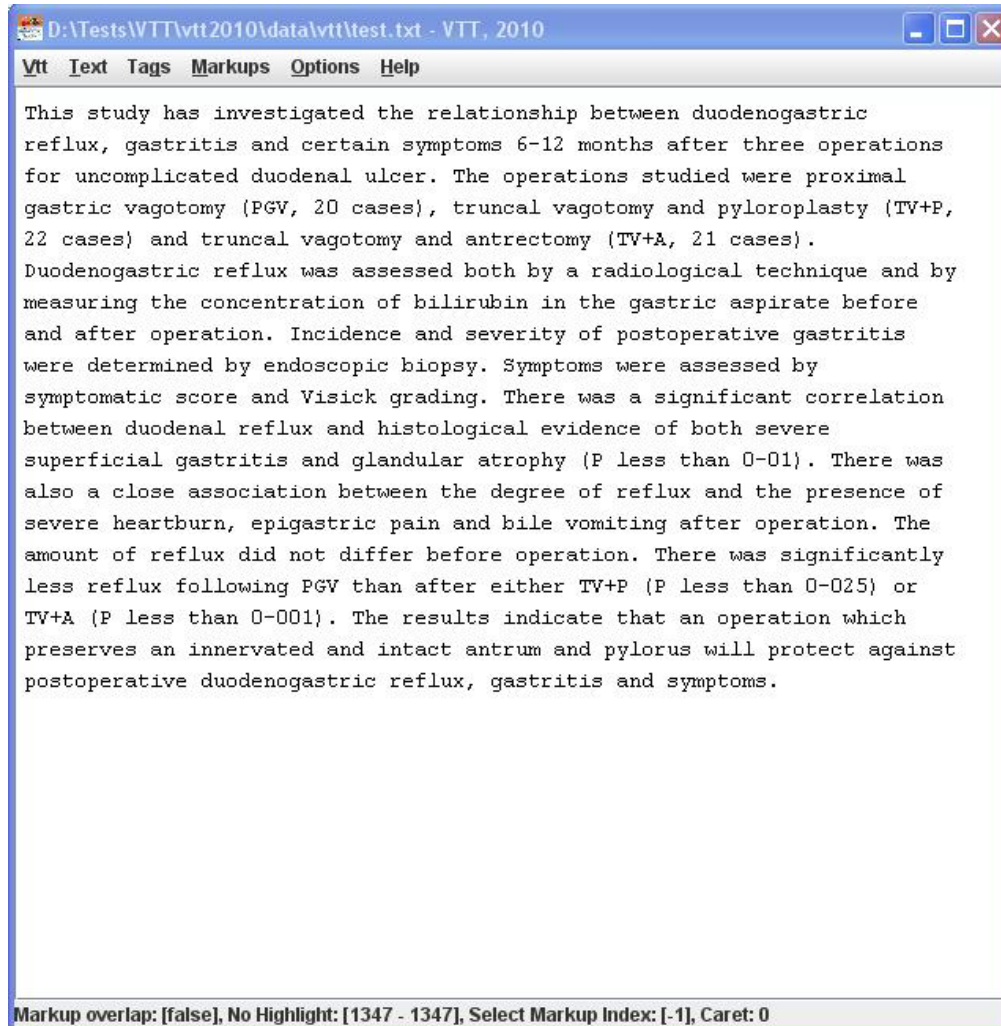
Basic Steps to Use VTT?

- 1) Open a plain unmarked text
 - Vtt -> Open (O)
- 2) Define or import tags
 - Tags -> Setup
 - Tags -> Setup -> Import
- 3) Start to markup (assign tag to plain text)
 - Select text: smear, double clicks, triple clicks, quick keys, etc.
 - Assign tag: quick keys, pull-down menu

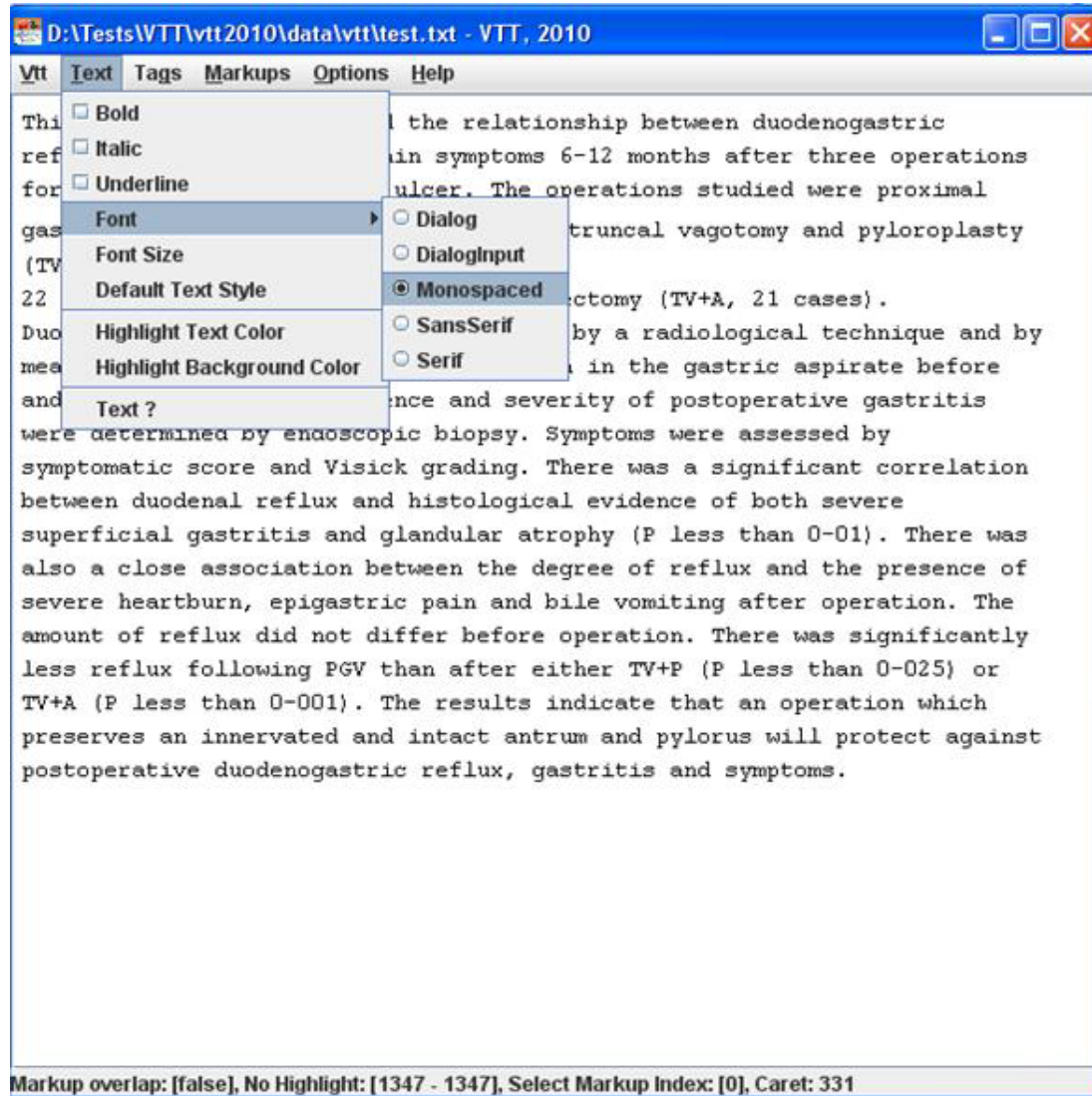
Text - Open



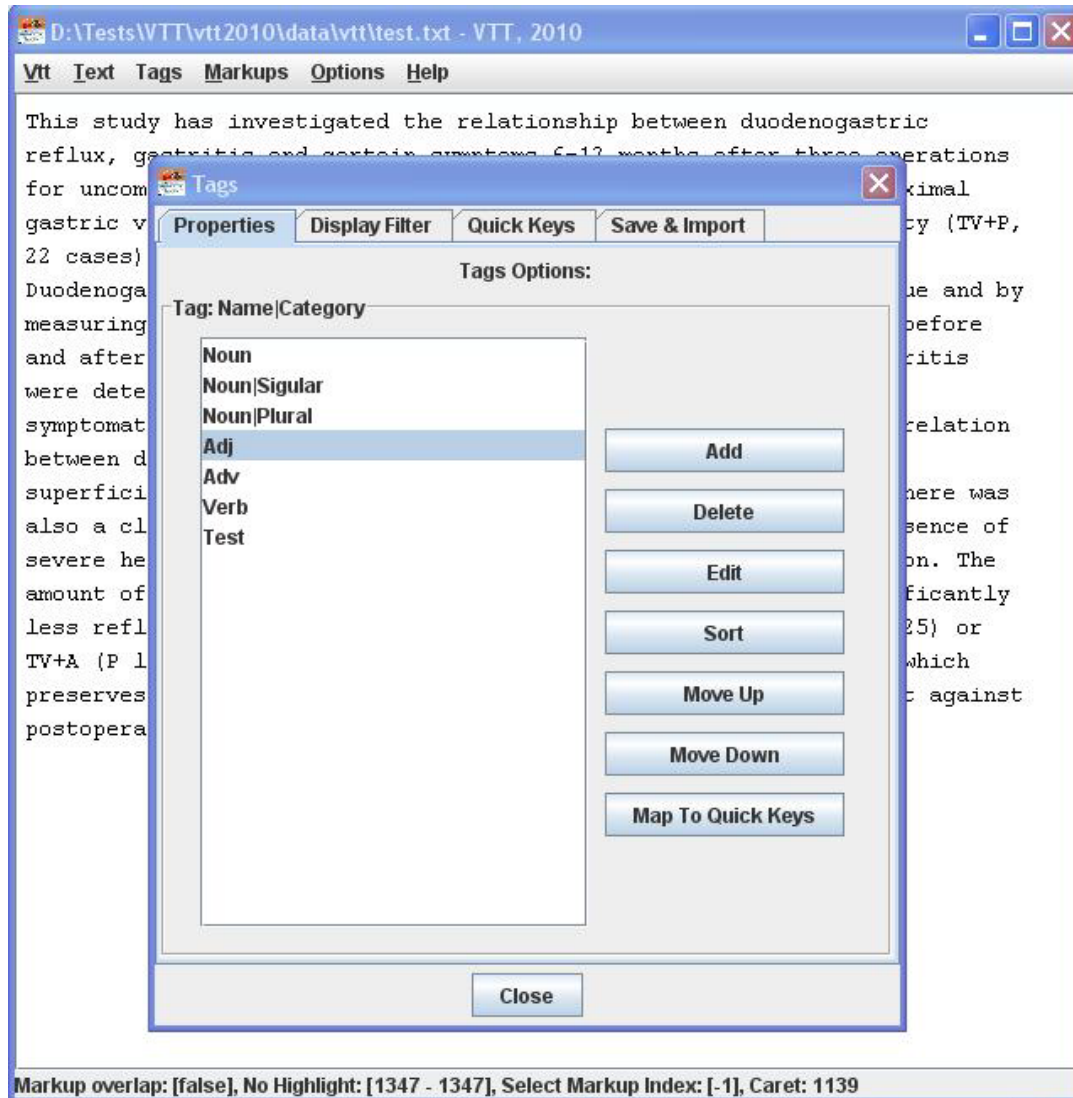
Text – Open (Cont.)



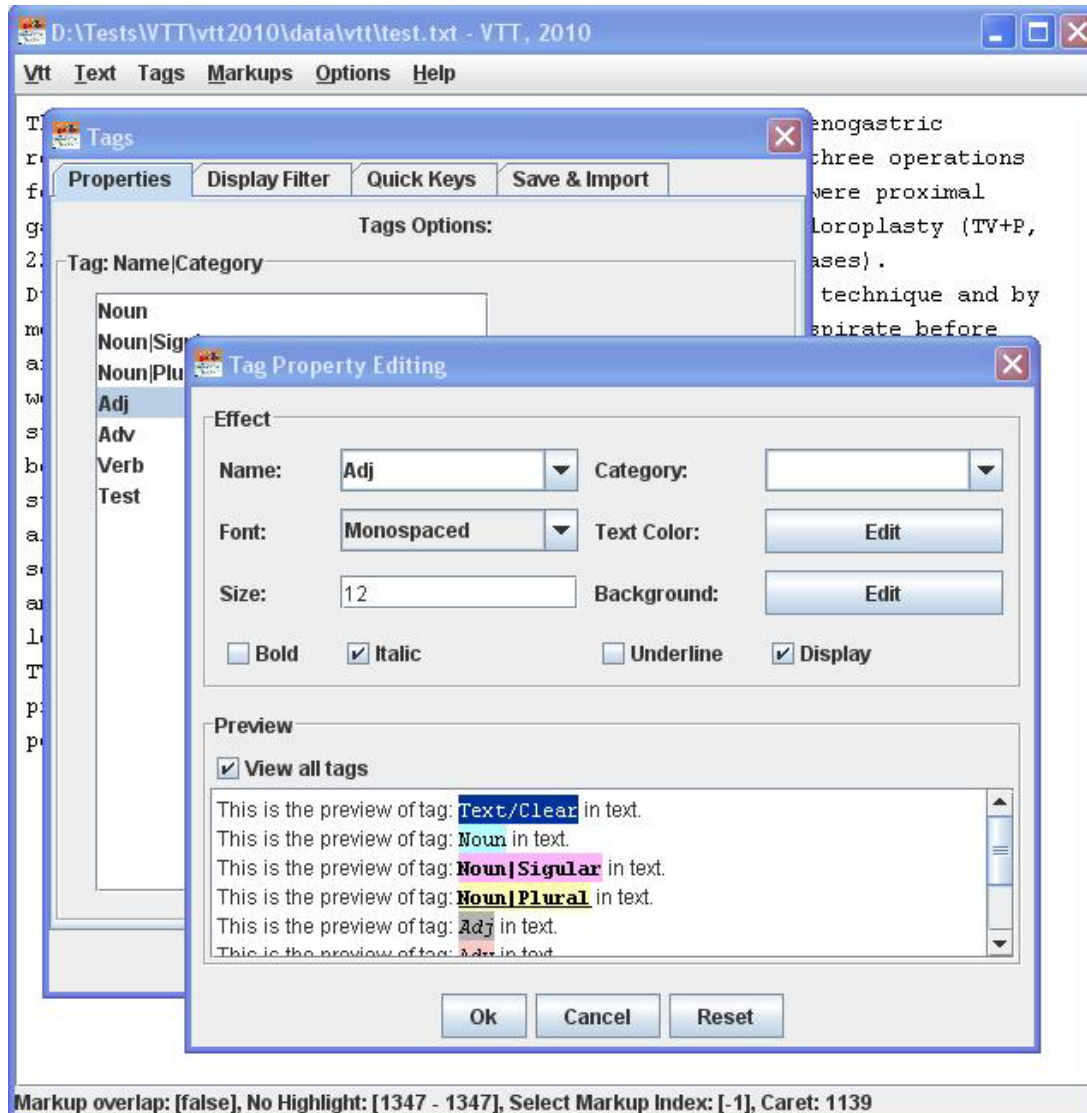
Text - Options



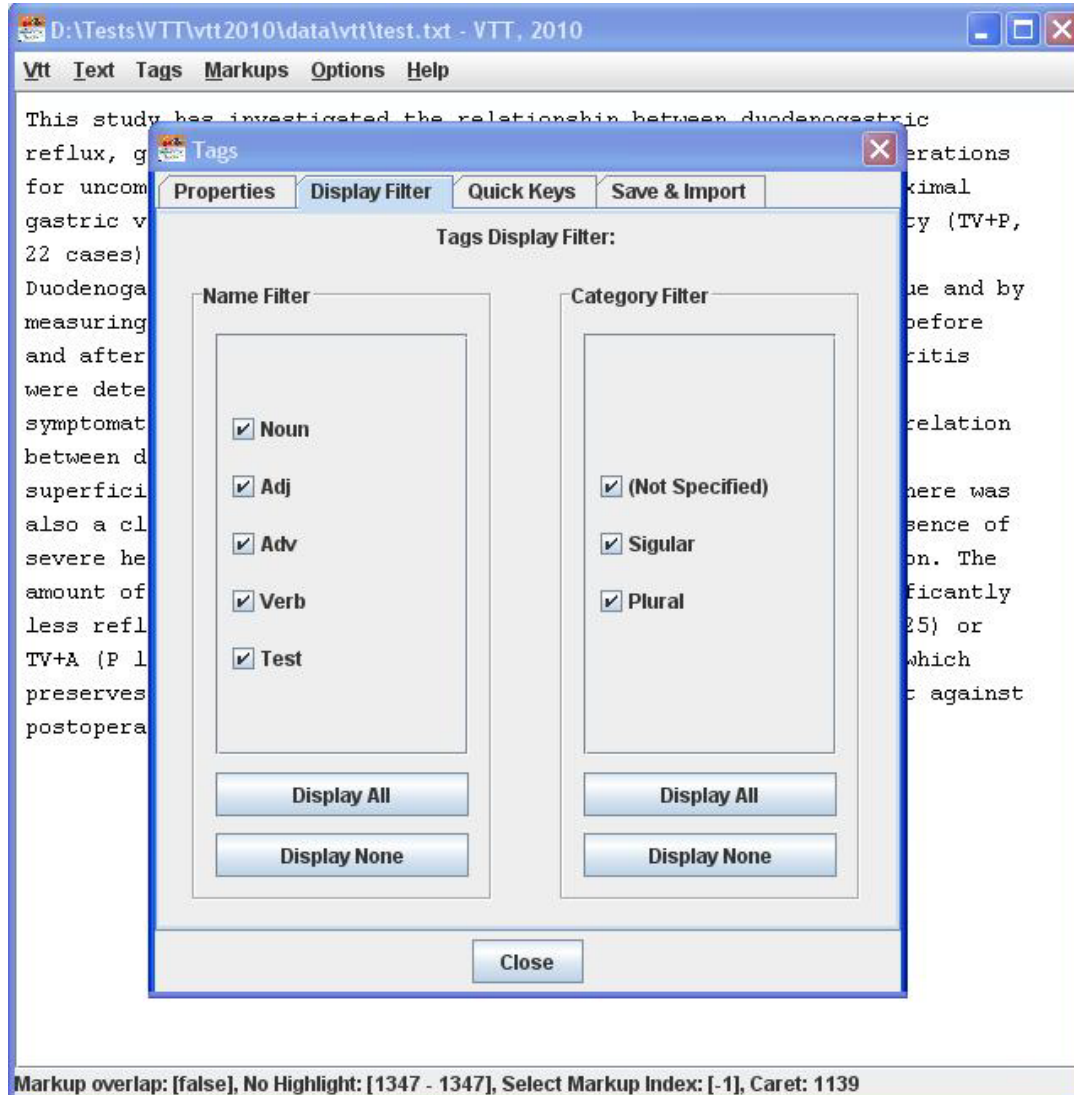
Tags - Setup



Tags - Edit



Tags – Display Filter



Tags - Quick Keys

D:\Tests\WTT\vt2010\data\vt\test.txt - VTT, 2010

Vtt Text Tags Markups Options Help

This study has investigated the relationship between duodenogastric reflux, g... operations
for uncom... kimal
gastric v... cy (TV+P,
22 cases) ue and by
Duodenoga... before
measuring... citis
and after... relation
were dete... here was
symptomat... sence of
between d... on. The
superfici... ficantly
also a cl... 25) or
severe he... which
amount of... c against
less refl...
TV+A (P l...
preserves...
postopera...

Tags

Properties Display Filter Quick Keys Save & Import

Tags: Quick Keys Mapping:

Quick Key Mapping

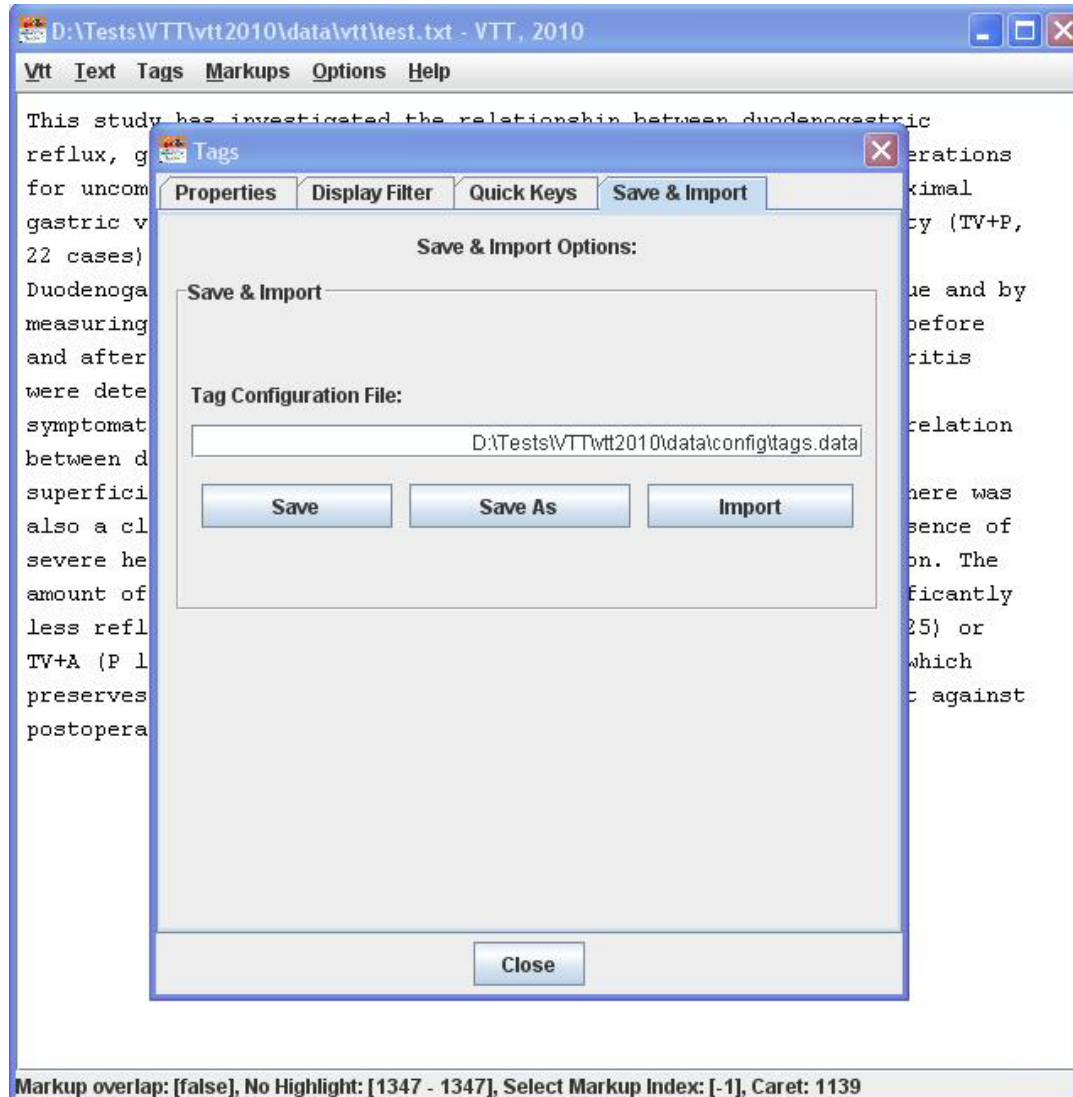
Key [1]:	Noun
Key [2]:	Noun Sigular
Key [3]:	Noun Plural
Key [4]:	Adj
Key [5]:	Adv
Key [6]:	Verb
Key [7]:	Test
Key [8]:	
Key [9]:	

Apply Reset

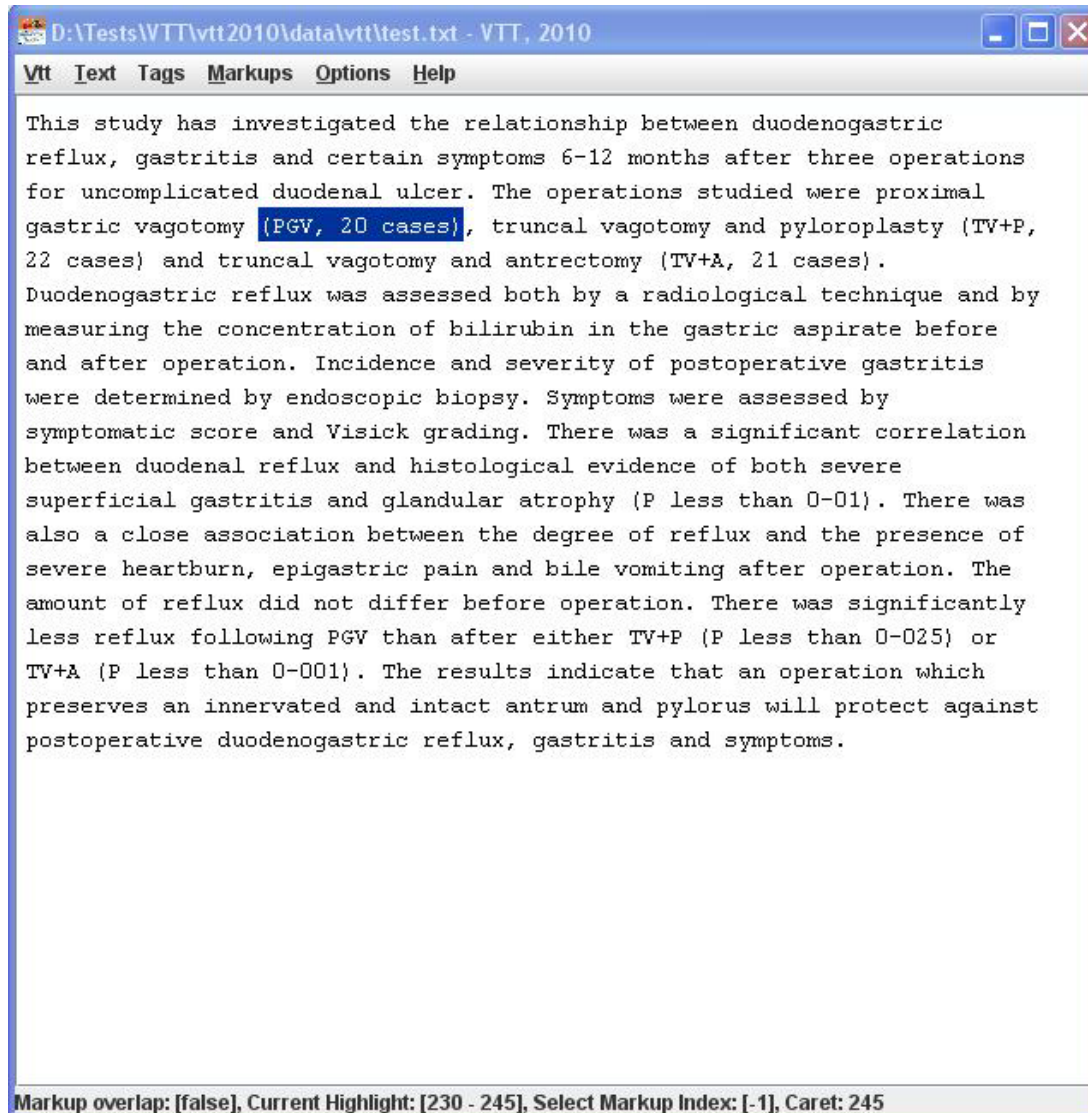
Close

Markup overlap: [false], No Highlight: [1347 - 1347], Select Markup Index: [-1], Caret: 1139

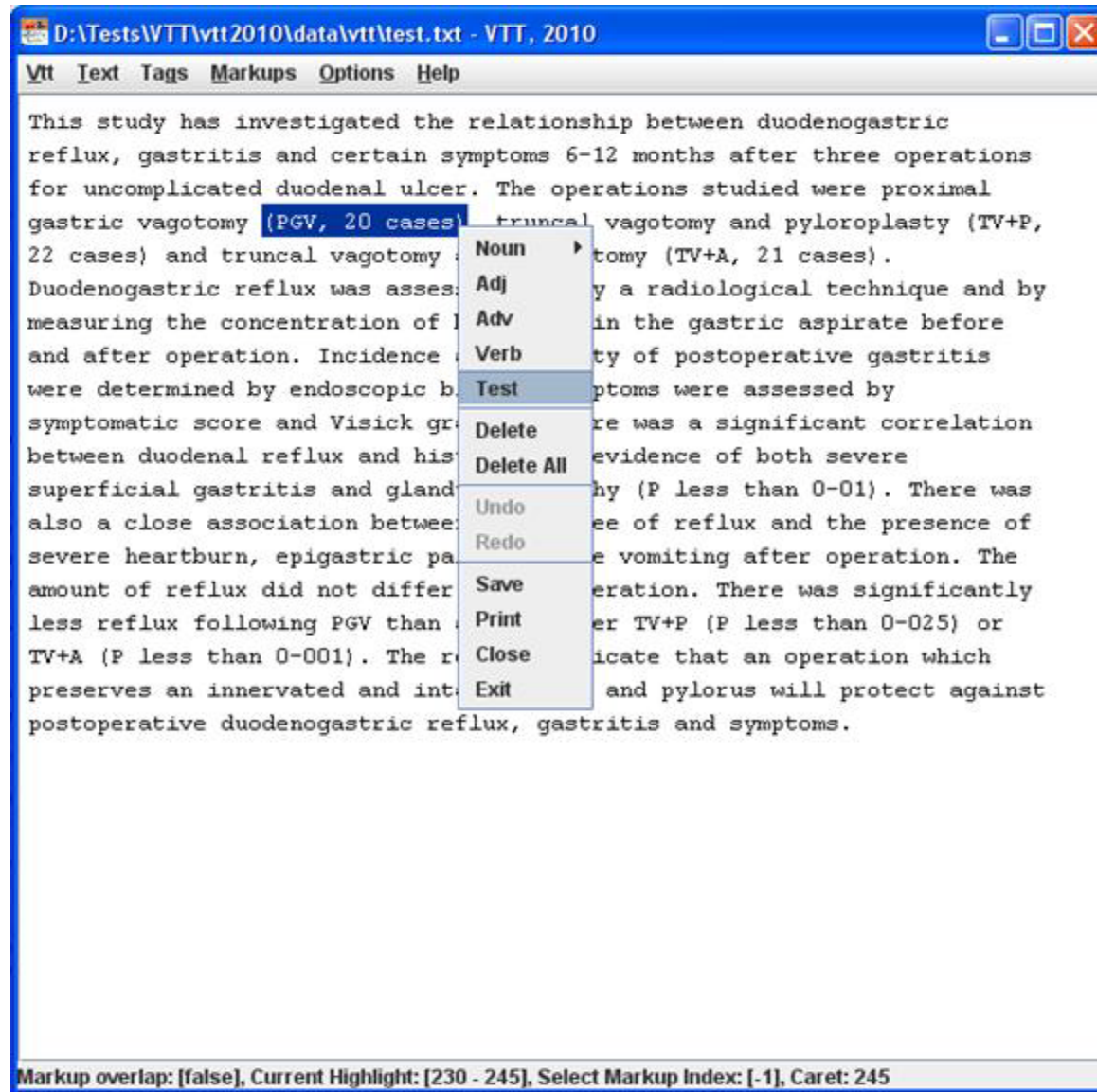
Tags – Save & Import



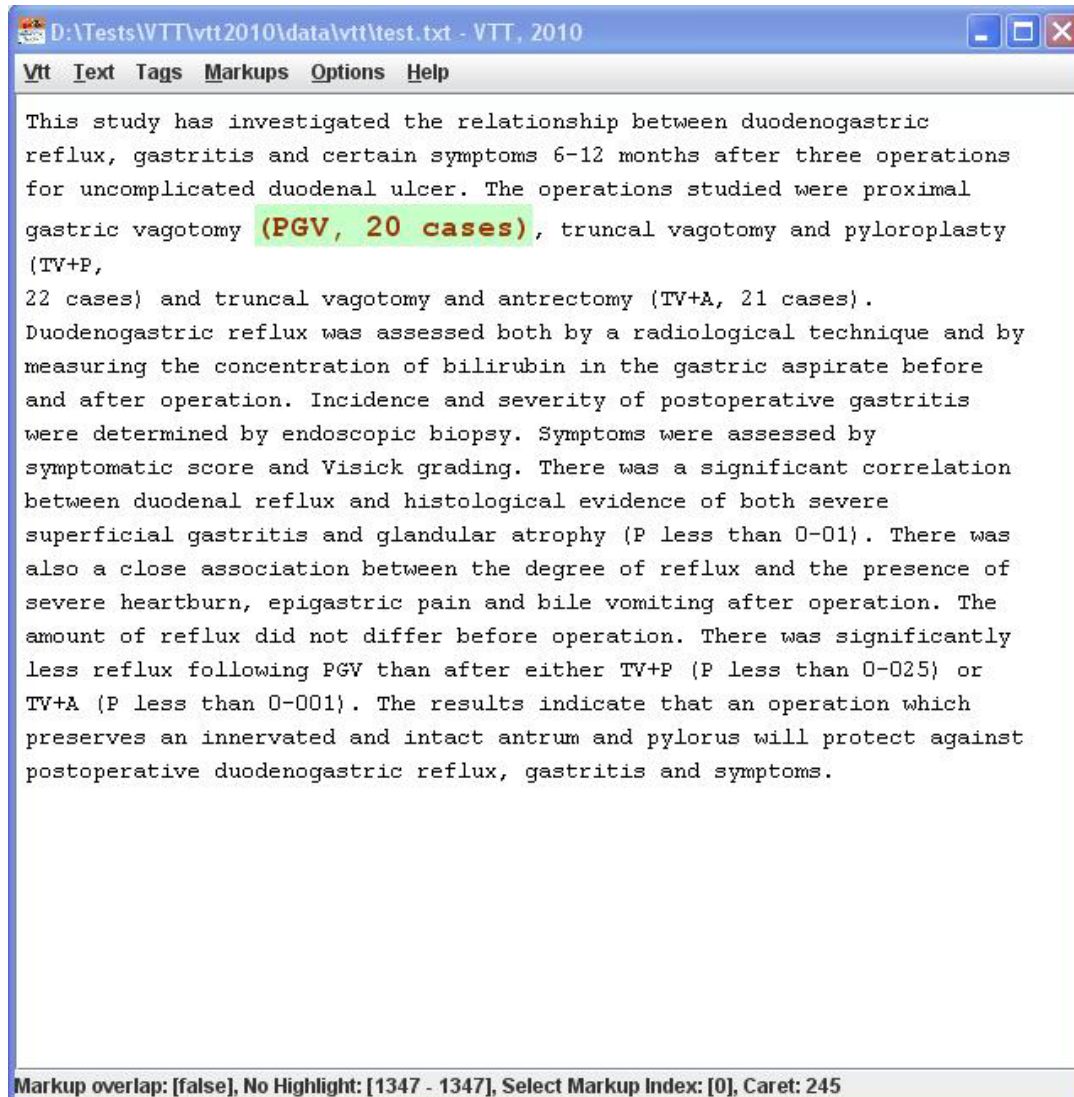
Markups – Select Text



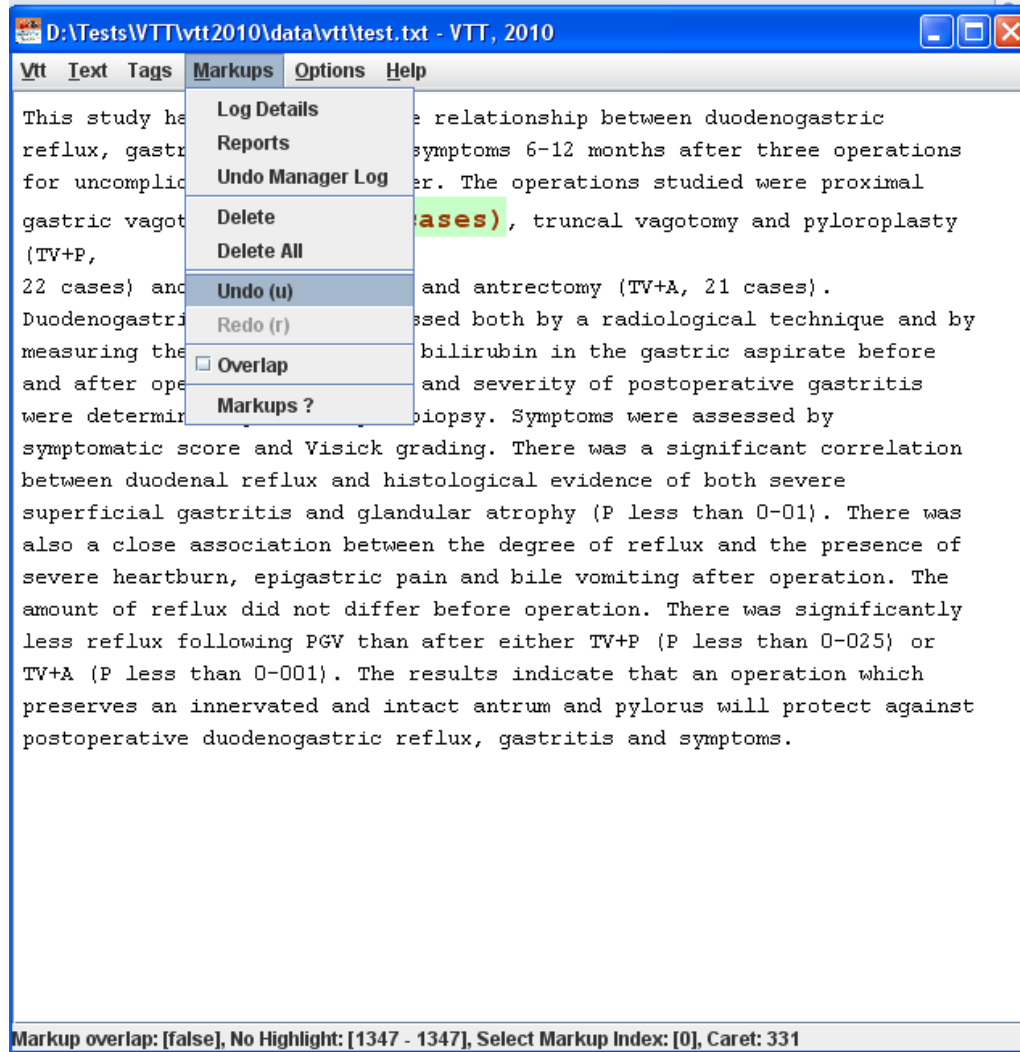
Markups - Assign a Tag



Markups - Assign a Tag



Markups - Options



The screenshot shows a text editor window titled "D:\Tests\WTT\vt2010\data\vt2\test.txt - VTT, 2010". The menu bar includes "Vtt", "Text", "Tags", "Markups", "Options", and "Help". The "Markups" menu is open, displaying the following options: "Log Details", "Reports", "Undo Manager Log", "Delete", "Delete All", "Undo (u)", "Redo (r)", "Overlap" (with an unchecked checkbox), and "Markups?". The text in the background is a medical study abstract. A status bar at the bottom reads: "Markup overlap: [false], No Highlight: [1347 - 1347], Select Markup Index: [0], Caret: 331".

This study has investigated the relationship between duodenogastric reflux, gastric symptoms 6-12 months after three operations for uncomplicated peptic ulcer. The operations studied were proximal gastric vagotomy (PGV), truncal vagotomy and pyloroplasty (TV+P, 22 cases) and truncal vagotomy and antrectomy (TV+A, 21 cases). Duodenogastric reflux was assessed both by a radiological technique and by measuring the concentration of bilirubin in the gastric aspirate before and after operation. The severity of postoperative gastritis were determined by biopsy. Symptoms were assessed by symptomatic score and Visick grading. There was a significant correlation between duodenal reflux and histological evidence of both severe superficial gastritis and glandular atrophy (P less than 0.01). There was also a close association between the degree of reflux and the presence of severe heartburn, epigastric pain and bile vomiting after operation. The amount of reflux did not differ before operation. There was significantly less reflux following PGV than after either TV+P (P less than 0.025) or TV+A (P less than 0.001). The results indicate that an operation which preserves an innervated and intact antrum and pylorus will protect against postoperative duodenogastric reflux, gastritis and symptoms.

Markup overlap: [false], No Highlight: [1347 - 1347], Select Markup Index: [0], Caret: 331

Markups – Logs

The screenshot shows the VTT software interface. The main window displays a text document with the following content:

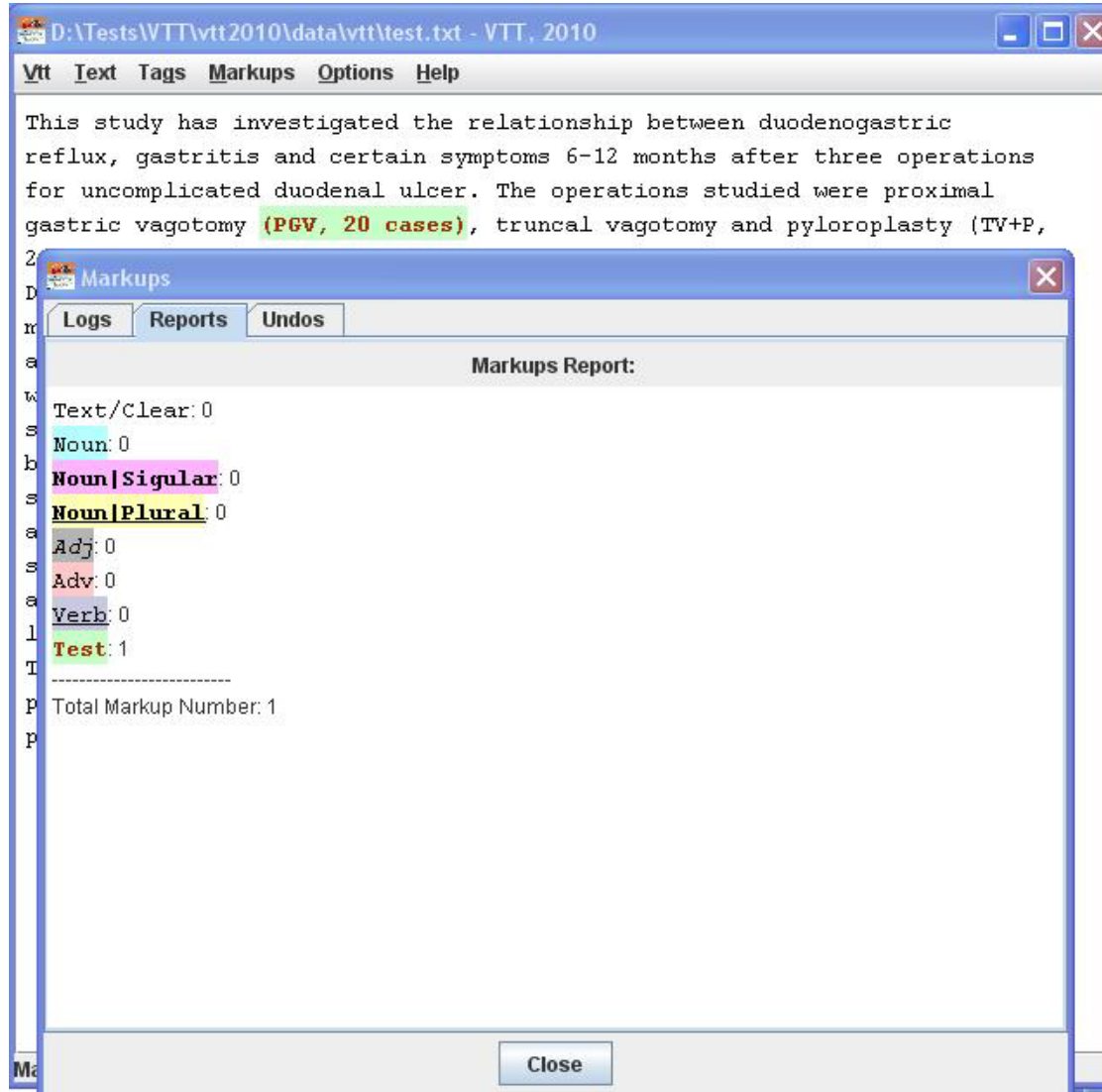
This study has investigated the relationship between duodenogastric reflux, gastritis and certain symptoms 6-12 months after three operations for uncomplicated duodenal ulcer. The operations studied were proximal gastric vagotomy (PGV, 20 cases), truncal vagotomy and pyloroplasty (TV+P,

The 'Markups' dialog box is open, showing the 'Logs' tab. The 'Markups Log Details' section displays the following information:

```
#<----->
#<MarkUps Information>
#<Offset|Length|TagName|TagCategory|Annotation|TagText>
#<----->
230|15|Test          |          ||(PGV, 20 cases)
```

The dialog box also includes a 'Close' button at the bottom.

Markups – Reports



Markups – More Features

- Delete selected markup (0)
- Assign a tag by quick keys (1~9)
- Markup redo (r) /undo (u)
- Markups join (j)
- Markups movement control by keys
- Selected Markup Information (m)

Markups – Informations

The screenshot shows the Vtt 2010 application window. The main text area contains a paragraph with several words highlighted in different colors: 'This' (yellow), 'study' (red), 'certain' (green), and 'months' (purple). A 'Markup' dialog box is open over the word 'study'. The dialog box has a title bar 'Markup' and a close button. The main title of the dialog is 'Modify Markup:'. It contains the following fields:

- Selected Markup:** A text box containing '[study]'
- Tagged Text:** A text box containing '[study]'
- Tag Name:** A dropdown menu with 'Noun' selected.
- Tag Category:** A dropdown menu with 'Singular' selected.
- Start Pos:** A text box containing '5'
- End Pos:** A text box containing '10'
- Annotation:** A text box containing 'Add notes here'

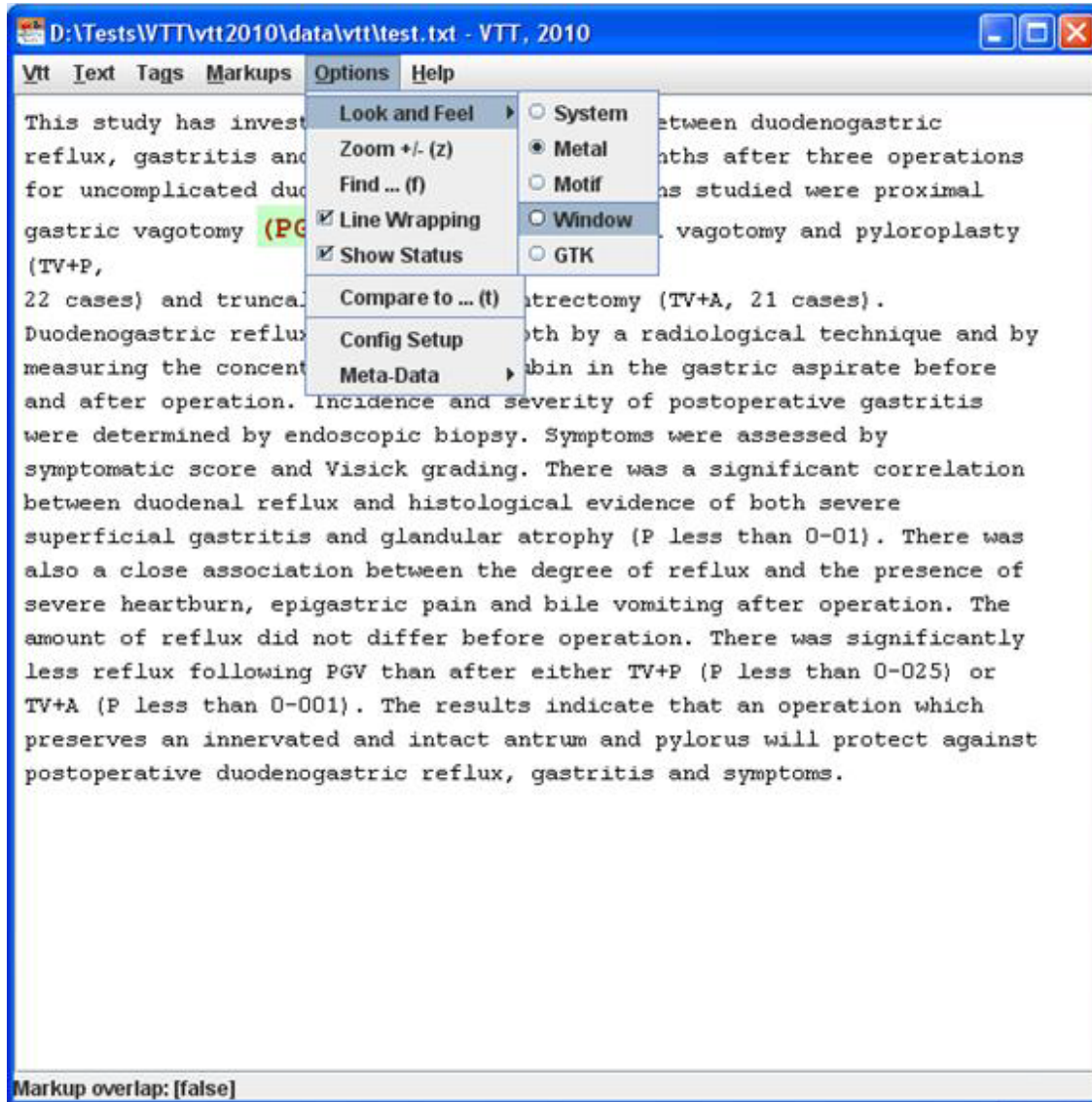
At the bottom of the dialog box are five buttons: 'Ok', 'Cancel', 'Apply', 'Reset', and 'Delete'. Below the main text area, a status bar displays: 'Markup overlap: [false], No Highlight: [10 - 10], Select Markup Index: [1], Caret: 10'.

Markups – Movement

- Markups related

Key	Action
[Set selected markup to the first markup
]	Set selected markup to the last markup
.	Backward selected markup
.	Forward selected markup
{	Set selected markup to the first displayed markup
}	Set selected markup to the last displayed markup
<	Backward selected displayed markup
>	Forward selected displayed markup
j	Join the selected markup and the next markup
m	Shows markup detail dialog on the selected markup
r	Redo markup operation
u	Undo markup operation
0	Delete selected markup
1 ~ 9	Set selected markup/highlight text to tag (1 ~ 9)

Other Options



Compare Option

The screenshot shows a software window titled "VTT Compare: Different Markups from Source and Target Files". The window has a menu bar with options: "View Markups Diff", "View Data Diff", "Stats Reports", "Difference", and "Raw Data". The main area is split into two columns: "Markups in Source File" and "Markups in Target File". Both columns contain the same text, which is a medical study abstract. The word "symptoms" is highlighted in yellow in both columns. At the bottom of the window, there are two buttons: "Ok" and "Open Target File".

View Markups Diff **View Data Diff** **Stats Reports** **Difference** **Raw Data**

Markups in Source File

This study has investigated the relationship between duodenogastric reflux, gastritis and certain symptoms 6-12 months after three operations for uncomplicated duodenal ulcer. The operations studied were proximal gastric vagotomy (PGV, 20 cases), truncal vagotomy and pyloroplasty (TV+P, 22 cases) and truncal vagotomy and antrectomy (TV+A, 21 cases). Duodenogastric reflux was assessed both by a radiological technique and by measuring the concentration of bilirubin in the gastric aspirate before and after operation. Incidence and severity of postoperative gastritis were determined by endoscopic biopsy. Symptoms were assessed by symptomatic score and Visick grading. There was a significant correlation between duodenal reflux and histological evidence of both severe superficial gastritis and glandular atrophy (P less than 0-01). There was also a close association between the degree of reflux and the presence of severe heartburn, epigastric pain and bile vomiting after operation. The amount of reflux did not differ before operation. There was significantly less reflux following PGV than after either TV+P (P less than 0-025) or TV+A (P less than 0-001). The results indicate that an operation which preserves an innervated and intact antrum and pylorus will protect against postoperative duodenogastric reflux, gastritis and symptoms.

Markups in Target File

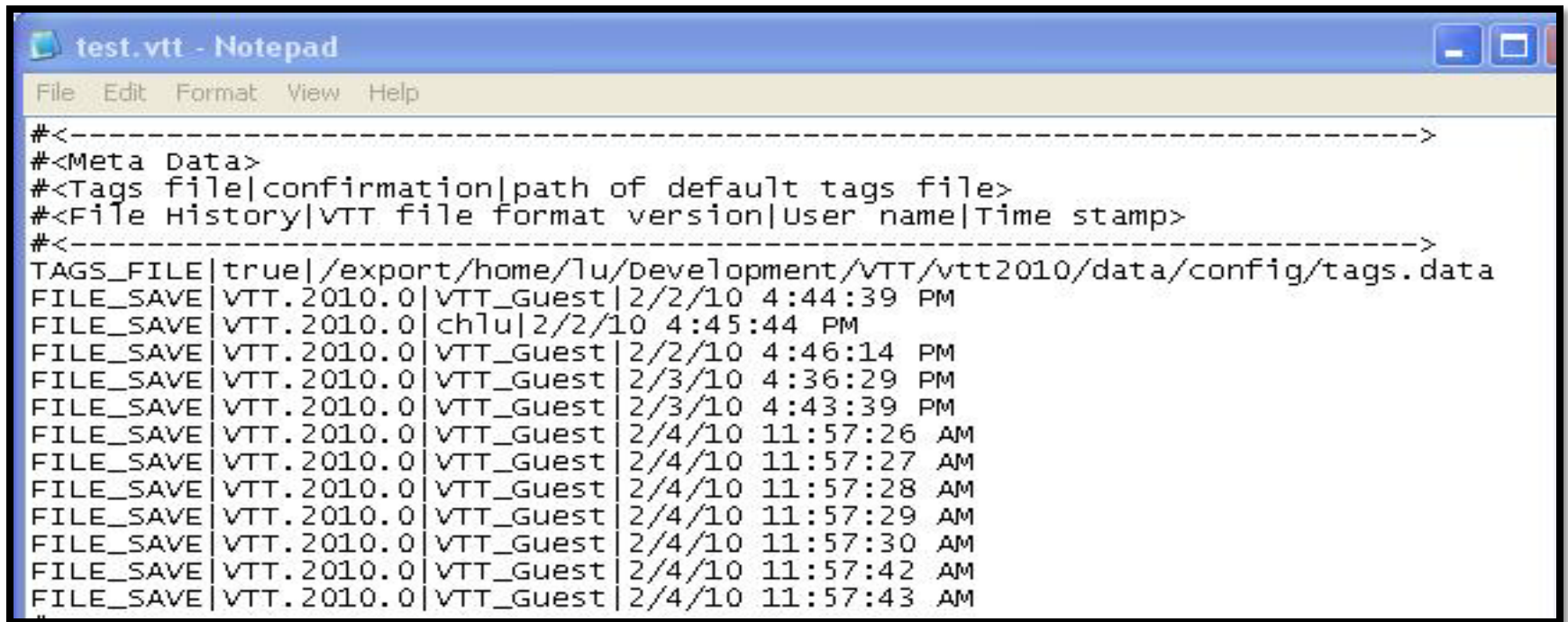
This study has investigated the relationship between duodenogastric reflux, gastritis and certain symptoms 6-12 months after three operations for uncomplicated duodenal ulcer. The operations studied were proximal gastric vagotomy (PGV, 20 cases), truncal vagotomy and pyloroplasty (TV+P, 22 cases) and truncal vagotomy and antrectomy (TV+A, 21 cases). Duodenogastric reflux was assessed both by a radiological technique and by measuring the concentration of bilirubin in the gastric aspirate before and after operation. Incidence and severity of postoperative gastritis were determined by endoscopic biopsy. Symptoms were assessed by symptomatic score and Visick grading. There was a significant correlation between duodenal reflux and histological evidence of both severe superficial gastritis and glandular atrophy (P less than 0-01). There was also a close association between the degree of reflux and the presence of severe heartburn, epigastric pain and bile vomiting after operation. The amount of reflux did not differ before operation. There was significantly less reflux following PGV than after either TV+P (P less than 0-025) or TV+A (P less than 0-001). The results indicate that an operation which preserves an innervated and intact antrum and pylorus will protect against postoperative duodenogastric reflux, gastritis and symptoms.

Ok **Open Target File**

VTT – File Format

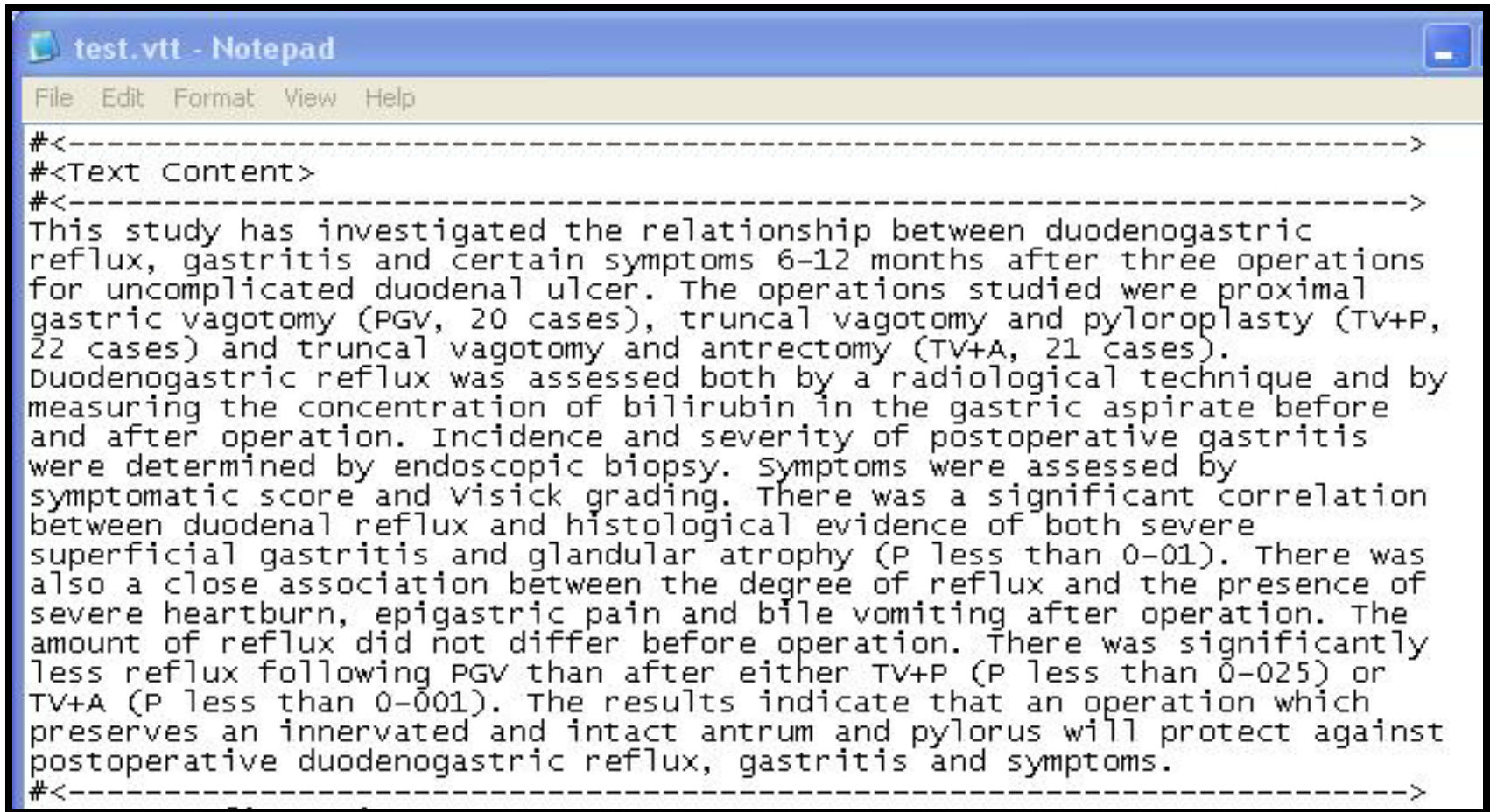
- **Meta Data**
 - Default tag file (3 fields)
 - History (4 fields)
- **Text Content**
 - Original plain text
- **Tags Configuration (14 fields)**
 - Name and category
 - Bold, italic, underline, display, font family, font size
 - Foreground and background colors (RGB)
- **Markups Information (6 fields)**
 - Offset and length
 - Assign tag (name and category)
 - Annotation and tagged text

File Format – Meta Data

A screenshot of a Notepad window titled "test.vtt - Notepad". The window has a menu bar with "File", "Edit", "Format", "View", and "Help". The text content is as follows:

```
#<----->
#<Meta Data>
#<Tags file|confirmation|path of default tags file>
#<File History|VTT file format version|User name|Time stamp>
#<----->
TAGS_FILE|true|/export/home/lu/Development/VTT/vtt2010/data/config/tags.data
FILE_SAVE|VTT.2010.0|VTT_Guest|2/2/10 4:44:39 PM
FILE_SAVE|VTT.2010.0|chlu|2/2/10 4:45:44 PM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/2/10 4:46:14 PM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/3/10 4:36:29 PM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/3/10 4:43:39 PM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:26 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:27 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:28 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:29 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:30 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:42 AM
FILE_SAVE|VTT.2010.0|VTT_Guest|2/4/10 11:57:43 AM
```

File Format – Text



```
test.vtt - Notepad
File Edit Format View Help
#<----->
#<Text Content>
#<----->
This study has investigated the relationship between duodenogastric
reflux, gastritis and certain symptoms 6-12 months after three operations
for uncomplicated duodenal ulcer. The operations studied were proximal
gastric vagotomy (PGV, 20 cases), truncal vagotomy and pyloroplasty (TV+P,
22 cases) and truncal vagotomy and antrectomy (TV+A, 21 cases).
Duodenogastric reflux was assessed both by a radiological technique and by
measuring the concentration of bilirubin in the gastric aspirate before
and after operation. Incidence and severity of postoperative gastritis
were determined by endoscopic biopsy. Symptoms were assessed by
symptomatic score and visick grading. There was a significant correlation
between duodenal reflux and histological evidence of both severe
superficial gastritis and glandular atrophy (P less than 0-01). There was
also a close association between the degree of reflux and the presence of
severe heartburn, epigastric pain and bile vomiting after operation. The
amount of reflux did not differ before operation. There was significantly
less reflux following PGV than after either TV+P (P less than 0-025) or
TV+A (P less than 0-001). The results indicate that an operation which
preserves an innervated and intact antrum and pylorus will protect against
postoperative duodenogastric reflux, gastritis and symptoms.
#<----->
```

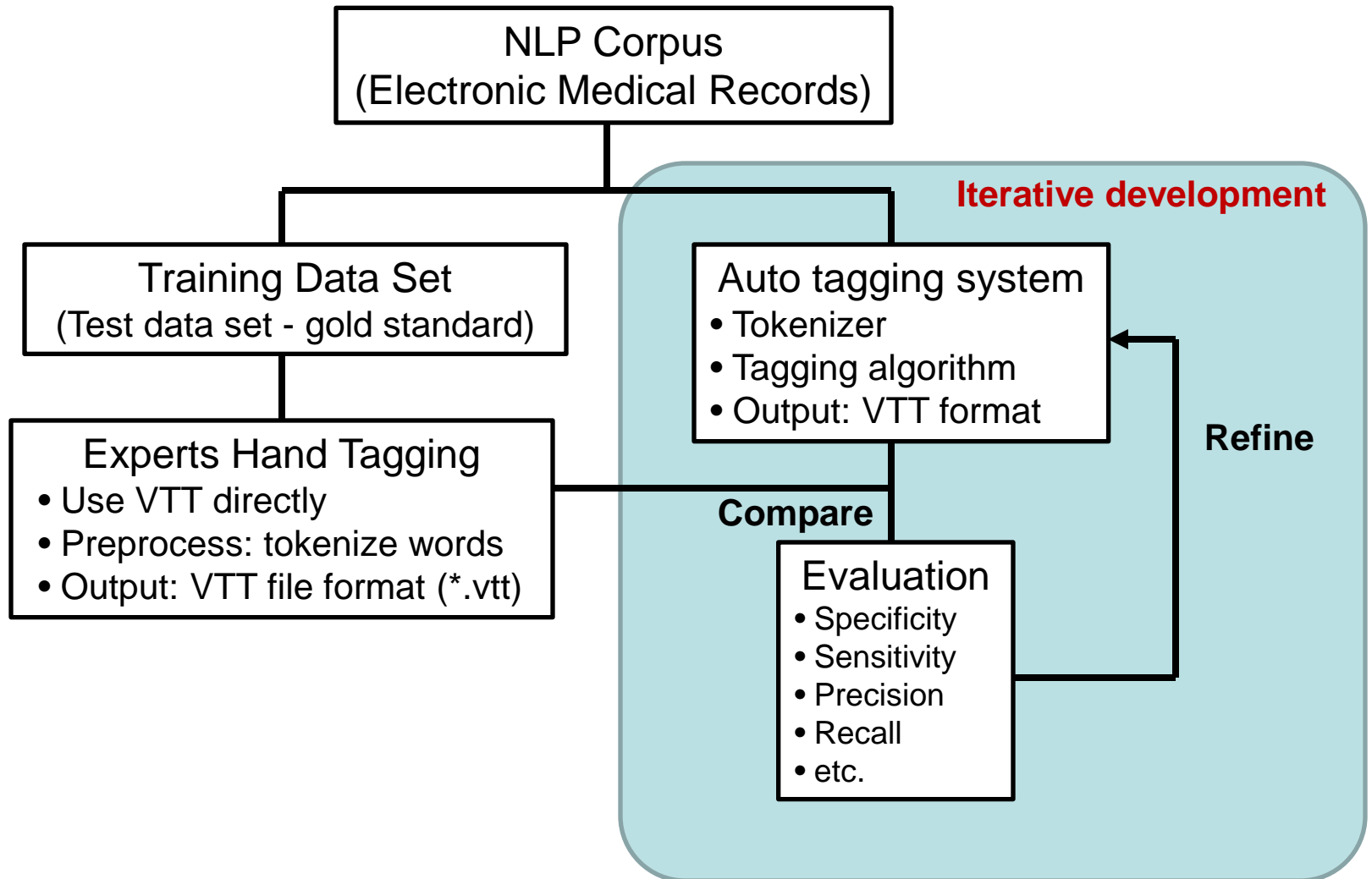
File Format – Tags

```
test.vtt - Notepad
File Edit Format View Help
#<----->
#<Tags Configuration>
#<Name|Category|Bold|Italic|Underline|Display|FR|FG|FB|BR|BG|BB|FontFamily|FontSize>
#<----->
Text/Clear||false|false|false|true|255|255|255|0|51|153|Monospaced|12
#<----->
Noun||false|false|false|true|0|0|0|255|180|180|Monospaced|12
Noun|singular|true|false|false|true|0|0|0|255|180|180|Monospaced|12
Noun|Plural|true|false|true|true|0|0|0|255|180|180|Monospaced|12
Adj||false|true|false|true|0|0|0|180|225|180|Monospaced|12
Adv||false|false|false|true|0|0|0|204|204|255|Monospaced|12
Verb||false|false|true|true|0|0|0|150|196|225|Monospaced|12
Test||true|false|false|true|153|51|0|255|255|204|Monospaced|12
#<----->
```

File Format - Markups

```
test.vtt - Notepad
File Edit Format View Help
#<----->
#<MarkUps Information>
#<Offset|Length|TagName|TagCategory|Annotation|TagText>
#<----->
  0 | 4 | Test          |          | First Test|This
  5 | 5 | Noun         | singular | study
 90 | 7 | Adj          |          | certain
112 | 6 | Adv          |          | months
247 | 6 | Verb         |          | trunca
367 | 6 | Adv          |          | reflux
373 | 1 | Noun         | singular |
374 | 3 | Adv          |          | was
377 | 1 | Noun         |          |
378 | 8 | Test         |          | assessed
387 | 4 | Adv          |          | both
392 | 2 | Noun         | plural   | by
397 |12 | Adj          |          | radiological
509 |12 | Adj          |          | operation. I
519 | 2 | Text/clear   |          | I
600 | 7 | Adv          |          | Extra space and S, need to
delete|biopsy.
617 | 4 | Noun         | singular | were is a be verb|were
736 |12 | Noun         | singular | histological
894 | 6 | Noun         | plural   | reflux
939 |10 | Noun         |          | epigastric
955 | 3 | Verb         |          | and
1337| 9 | Adj          |          | Last test|symptoms.
```

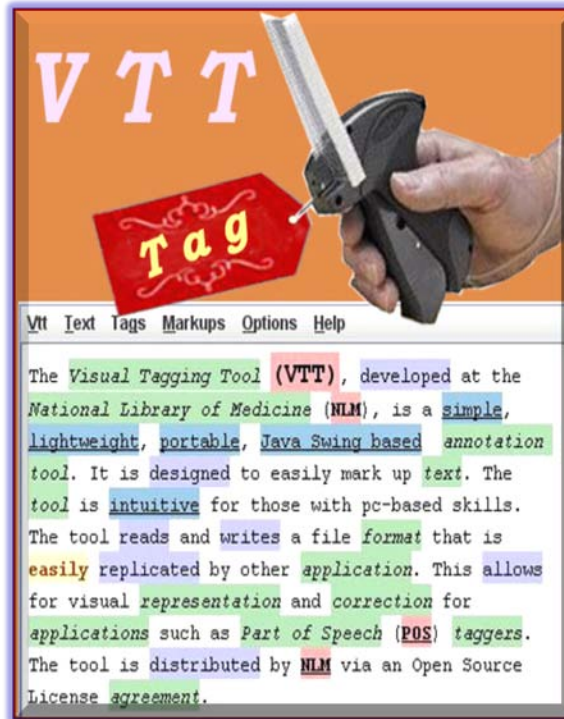
VTT – Example



VTT – Integration

- Use VTT file format for the NLP applications
- Integrate with words/terms tokenizer (text tools) to generate pre-tokenized VTT files for less key strokes
- Read in VTT files and apply evaluation software:
 - File level
 - Corpus level

Visual Tagging Tool (VTT)



<http://SPECIALIST.nlm.nih.gov/vtt>

Questions



- Lexical Systems Group: <http://umlslex.nlm.nih.gov>
- The SPECIALIST NLP Tools: <http://specialist.nlm.nih.gov>