

**Growth in Grammar Project
Transcription Manual
<STAGE ONE/>**

Mark Brenchley

Phil Durrant

Contents

Preparation	3
General Principles	4
General Practice	5
Our Recommendations.....	6
The "Mark Up" Language.....	7
The Conventions.....	9
Generic Conventions and Tags	11
Anonymisations.....	14
Formatting.....	16
Graphical Material	18
Punctuation & Other Marks	19
Word Level Changes	21

Preparation

Ethical Responsibilities

Before transcribing any project texts, you must have first read, understood, and signed the *transcriber agreement* form. You may not transcribe any texts until you have done so.

Whilst there is little risk of harm coming to the study's participants, our texts have been obtained solely on the basis of the participants' voluntary, informed consent. As one of our transcribers, you have an ethical duty to respect this consent, and the agreement form outlines your specific responsibilities here.

Materials

In general, we will provide you with all the materials necessary for transcribing our texts. This will include project specific laptops, memory sticks, and notebooks.

Queries

Whilst we aim to have provided you with all the information necessary to transcribe our texts, this information may not always be fully sufficient. Generally, such cases will be minor, and you will simply be asked to insert a specific code into your transcriptions; we can then identify these cases later and make a final decision ourselves.

However, if you have any substantial doubts regarding any part of the transcription process, please don't hesitate to contact us. You should also not hesitate to contact us if you have any doubts regarding the ethical dimensions of the research, especially when it comes to participant anonymity/confidentiality.

General Principles

Throughout the transcription process, your goal is to produce an *authentic* record of what the student has written. In general, this means that you should seek to transcribe the words that appear on the page exactly as they appear on the page. In practice, however, there are certain exceptions to this goal. In particular, you will need to pay particular attention to the following types of exception.

- (1) **Anonymisations.** This is material which requires anonymisation in order to ensure the confidentiality of the student authors, as well as anyone directly associated with either themselves or the school.
- (2) **Word Level Changes.** This is material where the student has written words in a way that would generally be considered an “error” according to the formal conventions of written Standard British English.

In all of the above cases, you will need to “amend” the text by replacing, removing or inserting material at the appropriate points during the transcription process.

Note that the word “amend” has been explicitly placed in inverted commas. This is to emphasize that your overall goal is still to create an *authentic* record of what the student wrote. In other words, “amendments” should only be made in accordance with the conventions set out in this manual. If the manual specifies no such amendment, then you should not alter the text and simply transcribe the material as the student has written it.

General Practice

Both during and after the transcription process, you must respect the anonymity and confidentiality of the students and their participating schools, as well as uphold the transcriber agreement signed at the outset.

In particular, you should not leave any of the original writing lying around or transcribe them in such a way that anyone other than a fellow transcriber or member of the project team can see them. Moreover, you should not discuss any of the confidential content of the texts with anyone other than a fellow transcriber or member of the project team.

For each text, you must:

- (1) **TRANSCRIBE** the text using *Microsoft Word*.

Make sure that the automated correction tools are switched OFF, but that the spellcheck and grammar highlighting tools are switched ON. You should also make sure the language setting has been switched to ENGLISH (UK).

NEVER add your own markings to the original pieces of writing. This is because we need to preserve the integrity of the original texts for the duration of the project. We will supply you with a notebook in which to make your own notes and comments, and we will collect this notebook back in once the transcription process is complete.

- (2) **TRANSCRIBE** the text using the conventions specified in this manual, *and only in this manual*.
- (3) **SAVE** the transcribed version as a .txt document, using the following format:

batch#_text#_your_name.txt

For example,

005_345a_Mark_Brenchley.txt

014_87_Philip_Durrant.txt

The batch and text number will appear as a highlighted single group of numbers on the text itself. You should record this exactly as it appears on the original piece of writing.

When you save to .txt, make sure the encoding is set to UTF-8.

Each transcription should always and only be saved directly to the encrypted memory stick which we will provide you. At no point, should a transcription be saved to your computer's desktop or some other external drive. Again, this is part of your commitment to respecting the consent of our participants.

- (4) **REVIEW** your transcription to make sure that it tallies with both the original text and the transcription conventions set out in the following section.

Our Recommendations

Ultimately, it is up to you what approach you adopt for transcribing and reviewing the texts, and we trust you to make the decisions you feel right. However, based on our own experience, we recommend the following strategies.

Creating Transcriptions

- (1) Take the first paragraph within that text and transcribe the original wording of that paragraph, *without* anonymising any material and *without* adding any material.
- (2) Check that the wording of your transcribed paragraph matches that of the original (e.g. make sure that any words are spelled as they are in the original).
- (3) Go through the transcribed paragraph and anonymise the material as appropriate.
- (4) Go through the paragraph, and see whether there is any other material that requires special treatment according to the other conventions specified in this manual (e.g. spelling errors, graphical material, special characters).

Reviewing Transcriptions

- (1) Run the *Spelling and Grammar* checker over your transcription to see if it can help you detect any errors. Remember that this checker is only a guide: it is up to you to make the final decision as to whether anything actually needs to be amended.
- (2) Check that there are no extra spaces that have accidentally crept into the transcript during the insertion of any tags. For example:

correct	<tag>amazing</tag>
incorrect	<tag> amazing </tag>
incorrect	<tag >amazing</tag >
incorrect	<tag>amazing< /tag>

Remember that a special macro has been provided that will help you do this!

- (3) Check that that the special codes “tally”. Specifically, in the case of tags that are inserted as pairs, your transcription should have equal numbers of each pair:

correct	<tag>amazing</tag>
incorrect	<tag>amazing<tag>

- (4) Check that the number in the file name matches that of the original text.

The "Mark Up" Language

The remainder of the manual details the specific conventions that you must follow when transcribing the texts, including any special material which you may need to add. This material may seem daunting at first, but it shouldn't take too long to get used to.

The most important thing to get to grips with is the mark up language. This "language" is simply the set of tags that you will add to a text, such as "`<sp >...</sp>`". Such tags are critical for maintaining an authentic record of the original texts. They are also critical for helping us analyse the resulting transcriptions in the most effective way possible. For example, in the case of spelling errors, they ensure that we can include both the original spellings and the amended versions, and then automatically filter out which version we analyse.

More importantly, the basic format of the language is very straightforward. In fact, there are only three considerations to bear in mind:

1) The majority of tags come in pairs. These are inserted *around* the material to which they apply. The first member of the pair tells us where the relevant bit of material begins, and the second tells us where it ends:

So I thought that was the most `<u>`amazing`</u>` thing I had ever seen.

As you can also see, there should be no space between the section of the text that begins the relevant bit of material and the section that ends the relevant bit of material. You should also preserve the internal spacing of the material. Thus, the following would be correct:

So I thought that was `<u>`the most amazing thing`</u>` I had ever seen.

But the following would not:

So I thought that was `<u>` the most amazing thing`</u>` I had ever seen.

So I thought that was `<u>`themostramazingthing`</u>` I had ever seen.

2) You may find that exactly the same piece of text requires two distinct tags. Here, there are two things to remember.

Firstly, in such cases, the order in which you insert each pair is not critical. What is critical is that the *relative* ordering of the second member of each pair mirrors that of the first. The best way to think of this is in terms of going up-and-down a set of stairs. To do this, you take stair one, then stair two, then stair three; when you come back down, you take these stairs in reverse order. It is exactly the same with our tags. Thus, both of the following would be absolutely fine for our purposes:

✓ `<caps>`amazing`</caps>`

✓ `<caps>`amazing`</caps>`

But the following would not:

✗ `<caps>`amazing`</caps>`

✗ `<caps>`amazing`</caps>`

Secondly, where the tags are being used to “correct” certain kinds of text, so that the original text is placed in the first member of the tag pair, you must remember to put the original text in each of the tag pairs. For example, suppose you needed to correct the spelling of a word and change the capitalization of this word. Then either of the following would be fine:

- ✓ `<sp AMAZZING><caps AMAZZING>amazing</caps></sp>`
- ✓ `<caps AMAZZING><sp AMAZZING>amazing</sp></caps>`

3) The final thing to keep in mind is that the tags come in three formats. These are as follows:

Format One

Here, the mark up is a *pair* of tags, where the first member of the pair consists of the relevant tag surrounded by angled brackets (e.g. ``). The second member repeats this notation exactly, but adds a slash (e.g. ``). For example:

`amazing`

Format Two

Format two is essentially the same as format one. Here, however, an additional bit of information is added after the tag in first brackets. This information is simply the original piece of text that you have just corrected. Critically, there must be space between the code that identifies the tag and the original piece of text that you have inserted inside the first bracket. For example:

`<sp their>there</sp>`

Format Three

Format three applies in cases where the mark up is not a pair of tags, but a single tag. Here, you simply insert the first member of what would otherwise be a pair, followed by a slash. For example:

`<symbol/>`

And that’s basically it. Once you have grasped these key features of the mark up language, it is simply a matter of getting the hang of which particular tags you need to use and when you need to use them. All of this information is detailed overleaf.

The Conventions

Remember: Your overall goal is to produce a transcription that is as *authentic* a record of the original text as possible. To this end, you should transcribe the content of each text as it appears on the page, in the order in which it appears on the page, proceeding word-by-word, line-by-line, paragraph-by-paragraph. There are only two exceptions to this general rule.

Exception One: This concerns blocks of writing arranged in a non-linear way. Here, you should simply transcribe this material as if it had been linearly, using the content of the text as your primary guide. For example, you would transcribe this portion of text:

First, take a cauldron and fill it with boiling hot blood.

Second, add three frogs and two toads born under a blue moon.

Third, cook the potion for three suns and sprinkle with children's toes until it gurgles.

and this portion of text:

First, take a cauldron and fill it with boiling hot blood.

Third, cook the potion for three suns and sprinkle with children's toes until it gurgles.

Second, add three frogs and two toads born under a blue moon.

identically:

First, take a cauldron and fill it with boiling hot blood.

Second, add three frogs and two toads born under a blue moon.

Third, cook the potion for three suns and sprinkle with children's toes until it gurgles.

Where the content of the text offers no clear guide to the ordering, then you should simply transcribe that particular portion of text on a left-to-right basis. For example, you would transcribe the following portion of text:

Take a cauldron and fill it with boiling hot blood.

Add three frogs and two toads born under a blue moon.

Cook the potion for three suns and sprinkle with children's toes until it gurgles.

as

Take a cauldron and fill it with boiling hot blood.

Cook the potion for three suns and sprinkle with children's toes until it gurgles.

Add three frogs and two toads born under a blue moon.

Exception Two: This concerns special kinds of material that you will need to replace, remove, or insert. In order to make this process more efficient and reliable, we have set up special macros. As such, your job here is basically to do two things.

Firstly, decide which macro to select. Having made this decision, simply highlight the relevant piece of material, whether this material is a single word or a whole group of words, and click the relevant macro button.

Secondly, if you have selected a macro for a pair of tags that requires you to "correct" a portion of text whilst adding the original material to the first member of the pair, then you will need to do so manually. For example, suppose you came across the following piece of material:

Their is no hope.

Step one would be to highlight the relevant word and then click the *Spelling* macro. This gives you:

<sp >Their</sp> is no hope.

Step two is then to correct the word manually and add the original piece of text to the first member of the pair. This gives you:

<sp Their>There</sp> is no hope.

Generic Conventions and Tags

Be Efficient!

Because our resources are finite, we ask that you do not spend huge amounts of time making a decision. As a rough guide, you should not spend more than 1 minute on any particular decision; at which point you should simply make an educated guess as to the best way to transcribe that particular piece of material. Of course, we understand that it will take you a bit of time to build up your transcribing experience and confidence, so we will give you plenty of time before the 1-minute rule "begins".

Content Only!

Remember that we are only interested in the *content* of the text. For example, you should not transcribe page numbers, the dates of the text or any grade that might have been assigned to the text. We will keep a separate record of this information, and it is not something you should worry about. You should, however, make sure to transcribe any headings, and mark these using the `<head>...</head>` tag.

Corrections

As these are student texts, you may find that some (maybe all!) bits of a text are incorrect, usually because the student has used the wrong form of a word (e.g. they wrote "begin" instead of "began", where it is the latter form that is correct). Unless otherwise specified, you should simply transcribe the material that is there – do not attempt to correct this kind of material by, for example, adding your words. If in doubt, ask!

The Benefit of the Doubt

If you can credit the student with having written a portion of text correctly, then transcribe it as having been written correctly. In other words, only "correct" a piece of text when you are confident it has been written incorrectly. For example, if a word is hard to read but looks as if it is probably spelled right, then you can just assume that it has been spelled right.

Ignore Crossed Out Text

We are only interested in material the student intended to be part of the content of their text. Hence, you should generally ignore any material that they have crossed out. Occasionally, however, such material may be a word or sequence of words that leaves the sentence incomplete (i.e. as if they crossed out a bit, meant to add something in its place, but forgot). In such cases, transcribe the whole sentence, whilst leaving out the crossed out elements, and then bracket the whole sentence with the `<garble>...</garble>` tag.

Ignore Student/Teacher Commentary

We are only interested in the material that forms the actual content of the text itself. As such, you should not transcribe any comments made on the text by either the student or their teacher. Simply treat this material as if it wasn't there.

Note that we also count as "teacher commentary" instances where the teacher themselves has corrected the students' spelling/punctuation, or if they have added extra words to a sentence in order to improve it! You should ignore all such teacher corrections and simply transcribe the material as originally written by the students. That is, by using the conventions for the `<sp>...</sp>` tag as described in the *Corrections* section below.

Missing Material

Unfortunately, because of the nature of our texts, you may come across material where parts of a word or parts of a sentence are missing. To mark this material, you should use one of two tags.

The first is the **<guess>...</guess>** tag. Use this tag where part of a word is missing but where you feel confident enough to guess what that word actually is. For example,

He is leav now --> He is <guess leav>leaving</guess> now

The second is the **<garble>...</garble>** tag. Use this tag where a sentence is incomplete because a word or sequence of words is missing, placing the first member of the tag at the beginning of the sentence and the second wherever the sentence "finishes". For example,

He wanted to [...] a book. --> <garble>He wanted to a book.</garble>

He wanted to [...]. So I left. --> <garble>He wanted to.</garble> So I left.

Uncertain/Unclear Material

Occasionally, you will come across portions of text about which you feel *very* uncertain because the text itself is hard to read. In such cases, you have two options.

Firstly, we ask that you transcribe this material as best you can, and then bracket the material using the **<guess>...</guess>** tag. Crucially, this tag should go around the whole of the material, including any tags which you may have just applied and about which you are uncertain. For example, where you are uncertain about a word but believe it to have been misspelled, you would transcribe as follows:

I am <guess><sp beet>beat</sp> </guess>.

Secondly, if you feel completely unable to identify any part of the material at all, then you should simply insert the single **<illegible/>** tag in its place, and move on. For example,

I am <illegible/>.

Quotations

Material which you have identified as a direct quotation should be bracketed using the special **<quot>...</quot>** tag. Note that this bracketing should only be applied to the actual quote itself. For example,

Juliet says <quot>"O Romeo Romeo, wherefore art thou Romeo?"</quot>

Material for Our Attention

Occasionally, you may wish to mark a specific portion of our text as requiring our attention. To do so, simply bracket this material using the **<comment>...</comment>** tag. As suggested by the tag itself, the tag is set up so that you can insert your own commentary inside the first member of the pair. For example,

<comment I think this is written by the student, but I am not sure>portion of text for our attention</comment>

GENERIC TAGS & CONVENTIONS				
MARK UP	GUIDANCE	MACRO NAME	EXAMPLES	
			ORIGINAL	TRANSCRIPTION
<comment> ... </comment>	Marks material as requiring our attention – You can insert your own comment into the first member of the tag pair here	Comment	n/a	<comment This word may have been inserted by the teacher>portion of text for our attention</comment> <comment Text is unfinished></comment>
<garble> ... </garble>	Marks a sentence where a word or sequence of words is missing or the material is ungrammatical	Garble	I wanted to a book.	<garble>I wanted to a book.</garble>
<guess> ... </guess>	Marks an instance where the material is hard to read or a part of a word is missing but you feel confident enough to guess it	Guess	He is leav now.	He is <guess>leaving</guess> now.
<illegible/>	Replaces material which you cannot transcribe at all because the text itself is impossible to read	Illegible	n/a	<illegible/>
<quot> ... </quot>	Marks an instance of quoted material	Quotation	Martin Luther King said "I have a dream"	Martin Luther King said <quot>"I have a dream"</quot>

Anonymisations

Anonymisation requires *replacing* all pieces of text which might *directly* identify either the student author or anyone directly connected with themselves or the school.

Most such instances will be readily identifiable: they will usually be the names of individuals, locations, and institutions which would not be known to the general public. In other words, they will usually be the names of people or things with which you yourself are unfamiliar. As such, you would not generally anonymise names where these refer to public figures or public places that do not in themselves directly identify the writer. For example, do not anonymise

Leonardo diCaprio is an amazing actor

London is the world's most important city

but do anonymise

My brother Leonardo DiCaprio is an amazing actor

I live at 23 Marple Street.

However, you may find instances which are not so easy to identify. In which case, please use the `<anon>...</anon>` tag. You should also use this tag for any extended portion of text that you feel might identify/compromise the student or anyone directly connected with them or the school. This is unlikely, but not impossible. For example, you might find a passage where the student tells a highly personal story which in itself might directly identify them. In such instances, please bracket the entire portion of material using the same `<anon>...</anon>` tag. Note that we do not expect you to spend a huge amount of time attempting to identify instances of the latter! We simply ask that you flag instances that strike you when transcribing so that we can make a final decision ourselves.

Anonymisation Conventions

- 1) Replace all names that might identify either the participant or anyone directly connected with them (e.g. friends, family) or their school (e.g. the school's name, staff members). You should do so using one of the three MARK UPs in the table below.
- 2) In the case of narrative text, anonymise a person's name only if (a) it includes a surname or (b) it matches the student's name at the top of the text.
- 3) Remember to replace the name *as a whole*, even if the name includes several words and even if some of these component words are not themselves names. For example,

Mark	-->	Person_name
Mark Brenchley	-->	Person_name
Exeter School	-->	Institution_name
The School of Exeter	-->	Institution_name
St John the Baptist Academy	-->	Institution_name
Heavitree Road	-->	Location_name
Prince of Wales Road	-->	Location_name
22 Eccles Street	-->	Location_name

- 4) If the original material has been apostrophized, you should also apostrophize the anonymization. For example,

Mark's manual	-->	Person_name's manual
Exeter School's success	-->	Institution_name's success

ANONYMISATIONS				
MARK UP	GUIDANCE	MACRO NAME	EXAMPLES	
			ORIGINAL	TRANSCRIPTION
Person_name	Replaces person's name	Anonymise Person	Mark Brenchley	Person_name
Location_name	Replaces name of specific location	Anonymise Location	Baring Court	Location_name
Institution_name	Replaces name of specific institution organisation	Anonymise Institution	The School of Exeter	Institution_name
<anon> ... </anon>	Marks a piece of text that you think may require anonymisation but about which you are unsure	Uncertain Anonymisation	n/a	<anon>text containing uncertain case of anonymisation</anon>

Formatting

As noted, your goal is to transcribe each text on a word-for-word basis, doing so in the order in which these words appear on the page and as they appear on the page. In line with this general goal, you will also need to capture certain aspects of each text's formatting, *but only these aspects*. There are two broad kinds of formatting to be transcribed, one relating to the "font" of the text and one relating to the "type" of text included within the original.

Font

Here, you will simply need to insert the various tags as appropriate and as listed in the table below. In all cases, this is simply a straightforward matter of bracketing the relevant material using the relevant tag.

Note, however, that the font tags should only be applied to the stretch of text that has been specifically formatted in one of these ways. Thus, the following material, for example

I think that is great.

becomes

<u>I think that is great</u>.

But the following material

I think that is great.

becomes

<u>I</u> <u>think</u> <u>that</u> <u>is</u> <u>great</u>.

FONT				
MARK UP	GUIDANCE	MACRO NAME	EXAMPLES	
			ORIGINAL	TRANSCRIPTION
 ... 	Marks bolded text	Bold	go now	go now
<i> ... </i>	Marks italicised text	Italics	go <i>now</i>	go <i>now</i>
<u> ... </u>	Marks underlined text	Underline	go <u>now</u>	go <u>now</u>
_{ ... }	Marks subscripted text	Subscript	H ₂ O	H₂O
^{ ... }	Marks superscripted text	Superscript	$x^2 = z^2$	x² = z²

Text Type

The “type” of text basically refers to two aspects of the texts you are transcribing:

- 1) how the text is arranged on the page (e.g. as a line or a paragraph)
- 2) whether it constitutes a “bullet point”, a “footnote”, a “heading”, or a “title”

Generally, the information required here is fairly straightforward, and is set out in the table below. The exception here concerns the presence of footnotes, which require two separate tags. Firstly, use the `<fnr>...</fnr>` tag to bracket the number that marks the footnote reference:

I have discussed this earlier`<fnr>1</fnr>`

Secondly, use the `<footnote>...</footnote>` tag to bracket the content of the whole footnote, remembering to transcribe this footnote at the end of the transcription (i.e. as an endnote). For this part of the footnote, you do not need to include the number that marks it; you simply need to make sure the footnotes appear at the end of the transcript in the same order that they appear in the original text, with the first footnote coming first, then the second, and so on:

`<footnote>See Chomsky (2012).</footnote>`

`<footnote>This is also something discussed by Chomsky (2012).</footnote>`

TEXT TYPE				
TEXT TYPE	MARK UP	MACRO NAME	GUIDANCE	TRANSCRIPTION EXAMPLES
Blank lines	n/a	n/a	No tag required - simply leave a blank line as appropriate	n/a
Paragraph break	n/a	n/a	No tag required - simply start a new paragraph as you would normally	n/a
Headings and Titles	<code><head></code> ... <code></head></code>	Head	Marks any kind of heading/title within a text	<code><head></code> full text of heading <code></head></code>
Bullet points	<code><bullet></code> ... <code></bullet></code>	Bullet	Marks a bullet point	<code><bullet></code> full text of bullet point <code></bullet></code>
Footnote reference	<code><fnr></code> ... <code></fnr></code>	Fnr	Marks the number that references a footnote	Many people have noted this <code><fnr>1</fnr></code>
Footnote proper	<code><footnote></code> ... <code></footnote></code>	Footnote	Use this tag to mark the actual footnote itself	<code><footnote></code> full text of footnote <code></footnote></code>

Graphical Material

Graphical material constitutes an exception to the rule that you should transcribe the material in the text as it originally appears and in the order in which it originally appears. Instead, you should transcribe each bit of graphical material as if it had been written immediately under its closest paragraph and on a separate line, using the appropriate tag from the table below. (In some cases, of course, graphical material will actually appear directly above or below a paragraph; in which case, this material should simply be transcribed as is.)

There are four types of distinct graphical material, which you should mark as such. These are listed in the table below. Note that that some of this material contains text and some does not. In those cases which do contain text, we ask that you transcribe all of this text within the relevant tag pair, doing so as if it were a single paragraph, and doing so in the default left-to-right order noted earlier.

GRAPHICAL MATERIAL				
MARK UP	MACRO NAME	GUIDANCE	XAMPLES	
			ORIGINAL	TRANSCRIPTION
<graphic> ... </graphic>	Graphic	Marks any graphical material that contains words but is not a table (e.g. a graph or diagram)	n/a	<graphic>full text of material included within the graphic </graphic>
<table> ... </table>	Table	Marks any material recorded in a table (i.e. rows & columns)	n/a	<table>full text of material included within the table </table>
<image/>	Image	Marks any graphical material which contains no text at all (e.g. an unlabelled illustration)	n/a	<image/>
<symbol/>	Symbol	Replaces any graphical material that appears within a paragraph and is used instead of a word or set of words	I would give it ★★★	I would give it <symbol/>

Punctuation & Other Marks

All punctuation and other orthographic marks should generally be transcribed as they appear in the original texts. You should not attempt to alter or correct such marks, *even where they are clearly wrong!!* For example:

.	remains	.
,	remains	,
-	remains	-
%	remains	%
£	remains	£
(remains	(
]	remains]

There are, however, five important exceptions to this general rule:

1) You should correct any cases where a punctuation mark has incorrectly been used to mark a particular word (e.g. where an apostrophe has been used but shouldn't). You should do so using the `<sp>...</sp>` tag from the next section in this manual. In effect, what this means is that we treat such mistakes as a spelling error. This is not ideal, but it is something we need to do for the purposes of our later analysis.

2) All quotation marks must be transcribed using the double quotation marks provided by *Microsoft Word*. This is to make sure that our analytical software can reliably distinguish such marks from the possessive apostrophe. Thus, for example, you would transcribe *all* of the examples below as follows:

I said "I love you"	-->	I said "I love you"
I said 'I love you'	-->	I said "I love you"
That book is "great"	-->	That book is "great"
That book is 'great'	-->	That book is "great"

3) You may find cases where an asterisk (*) or other mark has been used to mark text which appears later but which is meant to be read at this point. In such cases, you should ignore this mark and simply insert the related piece of text as if it actually appeared at that point in the text.

4) All fractions should be transcribed using the following format:

numerator/denominator

leaving a space between the fraction and any associated whole number (where present). Thus, for example, you would transcribe the following fractions as follows:

$\frac{1}{5}$	-->	1/3
$\frac{4}{5}$	-->	4/5
$1\frac{3}{4}$	-->	1 3/4
$21\frac{1}{2}$	-->	21 1/2

5) Finally, there are three specific replacement symbols that you must use. These are required as the original symbols interfere with the software that we will eventually use to analyse the transcriptions. These symbols should be transcribed as follows, using the relevant macros:

<	-->	<
>	-->	>
&	-->	&

Word Level Changes

There are three types of word level "changes" which you will need to make, primarily in order to bring the texts in line with the orthographical conventions of Standard British English. Most importantly, in each case, although we are asking you to "amend" the text, you must also maintain a record of what was amended, using tag *format two*. For example,

john --> <caps john>John</caps >
That's so rong --> that's so <sp rong>wrong</sp>

The three types of "correction" are as follows:

1) *Abbreviations* - Pieces of text where a word or sequence of words have not been spelled out in full. For example,

See q.1 --> See <abbr q.1>question</abbr> 1
C u soon --> <abbr C>See</abbr> <abbr u>you</abbr> soon

Note, however, that the following *do not* count as abbreviations, and should be transcribed as they appear in the original text:

Standard Metrics (for example, *cm, mm, km, ml*)
Standard Abbreviations (for example, *i.e., NB., e.g.*)
Acronyms (for example, *RSPCA, UoE, BAE*)

2) *Capitalisations* - Pieces of text where a student has (a) used a capital letter incorrectly, (b) forgotten to use a capital letter where there should be one, OR (c) used capitals to emphasize a word/sequence of words. For example,

I LOVE you --> I <caps LOVE>love</caps> you
I love You --> I love <caps You>you</caps >
I love yOu --> I love <caps yOu>you</caps >
i love you --> <caps i>I</caps > love you
I love tim --> I love <caps tim>Tim</caps>

3) *Spelling Errors* - Pieces of text which have not been spelled according to the conventions of standard British English. For example,

That's so rong --> that's so <sp rong>wrong</sp>
That's alotof errors --> That's <sp alotof>a lot of</sp> errors
He's mis-taken --> He's <sp mis-taken>mistaken</sp>
He's mis taken --> He's <sp mis taken>mistaken</sp>

Note from the examples above that we *do* count the following as spelling "errors":

"Exaggerated" spellings: sooooo --> <sp sooooo>so</sp>
"Split" words: with out --> <sp with out>without</sp>
"Merged" words: alotof --> <sp alotof>a lot of</sp>

Note, however, that the following *do not* count as spelling errors, and should be transcribed as is:

Contractions (e.g. I'm, she's, they're, don't, can't, must've, would've)

Compounds with/without hyphens (e.g. scary sounding, scary-sounding)

Please also note that you should also be very wary of oddly spelled words in fictional narratives, especially if these refer to special creatures or actions - these may well be deliberate, and you should give them the benefit of the doubt here.

Finally, if in doubt, all spelling errors should be checked using the English dictionary component of Cambridge Dictionaries Online, making sure to check the spelling against the British version of the word! The dictionary itself can be found here:

<http://dictionary.cambridge.org/dictionary/english/>

WORD LEVEL CHANGES				
MARK UP	GUIDANCE	MACRO NAME	EXAMPLES	
			ORIGINAL	TRANSCRIPTION
<abbr original> ... </abbr>	Use this to expand and record non-standard abbreviations	Abbreviation	Q1. A. two	<abbr Q1.>Question 1</abbr> <abbr A. two>Answer two</abbr>
<caps original> ... </caps >	Use this to make and record any capitalisation changes	Caps	john He's GREAT!	<caps john>John</caps> He's <caps GREAT>great</caps>!
<sp original> ... </sp>	Use this to make and record any spelling changes as licensed by this manual	Spelling	Luvv it! mis-take in stead of its terrible Trumps plan	<sp Luvv>Love</sp> it! <sp mis-take>mistake</sp> <sp in stead>instead</sp> of <sp its>it's</sp> terrible <sp Trumps>Trump's</sp> plan