Is it OK for a machine-readable text to become more complex than the original edition? -- The case of graphological study of the letters of Margaret Paston.
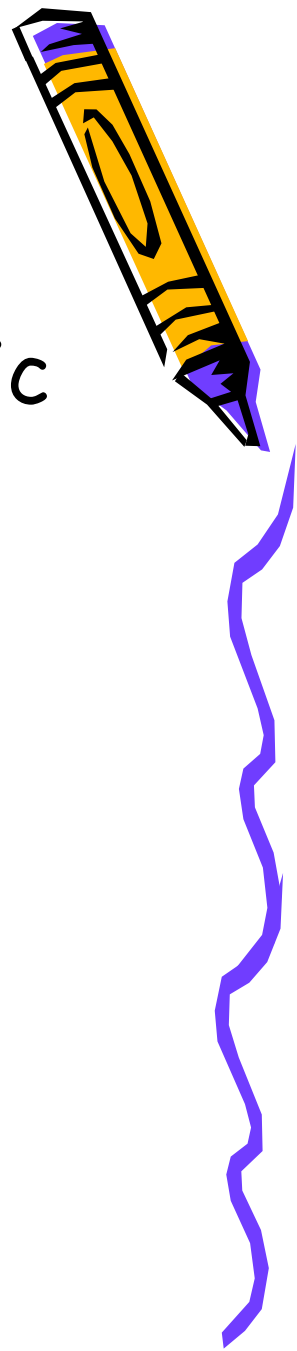
Nov. 2, 2007
TEI@20
Osamu Ohara
(Jikei University School of Medicine)

LP and GP

- Linguistic profile (LP) and graphetic Profile (GP) (McIntosh, 1974)

- LP: deals with graphemic choices.
- eg. *pat* vs. *that*
- GP: deals with the actual representation of each grapheme.

Cursive and non-cursive <w>

- Classified into three groups.
1. Amanuenses who use cursive {w}.
   → Gloys, Edmond II, Mowth.
2. those who use non-cursive {w}
   → John II, John III, Pampyng, Wykes.
3. and those who use both.
   → Calle, Daubeney, Gresham.

Allographs in GP

- Not only this {w}, but also other graphemes such as <e>, <g>, <r>, <s>, <v> and the "Tironian et" have various characteristic allographic forms according to each amanuensis.

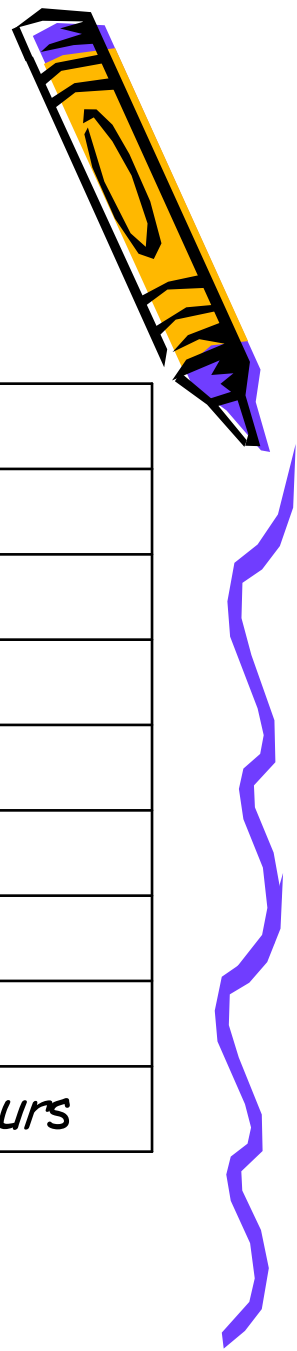- → also helpful for confirming the amanuensis of each letter.

Two sub-categories in GP

- graphetic variation
→ appearing as actually written forms of each allograph.

- allographic variation
→ resulting from the cursive and non-cursive distinction and the difference of handwriting styles.
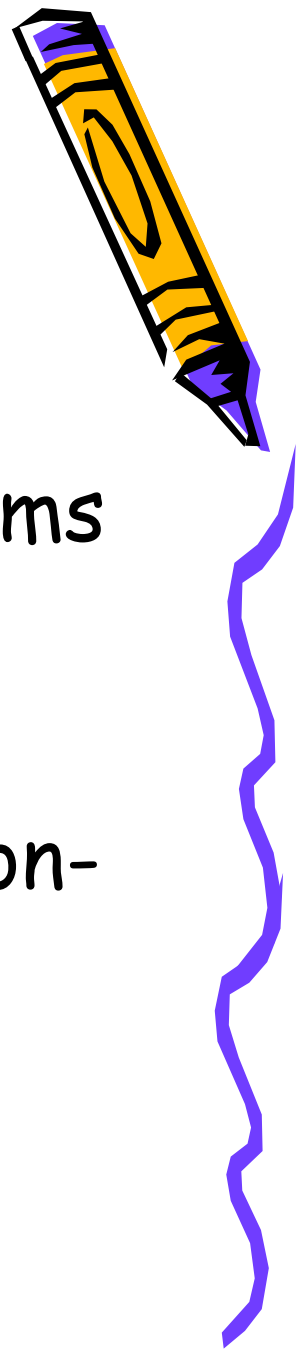
No. 158 in Margaret Letters

no.158	Gresham
haue	haue
right	right
seid	seid, seide, seyd
shall / shuld	shall / shuld, schulde
suyche	suyche
up-on	up-on
wherby / wherfor	wherby / wherfor, wherefor
ye / you / your / yours	ye / you, yow / your, your / yours

Gresham's short <s> at the end of the word

Two short <s> allographs in Margaret by Gresham.

	No.	129	158	197	Total
End	Short 1	84	11	12	107
	Shrot 2	0	12	0	12

Gresham's short <s> at the end of the word

Two short <s> allographs in Margaret by Gresham.

|	No.	129	158	197	Total
End	Short 1	84	11	12	107
	Shrot 2	0	12	0	12
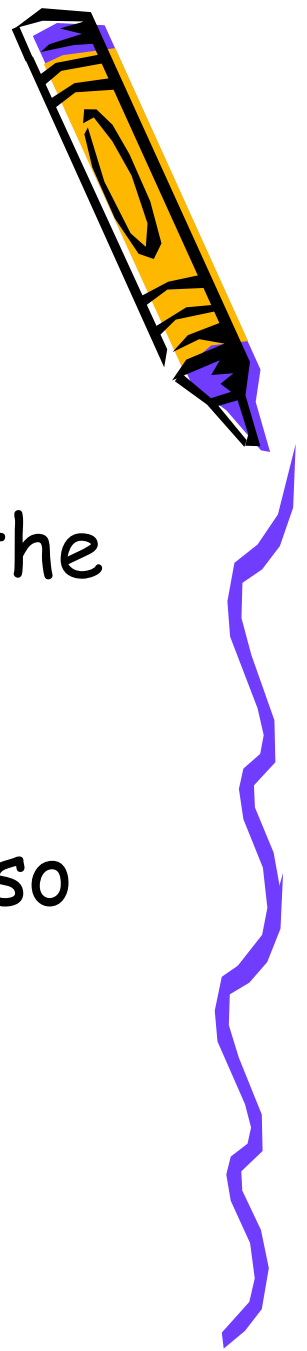
What is troubling us is ….

- Gresham may have known the different handwriting style from the one we usually see in his letters.
- Why do these very sporadically appearing characteristics appear so constantly in the letter no.158?

Graphological examination of the *Paston Letters*

- If we can reconfirm who wrote what part of the letters we can proceed into the study of idiolects,
- which will lead us to more profound understanding concerning the development of the Modern English.
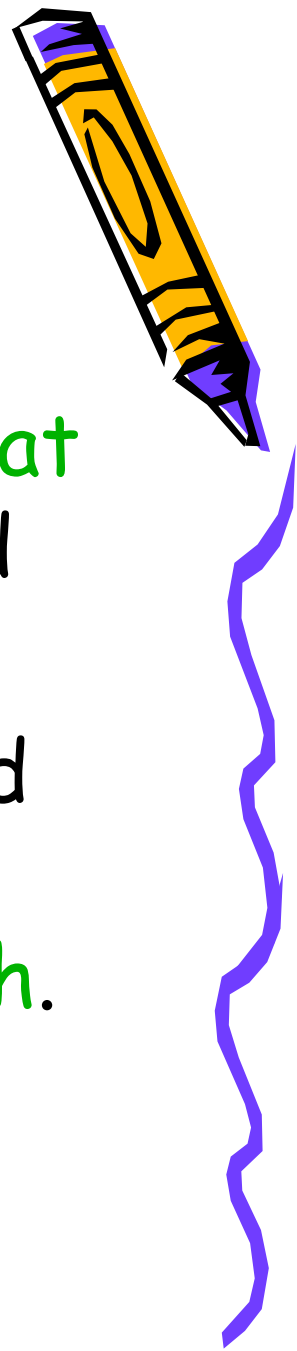
Graphological variants

```
<w type="adjective" lemma="great">
  <g ref="#GRg1">g</g><g ref="#GRr1">r</r>
  <g ref="#GRe1">e</e>t</w>


<w type="adjective" lemma="great">
  <g ref="#GRg2">g</g><g ref="#GRr2">r</r>
  <g ref="#GRe2">e</e>t</w>
```
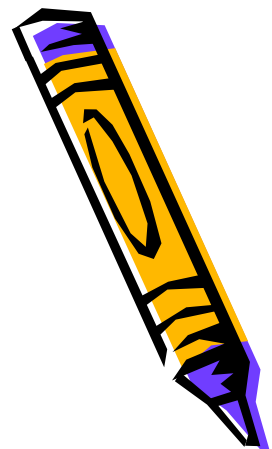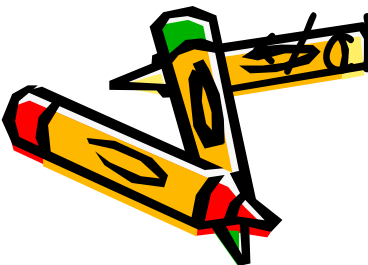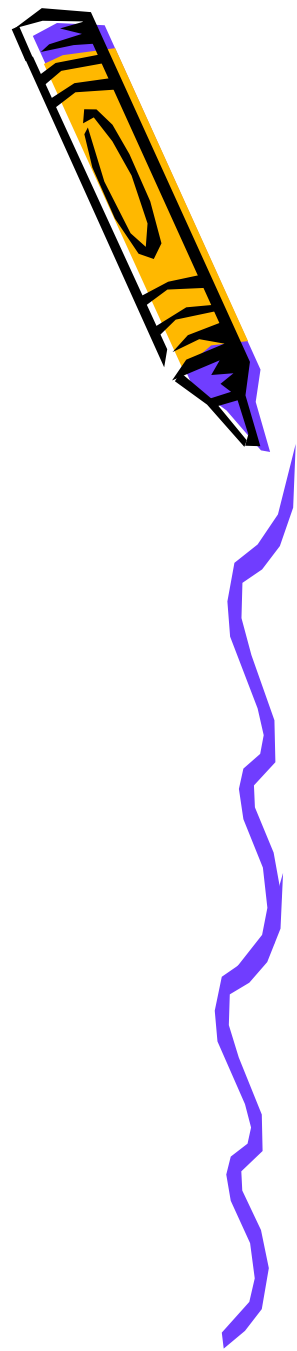
Graphological variants

```
<glyph xml:id="GRg1">
  <glyphName>
    Anglicana allograph of &lt;g&gt; by Gresham
  </glyphName>
  <charProp>
    <localName>entity</localName>
    <value>GRg1</value>
  </charProp></glyph>
```

Graphological variants

```
<glyph xml:id="GRg1">
  <glyphName>
    Anglicana allograph of &lt;g&gt; by Gresham
  </glyphName>
  <charProp>
    <localName>entity</localName>
    <value>GRg1</value>
  </charProp></glyph>
```
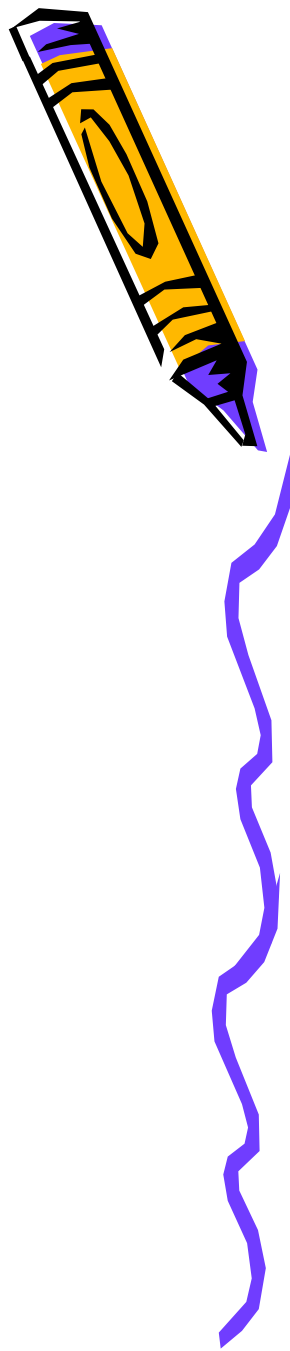
Graphological variants

```
<w type="adjective" lemma="great">
  <g ref="#GRg1">g</g><g ref="#GRr1">r</r>
  <g ref="#GRe1">e</e>t</w>

<w type="adjective" lemma="great">
  <g ref="#GRg2">g</g><g ref="#GRr2">r</r>
  <g ref="#GRe2">e</e>t</w>
```
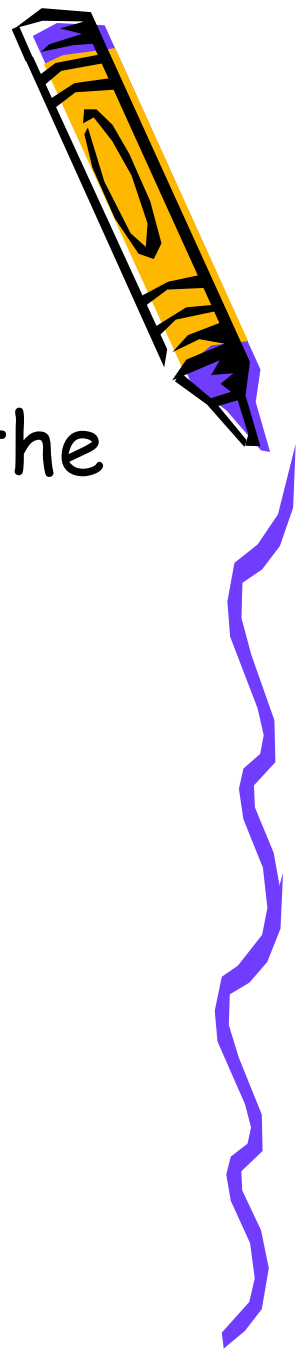
Making the Word List in XML

- Want to make the word list from the XML files of the *Paston Letters*.
- Will add as many information concerning each word available for our linguistic analysis as possible.

Abbreviation and Expansion
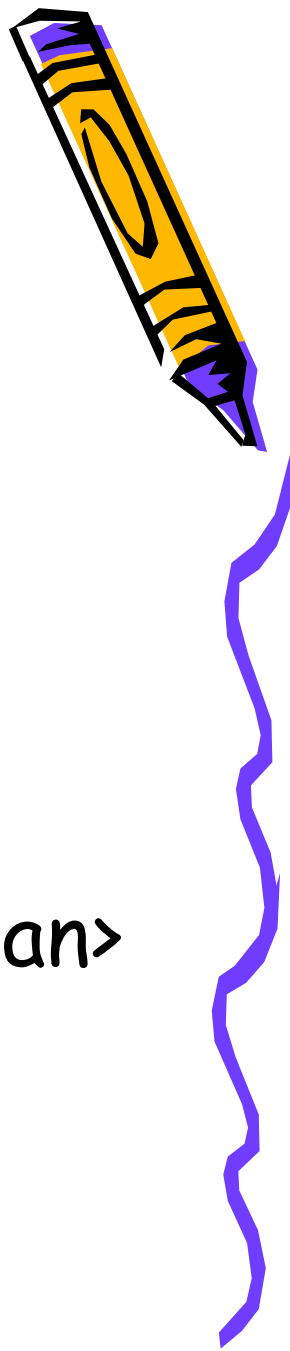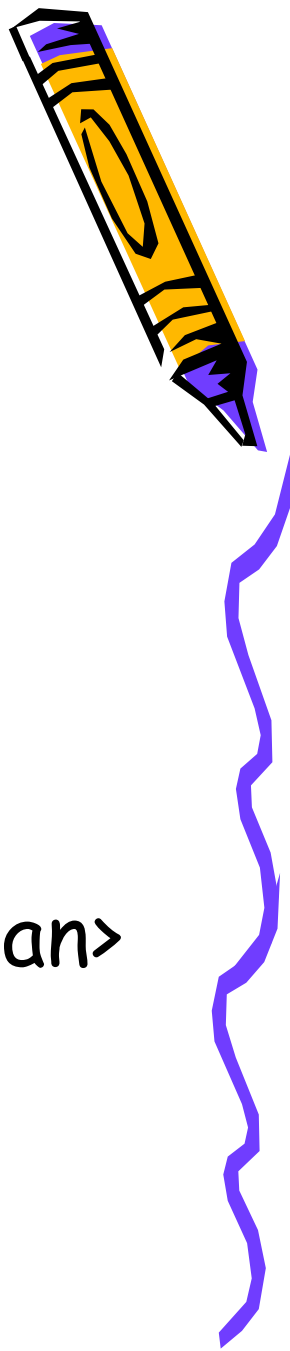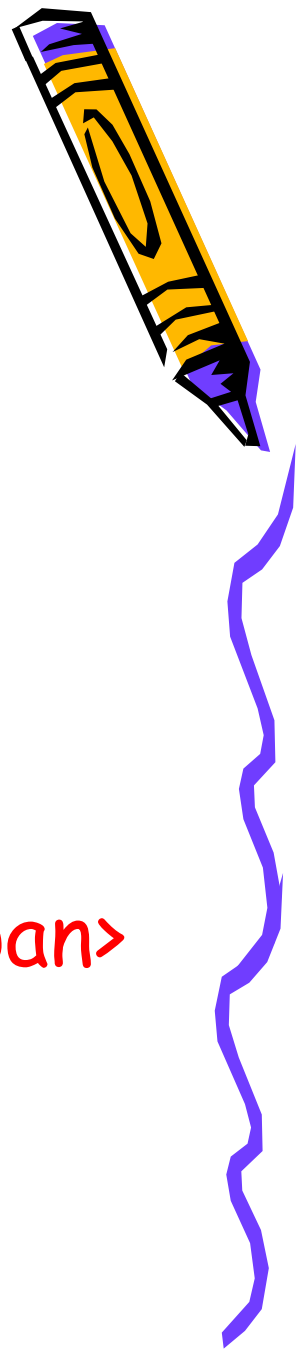
<w type="verb" lemma="speak">

spake

</w>

<w type="preposition" lemma="with">

 <expan>w<ex>i</ex>t<ex>h</ex></expan>

</w>

Abbreviation and Expansion

<w type="verb" lemma="speak">

spake

</w>

<w type="preposition" lemma="with">

 <expan>w<ex>i</ex>t<ex>h</ex></expan>
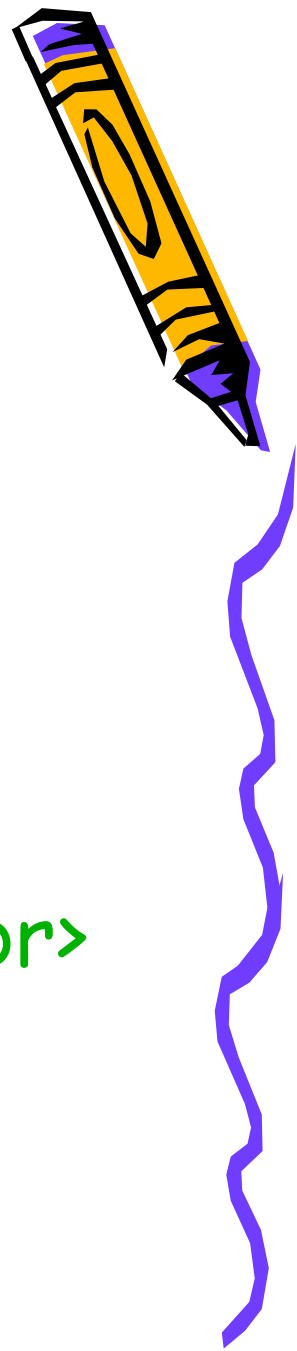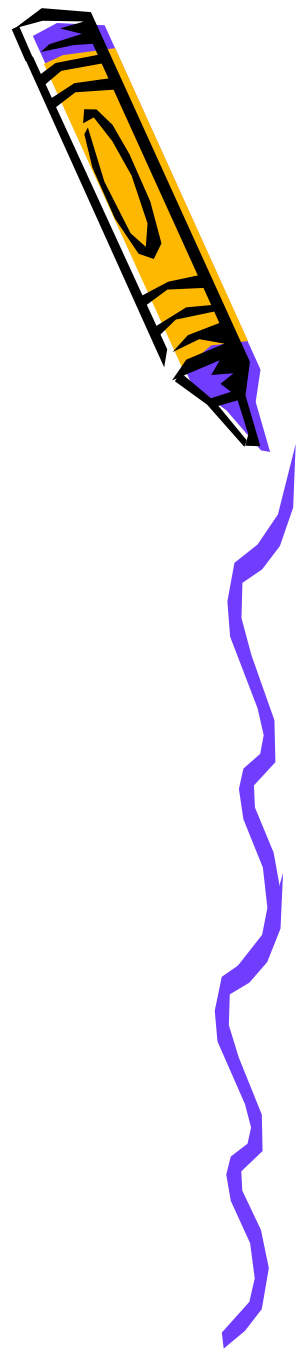
</w>

Abbreviation and Expansion

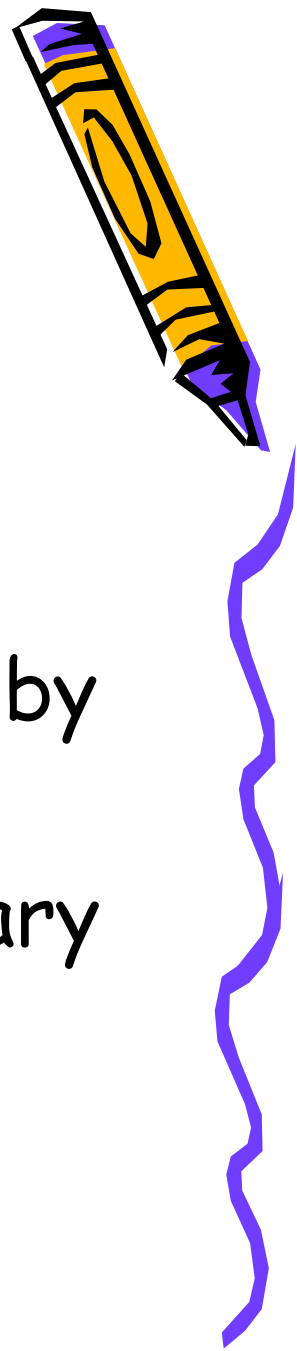<w type="verb" lemma="speak">

spake

</w>

<w type="preposition" lemma="with">

 <expan>w<ex>i</ex>t<ex>h</ex></expan>

</w>

Abbreviation and Expansion

```
<w type="verb" lemma="speak">
<nonabbr>spake</nonabbr>
</w>
<w type="preposition" lemma="with">
 <abbr>w<am>i</am>t<am>h</am></abbr>
</w>
```

Abbreviation and Expansion

```
<w type="verb" lemma="speak">
  spake
</w>
<w type="preposition" lemma="with">
  <g ref="#CLnc">w</g><ex>i</ex>
  <g ref="#t2">t</g><ex>h</ex>
</w>
```

HandShift

- We need the name of the sender
- and the name of the amanuensis.
- One letter was not always written by only one amanuensis.
- The amanuensis element is necessary even in the middle of the letter.

HandShift

<handShift/> marks the beginning of a sequence of text written in a new hand, or the beginning of a scribal stint.
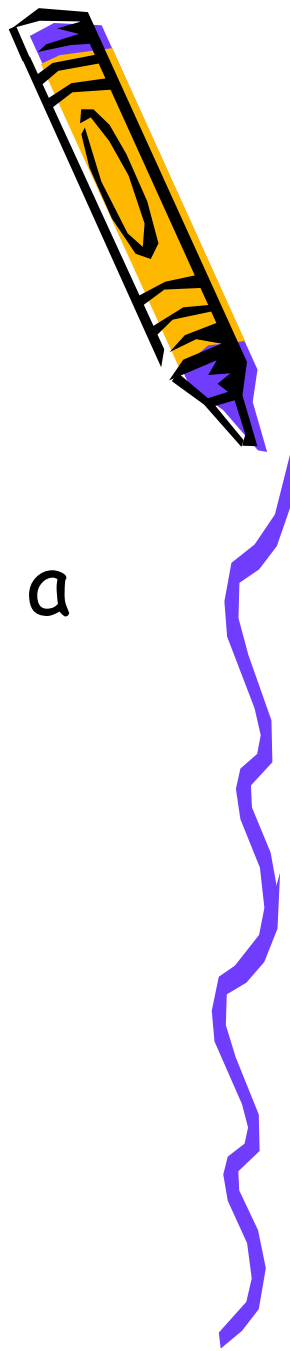
HandShift

<sender name="Margaret">

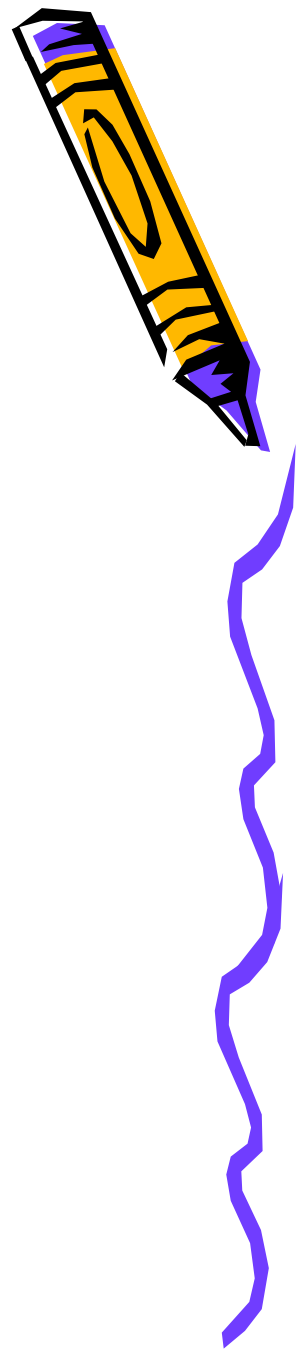 <amanuensis name="Gresham">

    …….

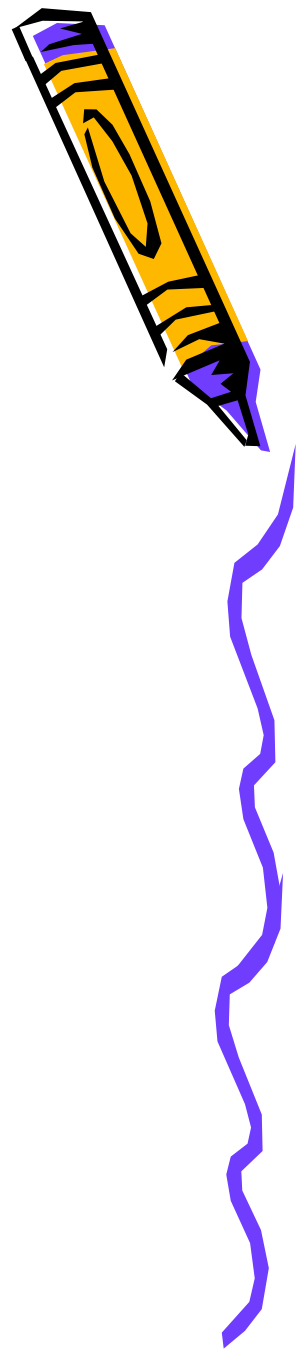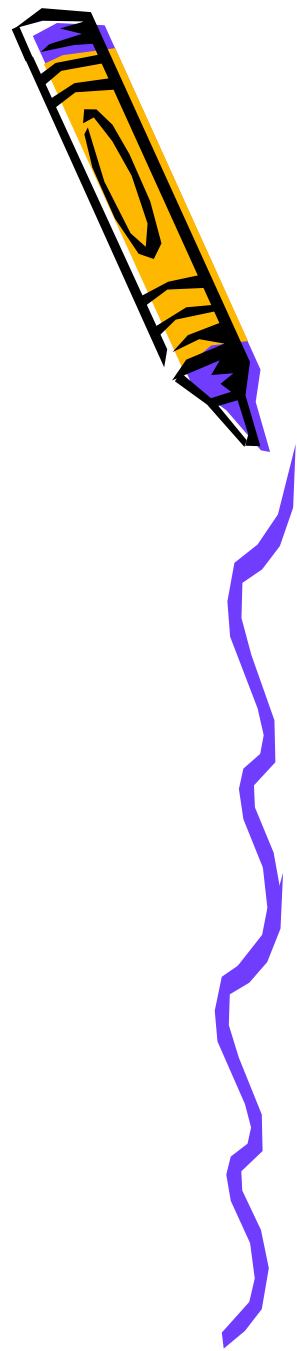   …….

    …….

  </amanuensis>

</sender>

P5-MS

- Ideally, a tagset for manuscript description should be able to accommodate the full range of eventualities.

(Driscoll 2006)

P5-MS

- Ideally, a tagset for manuscript description should be able to accommodate the full range of eventualities.

(Driscoll 2006)

P5-MS

- Ideally, a tagset for manuscript description should be able to be flexible enough to cope with every kind of objective.

- The Objective justifies the process.