

TRANSCRIPTION OF EARLY LETTER FORMS IN RARE MATERIALS CATALOGING

Prepared for the DCRM Conference, 10-13 March 2003

Deborah J. Leslie and Benjamin Griffin

5 March 2003 16:57

© The authors 2003

I.

General Comments on Transcription

It may surprise some readers to be facing such a hefty document, given that there are some 20 years behind us of standardized transcription practice for early printed materials. In that time there have naturally arisen some questions from the field about how to apply certain provisions (such as whether **q̄**/**q̄** should be transcribed as **[que]** or **q[ue]**), or confusion resulting from general rather than detailed directions (such as how to separate component letters of an uppercase ligature). More compelling, though, is the continuing controversy over the ISBD(A) / AACR2 / DCRB rules for transcription of **i/j u/v** and **vv/w**. This issue has erupted unexpectedly at a public hearing for the proposed DCRB Core, over online discussion groups, practically “wherever two or three catalogers are gathered together.” The predicament, really, is that there is no clearly superior solution to the problem; our task is finding the least painful solution. Hence this study, which attempts to consider transcription of rare materials in a general, overarching way, and then use those principles and insights to clarify our thinking about the really prickly problems, and perhaps to render the decisions less vulnerable to continuing question.

Then, too, it satisfies pleas that whatever is decided, discussion on the issue be recorded. The earlier generation of catalogers and catalog-rule-writers who decided on the current method of following the printer’s practice in the text when converting case for **i/j u/v vv/w** did not leave a record of their deliberations. We cannot guarantee that the next generation of catalogers and catalog-rule-writers will not also want to debate the issue anew, but we will at least have left them our deliberations (at least, the deliberations of the authors of this paper, who quite frankly duked out the issue between themselves).

Appendix 1 contains excerpts of pertinent rules regarding transcription from several standard documents: *Anglo-American Cataloguing Rules* (2nd ed.) (AACR2), the *Library of Congress Rule Interpretations*, *International Standard Bibliographic Description for Older Monographic Publications* (*Antiquarian*) (ISBD(A)), and *Descriptive Cataloging of Rare Books* (DCRB), the document under revision. These various documents give largely the same instructions for transcription; the differences that do occur are worth considering as well. Essentially, all require the preservation of word order and spelling while normalizing capitalization, diacritics, punctuation, and letter

forms where possible. AACR2 orders the separation of ligatures and digraphs, while the LCRI and DCRB make certain exceptions. AACR2 chapter 2's section on older monographic publications prohibits conversion of case from lower to upper, as does ISBD(A), although the LCRI and DCRB provide for complete normalization of case in both directions. There are some slight differences regarding the transcription of *i/j u/v vv/w*, but all require adherence to the upper- and lowercase usage in the text (to the extent possible) when converting case, with the exception of the LCRI to 1.0E, which then proceeds to eliminate itself for our purposes by referring users to DCRB for the bibliographic treatment of items published before 1801.

Before plunging ourselves fully into cataloging rules, let us consider the nature and development of early letter forms and spellings.

II.

Early Letter Forms and Orthography

Many new catalogers of early printed materials, that is, materials of the fifteenth through eighteenth centuries are daunted by the unfamiliar symbols and spellings they find on title pages; the earlier the printing, the more archaic the look and the more daunting the task of transcription. Even experienced rare materials catalogers may find themselves at a loss in dealing with unfamiliar abbreviations or contractions. Letter forms and spelling conventions have both changed over time, and the intention in DCRM to provide a relatively faithful representation of the title page requires it to provide clear instructions for transcription of pre-modern letter forms and orthography.

We can identify three general types of differences between early printed texts and modern ones that must be taken into account by the rare book cataloger: 1) graphical differences, 2) orthographic changes, and 3) contractions, ligatures, and digraphs.

1. Graphical differences: letter forms and punctuation.

Certain letter forms have changed over time, but the general trend has been toward fewer special forms. Obsolete letter forms nevertheless generally have unambiguous modern equivalents. Indeed, letter forms we consider archaic coexist side-by-side with modern forms of the same letter in the same word.

Let us take, for example, *s* and *f*, which are different shapes, or “graphs” (definition of the graph: what it looks like) of the letter *s*, convention governing which form was used where. An *f* was rarely used at the end of a word, for example; that was the place for *s*. But there was and is no confusion about their natures—they are two graphs of the letter *s*.

Conventions governing the use of punctuation marks were evolving during the hand-press period until they finally settled down into general regularity in the 19th century.¹ Before the 17th century especially, notions on the use of commas, semi-colons, and colons were subject to a certain amount of individual preference. A virgule, preferred by early black letter printers over the comma, looks like a slash or vertical line and is equivalent to a comma. The hyphen as a single straight line was in regular use by the 12th century in manuscripts, but some early printers, notably Gutenberg, preferred a double-stroke hyphen (not an equal sign) for word-breaks.

The primary problem here is that it is not a matter of archaic forms of punctuation, but rather of inconsistent conventions for the various symbols—we have both a straight-line hyphen then as now, but they also had a double-stroke hyphen. The modern equal sign is equivalent in form to the double-stroke hyphen, but entirely different in function; modern convention for separating words is by single-stroke hyphen. Although punctuation is normalized in transcription, we must have instructions for transcribing pre-modern punctuation uses for the occasions when punctuation is an integral part of a word.

The graphical differences and changes associated with the letters **i/j**, **u/v** and **vv/w** pose particular difficulties, and will be considered separately in the next section.

2. Orthographic changes.

Spelling of certain words in all European languages has changed over time. And of course, spelling in general before the later 17th century was not so standardized as it is now. **Vertue**, **justifiez**, **zwey** are common examples of changed orthography in English, French, and German respectively.

Since original spelling is to be preserved in transcription, pre-modern orthography in itself does not present a transcription problem. It is only when the cataloger is faced with pre-modern use of **i/j**, **u/v** and **vv/w** that it becomes necessary to sort out whether something, say the lowercase form of the word “DIVERS,” provides an instance of a graphical difference or an orthographic one, and the answer depends on when and where you ask it. But more on that later. For the time being, while we’re talking about orthography alone, it is enough to say that regarding orthographic changes, the primary task for a cataloger upon encountering an oddly-spelled word is to recognize whether it is a valid spelling of the time or whether it is a typographical error. The latter is to be noted or corrected, most often using “[sic],” while the former is transcribed without comment. A secondary difficulty is the brain’s tendency to see what it expects, tempting the cataloger to unconsciously supply modern spelling in place of archaic.

3. Contractions, ligatures, and digraphs

¹ I have relied on A.C. Partridge, *Orthography in Shakespeare and Elizabethan Drama*, London, 1964, Appendix VIII, “The historical development of punctuation marks” for this discussion of punctuation marks.

Contractions. Contractions and abbreviations are largely a feature of Latin, Greek, and Church Slavic books of the 15th and 16th centuries, continuing the custom of manuscript contraction. DCRB instructs the cataloger to expand the contractions, with the supplied letters within square brackets (0J2). Although it can be a difficult task for a cataloger unfamiliar with the language, contractional expansions do not generally provoke substantial difficulty. Lists of common contractions and other symbols, such as that found in McKerrow's *Introduction to Bibliography*², are generally sufficient as a guide to the cataloger in expanding contractions. Some problems recur, however, for which it would be beneficial to give specific guidance.

Ligatures and digraphs. A ligature is a single type-body containing two or more letters. Its transcription is an easy proposition: transcribe it into its component letters. Digraphs, on the other hand, are two letters written as one and considered to be a single letter and/or sound; sometimes also called diphthongs. This distinction between ligature and digraph is not made in cataloging rules, and both are called "ligatures." The difficulty for transcription here are the rules instructing us to separate ligatures into their component letters, even though the common ligatures are available in the character set, and the fact that the instruction in the Library of Congress' rule interpretations to AACR2, repeated in DCRB, instruct us to make exceptions by retaining ligatures in certain circumstances. For example, when faced with a title page with "Ædipus," the transcription as "Ædipus" or "Oedipus" or possibly "OEdipus" would depend on whether the title page was in French or English.

III. I/J, U/V, and VV/W

Without going far into the territory of historical linguistics, we may begin with the fact that European orthography presents us with a problem of terminology as regards "letters." We may ask, of any given period, "Were **i/j** **u/v** considered different letters?"; but what does it mean for them to be "different letters"? A 16th-century Englishman would agree that **u** (for example) was "a letter," but would deny that **v** was "a different letter". He would also not be able to say what sound the letter **u** makes without reference to its immediate context. The mere fact of the **w**'s English name "double u" is only the most obvious relic in our language of the earlier usage, when the name of a free-standing **v** was pronounced "you" and the **vv** as "double you".

To the modern mind, **u** and **v** (and **i** and **j**, henceforward, unless otherwise specified) are different letters because they represent different sounds. Since our discussion must consider both the archaic and the modern usage, it will be all right sometimes to refer to "letters," and sometimes necessary to refer to "graphs". The **u** graph is pronounced differently in different periods and contexts, but it remains the **u** graph throughout. For obvious reasons, these points

² R.B. McKerrow, *An introduction to bibliography for literary students*, New Castle, DE, 1994, p. p. 319-324.

are more easily grasped in writing than they are in speech. The determination of whether at any given time we are dealing with **u** and **v** as different graphs of the same letter or two different letters determines whether we are dealing with orthographic differences (which must be preserved in transcription) or graphical differences (which are not).

The best brief historical overview of the use of **u/v** can still be found in the work of R.B. McKerrow.

McKerrow helpfully provides a field guide to "the practice of the earliest printers" -- that is to say, printers of black-letter incunables:

1. There was an upper-case letter approximating in shape in Gothic fonts rather to the modern **J** than to **I** [3], but serving indifferently for modern **I** and **J**.
2. An upper-case letter approximating in shape in Gothic fonts to **U** [U], and serving indifferently for modern **U** and **V**.
3. A lower-case **i**, serving for both modern **i** and **j**.
4. A lower-case **j**, used for the second of two **i**'s in words like "perij," and in Roman numerals as "viiij."
5. A lower-case **u**, serving for both modern **u** and **v**, but only used medially or finally.
6. A lower-case **v**, serving for both modern **u** and **v**, but only used initially.³

The situation is much the same for printing in the later Roman fonts, except that in the Roman capitals, letters looking more like **I** and **V** replace the Gothic types looking like **J** and **U**; their function is the same as the Gothic types (that is, each form serves indifferently as consonant or as vowel, and its use is dependent only on position). McKerrow notes that in Italian (Roman-font) books, lowercase **u** is used in the initial position by several printers.

In this earliest printing, therefore, differences between **u** and **v** were simply graphical; they were different shapes for the same letter whose use depended on the position in the word, with some variation by convention. These graphical differences are analogous to the differences between **f** and **s**.

If the **u/v** problem were only that of graphical differences, we would have a 23-letter alphabet and our way as catalogers would be easy. Instead, the period between the cradle of printing and the mid-17th century witnessed the gradual differentiation between **u** and **v** as representing different sounds, in particular, **u** (and **i**) as vowels, and **v** (and **j**) as consonants. And if

³ R.B. McKerrow, "Some notes on the letters i, j, u and v in sixteenth-century printing," *Library*, 3rd ser., no. 1 (1910), 239-259, p. 25-26.

representing different sounds, they represent different letters—graphical differences transforming into orthographical differences (and in a most disorderly fashion). Our alphabet has 26 letters, and our way as catalogers is complicated.

We are indebted to McKerrow for his investigation of the historical movement through which the **u/v** forms came to stand for a phonetic difference. He found an example of the phonetic (modern) usage from a Dutch book of 1492, thus dating from the incunable period. The modern system was "the usual thing" on the Continent by about 1620.

In England from about 1570 there were apparent attempts to use the modern distribution, which would only become standard around 1630. But admixtures of the two "systems" and ad-hoc irregularities persist until a surprisingly late date; and commonly at that period there are differences, within the one book, between lowercase and uppercase usage. As late as 1660, one can find a printer using **v** initially at all times – but alternating between initial **i** and **j**. (Wing J541A)

As for the letter **w** and its analogues, scant attention has been paid compared to that given **i/j** and **u/v**. McKerrow states without making a point of it that "In early founts this is often represented by **vv**. In later times the same is often found in founts of extra large size ... and in ordinary founts when there happened to be a run on **w** and the compositor had not enough." ⁴ (p. 312). The implication by McKerrow, borne out by our extensive experience with pre-1801 English books, is that from the very beginning **vv** has represented **w**, and thus can be considered a graphical variant rather than an orthographical one. We are not suggesting necessarily that a **vv** should be transcribed as **w**, but we do want to question how **VV** is to be transcribed, whether it is retained in capitals or converted to lowercase.

In printing there are three ways of forming **w** and its analogues: **w** as a single type-body looking perfectly like a modern **w**; **vv** as two distinct type-bodies placed adjacent to each other with normal kerning; and a hybrid: two type-bodies with one of the pieces (usually the left) filed or rubbed away allowing the two pieces to sit more closely to each other. For the last, the intention seems clearly to form a **w**; one assumes that the font being used wanted enough **w** types, as McKerrow wrote. We assert that this is much more the case with uppercase types. That is, one is much likelier to find lowercase **w** as a single type-body than one finds an uppercase **W**, and we suggest that the majority of cases in which there is a mixture of distinct **VV** and filed-down **VV** with **W** in one piece shows a corresponding distribution of uppercase for the former two and lowercase for the last.

There is not much historical progress regarding these representations of **w**. In our examination of early title pages using examples to hand of early uses of **u/v**, we find a 1513 work⁵ using **w** consistently (all lowercase), and a 1648 work⁶ using a filed **VV** and a distinct **VV** (both

⁴ R.B. McKerrow, *An introduction to bibliography for literary students*, New Castle, DE, 1994.

⁵ *The hystorye, sege and dystruccyon of Troye* (STC 2nd ed. 5579)

⁶ *The full truth of the Welsh-affaires* (ESTC R204754)

uppercase). We believe all this is evidence that **vv** and **w** have always been equivalent, unlike the gradual transformation of **u/v** from graphical to phonetic differences. The consequences of such an understanding will be discussed below, with our recommendations.⁷

IV.

Modern Treatments and Rationales

Let us look briefly at the transcription practices of other standard bibliographies and catalogs, partly to see whether other bibliographical treatments might provide a model, but more importantly because they may provide the citation from which researchers search our own databases.

We first turn to the eminent pre-AACR enumerative bibliography for English printed works: Pollard and Redgrave's *A short-title catalogue of books printed in England, Scotland, & Ireland and of English books printed abroad, 1475-1640*. The transcription policies of both the original edition of 1926 and the revised edition of 1976-1991 engage in a fair amount of arbitrary regularization, which, as the second edition confesses, is an unfortunate aspect of this otherwise model bibliography:

In books up to 1600 the intention has been in every instance ... to print lower-case *j* as *i* in all positions, lower-case *v* for both *u* and *v* when used initially, and lower-case *u* for both *u* and *v* when used medially. This practice has been followed, as it was by the printers themselves, in transliterating upper-case *V* to the two lower-case letters After 1600 the modern practice as to *i* and *j*, *u* and *v*, has generally been followed, despite the slowness with which it gained ground, as the earlier use could not be assumed without verification in each case.

To proceed as if practice had changed unilaterally at a certain date was based on expediency — the last clause quoted above shows this, hinting that STC's complete change to modern practice after a certain date is a necessity brought about by the materials at hand (presumably the information-gathering stage had gone forward for some time without any reasoned approach to the letter-forms).

Additionally, Pollard explained that "in proper names **I** when followed by a vowel has usually been printed as **J**." We do not know which is more alarming, the rule itself, or that "usually";

⁷ Notes: Search of Hamnet "uu" retrieved 31 records, in which all but one were Italian or Latin forms of the word of "uulgo". The exception was German "uund". In all cases at least one of the **u**'s was needed to function as a vowel. Conclusion: no examples in European languages of **uu** representing **vv**. In the Folger ESTC database, 10 records were retrieved. Aside from the Dutch "uut," the remaining English words were all typographical errors for **un**, as in "uutruths." Conclusion: no English words with **uu**. Folger ESTC **vv** = 2951 entries. All English words with **vv** show it used as a consonant corresponding with modern w. We do find examples of English Latin printing of Reginald Wolfe's in the accusive as "Vuolfium."

likewise with “generally” in paragraph above. In any event, STC's compromise is neither consistent, historically accurate, nor representative either of the title pages nor the practice DCRB inaccurately calls the “practice of the printer.” (The “usage of the text” would be a more accurate way of stating it). STC does not provide us with any models for transcription.

Donald Wing's bibliography, taking up where Pollard & Redgrave left off, declined to follow the STC's idiosyncratic transcription practices. The first edition of *Short-title catalogue of books printed in England, Scotland, Ireland, Wales, and British America, and of English books printed in other countries, 1641-1700* fails to mention the transcription practices followed. The 1993 revision of volume 1 includes this not entirely satisfactory explanation:

Seventeenth-century spelling and punctuation have been retained: “vv” for “w,” “v” for “u,” “j” for “i” whenever they appear, although some inconsistencies may result from differing conventions by reporting libraries. (We have tried to see originals or photocopies of as many title pages as possible.) Capitals have been lowercased, except for proper nouns and adjectives.

But as the preface gives no word on the important point of *how* capitals have been lowercased, one does not know if the example of old-spelling title such as J137 in which an uppercase I is transcribed into a lower-case i instead of j is a mistake or a deliberate and implicit declaration of the lowercasing principle.⁸

Therefore, the Wing bibliography provides us with no better model. The struggles of STC and Wing are instructive in themselves, however; this is a difficult issue, the practice used in cataloging should be reasonably clear and consistent, and our solutions need to take into account these differing and unpredictable transcription practices of sources of bibliographical citations.

The English Short-Title Catalogue (ESTC) is the heir to both printed bibliographies, and includes the 18th century as well. Copies of all STC and Wing entries are in the process of examination by ESTC staff and are edited according to the ESTC transcription rules, which follow ISBD(A) exactly. That is, the usage in the publication being described is followed when converting from upper- to lowercase. Conversion from lower- to uppercase is not permitted; this marks the major divergence of AACR2 and DCRB from ISBD.

Still, the extent of the problem is perhaps less great than it appears. Current (that is, ISBD and codes based on it) cataloging theory and practice broadly agree that the spelling of title page information is to be reproduced as exactly as possible (that is, original orthography is to be preserved); so that there is no question of, for example, actually modernizing the spelling of old books (as is done, for example, with the texts of student editions of Shakespeare, or as is done in

⁸ KING IAMES His Iudgement of a KING ... is transcribed in the Wing entry as: King James his iudgement of a king.

creating standardized titles for use in creating thesauri). So transcription problems are limited to a few small (but vexing) areas.

V. Transcription Systems

We make reference to the working group 1 paper on [General Principles by Joe Springer](#) in evaluating the four transcription described below.

user convenience: decisions taken in the making of descriptions should be made with the user in mind.

representation: descriptions should be based on the way an information entity describes itself. A subprinciple is

accuracy: descriptions should faithfully portray the entity described.

principle of sufficiency and necessity: descriptions should be sufficient to achieve stated objectives and should not include elements not required for this purpose.⁹

principle of standardization: descriptions should be standardized, to the extent and level possible

In evaluating how well each system satisfies the standardization principle, we are considering the whole catalog as the context rather than how a single method can be standard in reference only to itself.

When catalogers and others discuss transcription issues, one recommendation sure to be voiced is “Transcribe it like it is.” What these speakers *probably* mean is what we call a “converted graphical transcription,” that is, when converting from upper- to lowercase, the graphical form of the capital is assigned to the lowercase regardless of the lowercase practice for that letter in context. Let us be clear right away that the only way a title page can be “transcribed as it is” is to use quasi-facsimile transcription. This form of transcription, standardized by Fredson Bowers in 1949 in his *Principles of bibliographical description*, distinguishes full capitals, small capitals and lowercase letters, type faces, and line endings of the original, along with rules, devices, ornaments and the like.

We do not believe quasi-facsimile transcription is a viable option, so let us give our reasons and be done with it. This type of transcription is taxing, laborious, and time-consuming, and is exponentially more vulnerable to error. The time required for proof-reading alone would be prohibitive. Our character set does not distinguish between large and small capitals, nor between Roman, italic, or black letter. And most persuasively, quasi-facsimile transcription offers the user nearly nothing as compared to a simplified transcription. Quasi-facsimile transcription, while garnering high marks for representation and accuracy (assuming that the cataloger has in fact produced an accurate transcription), fails miserably regarding user

⁹ The principle of “parsimony” is considered for our purposes to be included in the principle of sufficiency and necessity.

convenience (it is hard to read), sufficiency and necessity (it offers much more data than is necessary for the task), and standardization (no current and few former cataloging practices involve quasi-facsimile transcription).

Fredson Bowers sets forth in his *Principles* a simplified transcription style as well. I echo his assertion that the purpose of simplified transcription is to preserve the content, although not necessarily the form, of the title page.¹⁰ Simplified transcription preserves the content by retaining in transcription the printed text “exactly as to wording, order, and spelling, but not necessarily as to punctuation and capitalization.” (AACR2)

With reference to the DCRM General Principles, the use of simplified transcription rather than quasi-facsimile transcription for cataloging purposes satisfies the *principle of user convenience* (it is easier for users to read simplified transcription than quasi-facsimile), the *principle of representation* (by retaining with some faithfulness how the item represents itself, although less so than with quasi-facsimile), the *principle of sufficiency and necessity* (simplified transcription is generally sufficient for identification), and the *principle of standardization* (by modernizing letter forms and by taking already-established practice in cataloging and bibliographical documents as precedent).

Before we proceed any further, let us discuss what it means, when converting uppercase **i/j u/v** to lower, to “adhere to the pattern of uppercase/lowercase employed by the particular printer” (DCRB) or to follow “the usage in the publication being described” (ISBD(A)) or “convert to lowercase according to the usage of the text”. Appendix B in DCRB is quite helpful, and specifies that “the practice of the individual printer can usually be discovered by an observation of the internal text (preferably in the same type as the title page) of the publication.” The parenthetical statement is important; when converting Roman uppercase to lower, it is Roman lowercase usage in the text, not Italic, that provides the proper evidence for the printer’s practice. A printer may have different practices regarding the contentious letters between different typefaces. It is in a few cases unavoidable resorting to another type, such as when a title page is completely in uppercase Roman, but the text is entirely in Italic. Nevertheless, the basic principle should be Roman for Roman, Italic for Italic, and black letter for black letter. This is the first point I would like to emphasize.

The second is that none of the tables giving guidance when the practice cannot be discerned in AACR2, DCRB, or ISBD(A) make any mention of noticing whether the graph in question is vocalic or consonantal. And yet, that is one of the essential points to note and follow, especially if one is trying to discover whether or not pre-modern graphs or orthography is involved.¹¹ That is, use of a lowercase medial vocalic **u** does not provide evidence for determining the use of a

¹⁰ Bowers, F. *Principles of bibliographic description*. New Castle, DE, 1994, p. 138.

¹¹ The inability to tell at first glance is not as odd as it may seem. We are looking at a 1636 English title page reading “THE NAVIGATOR.” One might transcribe this as “The navigator” without a second thought or glance, but the printer gives lowercase medial consonants the **u** graph, which means that the title is to be transcribed “The nauigator.”

lowercase medial consonantal **u**. Granted, the illustration given in DCRB Appendix B of the transcription of IVS VERVM does make mention of a pattern of vocalic and consonantal use, but nowhere is it given as part of the task of discovering the use of the text.

The third point I would like to make is that frequently, the title page itself provides all or much of the needed evidence, even given the additional burden of distinguishing between use in different typefaces, and vowels from consonants. Appendix 2 contains a survey with discussion of several sets of early works with **i/j u/v** conversion needs. A cataloger need not necessarily consult the interior text of the book unless the evidence is not apparent from the title page. One exception is the lowercase textual use of **II**. Regardless of an implicit assumption in the conversion tables that a text would use either **ii** or **ij**, it is not unusual for a printer to use one of the forms medially and the other finally. For example, the text might use Venetijs, but Tullii, or vice versa. As we say in the survey discussion, this is a particularly difficult task because although it is quite usual for title pages to use the genitive and ablative forms of names (lots of **II** combinations), it may be difficult to find the corresponding lowercase textual use. The Folger continental cataloger says that uncovering the correct conversion of **II** is easily the most time-consuming part of transcription for him.

Let us now consider several different forms of simplified transcription that have been proposed. The first two involve complete normalization, but in different ways. One is to normalize according to contemporary norms; another, to normalize according to modern norms.

1. Normalization to modern usage

Follow the transcription method of the LCRI (see Appendix 1) or one similar. The cataloger transcribes **u** for vowels, **v** for consonants, and **w** for consonantal **vv** or **uu**.¹² No attempt is made to regularize **i** or **j**. A strong argument can be made for doing this even when *not* converting from upper- to lowercase, if one claims that since we are dealing with graphical differences (**u/v**) of the same letter, we convert them to the modern form of the letter. On the practical level, mere consistency would require that this method be applied throughout, in fact, regardless of case; otherwise the result would be an historical mishmash.

This transcription method assumes that all differences between pre-modern and modern graphs of the **i/j u/v vv/w** are simply graphical changes, and in so doing must ignore the legitimate shift of graphical into orthographical change that these letters make as they are transformed into vowels and consonants.

Advantages: Easy for catalogers to perform, easy for users to read, this would be the most predictable transcription method. High marks for user convenience. Moderate marks for standardization; this is the LCRI conversion table, and is presumably the method used for American non-DCRB rare book cataloging.

¹² The authors have been unable to find any instances of consonantal **uu** in texts, and would be grateful to have any brought to their attention.

Disadvantages: Ignoring legitimate orthographic changes is not only historically inaccurate but violates the cataloging principle of preserving original orthography. Practically, there would be one of two ways of determining what the modern use would be: 1) relying on how the word is pronounced today, or 2) verifying the modern spelling in a dictionary. Relying on how words are pronounced to guide the transcription may prove taxing if the cataloger does not have speaking ability in the language or there are different ways of pronouncing the word. The most prominent example is the word LIEVTENANT, which would be transcribed as either “lieutenant” in a system that intends to modernize letter forms, upon whether the V is construed as a vowel (as Americans speak the word today), or “lievtenant,” if construed a consonant (as do the British, who nowadays spell it “lieutenant” and pronounce it **leftenant**). Verifying from a dictionary would require recourse to an external source, the dictionary, thereby reducing its predictability slightly. However, we do not want to overstate this disadvantage. In our experience, most people can fairly quickly determine with good accuracy whether a medial **u** has vocalic or consonantal value, even in unfamiliar languages.

This transcription type would follow the principle of user convenience, or we should say “some users’ convenience,” depending on what use the user was making of the catalog entry. Modern normalized transcription fails in representation, accuracy, sufficiency and necessity. There is an important place for this type of normalization, however—in uniform titles.

2. Contemporary tabular normalization

Tabular normalization is based on the recognition that although there were a number of combinations of patterns a printer could use with **i/j u/v**, there was nevertheless overall a usual practice. The tables given in McKerrow quoted above, in AACR2 and DCRB are based on this understanding. Instead of identifying the particular use of the text, the cataloger would convert from upper- to lowercase solely by use of the “table of last resort.” This transcription method allows for the complicated nature of graphical and orthographical changes of **i/j u/v**, but does so in a general way rather than a way specific to the text.

Advantages: The transcription would go very quickly, and most of the time it would in fact turn out the same as transcribing based on textual use, if our survey sample is typical. (See Appendix 2). A similar transcription method is used by the standard printed English-language national bibliographies. It satisfies to a moderate degree the principles of user convenience, representation, accuracy, and sufficiency and necessity. In two ways it satisfies to a low-moderate degree the principle of standardization: 1) we are aware that some agencies apply this as a “table of **first** resort;” and 2) if the results we obtained in a survey of actual usage in books recorded of Appendix 2 can be duplicated by other institutions, tabular transcription may result in the same or nearly the same transcription as that based

on the text. But “nearly the same” is not the same, and it is overall a pretty rickety kind of standardization.

Disadvantages: There are two common patterns used by continental printers not accounted for in the table: initial **V/u** and medial or final **II/ij**. The conversion of **VV** in uppercase to **uu** in lowercase (**vv** in initial position) is an historical mistake. We cannot find evidence that **uu** was ever used to represent a **w**. The closest we have come is the name Wolff, printed in Latin imprints as **Vuolffium**. These are problems with this particular table. If one were to amend the table so that it would somehow be more responsive to actual contemporary printing patterns, this method will still not provide an accurate transcription, and will result in some strangely inconsistent spellings if transcription includes lowercase text of the pertinent letters showing a different pattern than the table’s. For example, **novvm** and **nouum** might appear together if the first occurrence of that word were in all capitals, and the second in lowercase.

In order for this method to achieve even a rough level of historical accuracy, it must not be applied to books, whatever the date, that uses a modern distribution. Therefore, the cataloger would need to make sure of the usage of the text. Granted, it may be a simple matter to ascertain that a particular text adhered to pre-modern graphs and proceed to use the table. As we point out earlier, however, the transformation into orthographical distinctions of the letter pairs did not always happen at the same time, so a text may show a modern use of **u/v** and an archaic use of **i/j**.

Question: if we were to recommend this method, should we consider separate tables for continental and English books?

3. Converted graphical transcription

When converting from upper- to lowercase, the cataloger takes the graphical form of the uppercase letter and assigns it to the lowercase for purposes of transcription. The polar opposite of normalized modern transcription, this method does not recognize that differences in **u/v** letterforms can be merely graphical, but assumes that all differences are orthographic.

Advantages: It is the fastest and easiest method for the cataloger to employ. No decisions have to be made, the text does not need to be consulted. The cataloger need not even be concerned about whether the text uses modern distribution or not. High marks for representation and accuracy; moderate marks for sufficiency and necessity.

Disadvantages: The result is sometimes incoherent; by ignoring the reality of graphical differences, this transcription results in forms and spellings one would never see in lowercase in the original. Few if any current operations use this transcription method,

meaning that the citation in hand is unlikely to match the transcription in the bibliographic entry. It is difficult to read. Low marks for user convenience; low marks for standardization.

4. Textual use

This is the method currently specified by DCRB, AACR2, and ISBD(A). It requires the cataloger to identify lowercase patterns used by the printer in the particular text being cataloged, and to use those patterns in transcribing from upper- to lowercase. The textual transcription method recognizes the differences between graphical differences and orthographic differences, and take them into account in a precise way for each book.

Advantages: This method provides the truest and most precise historical and linguistic representation of the title page. It is the current standard for transcription of early letter forms; its use in AACR2, DCRB, and ISBD(A) ensures that it is not only the standard for the Anglo-American world, but for the western world in general, which in turn predicts a high level of correspondence between the transcription form and the form searched by the users. It has high marks for the principles of user convenience, representation, accuracy, sufficiency and necessity, and standardization.

Disadvantages: It is the most taxing for the cataloger to apply. One can find examples (if one looks hard enough) of inconsistent patterns within one book. One must have the book in hand in order to provide an accurate transcription, rendering it unpredictable to users who do not.

VI. Recommendations

1. Punctuation

1.1. We recommend normalizing and modernizing all punctuation marks that must be transcribed. We give greater weight to the instruction to completely normalize and modernize punctuation within a transcription than we do to the instruction to transcribe letters as they appear. (Text of 0H: In general transcribe letters as they appear. Convert earlier forms of letters and diacritical marks, however, to their modern form). Indeed, since we are not considering letters at all here but punctuation marks, that instruction seems to have little relevance. In recommending this, we are considering punctuation marks primarily on the basis of their function and secondarily on their form. Although it is not a matter of transcribing archaic symbols according to their form, it is a matter of transcribing symbols according to their modern function with reference to their form. In one important respect the rendering of punctuation symbols is different from that of letters. For punctuation symbols, the underlying

coding is not significant since punctuation is disregarded in indexing. With the exception of the original punctuation option in DCRB, punctuation within the transcribed text is completely normalized and modernized.

Transcribe

Esq; Esq:	as	Esq.
Flete=strete	as	Flete-strete
M. ^r	as	Mr.
Exc ^{mo}	as	Excmo.
Excmo	as	Excmo

Move mention of the virgule from 0E to Appendix B.

1.2. Dates.

- 1.2.1. Periods with dates. We recommend the complete modernization of periods used in conjunction with both Arabic and roman numerals. They have no abbreviation function and their use is irregular.

Transcribe

January 10. 1641. till ...	as	January 10, 1641 ... or January 10 1641 ...
29. Nouembris. 1591.	as	29 Nouembris 1591
Decemb. 19. 1697.	as	Decemb. 19 1697
die xxvi. Iulij.	as	die xxvi Iulij
5th. time	as	5th time

- 1.2.2. Roman dates. They may appear as a succession of letters without spaces, with periods after the major portions of the date, with spaces, or a combination of all three. Example 13 in *Examples to Accompany DCRB* renders as M.D.LXXVIII a date appearing as **M. D. LXXVIII**. That is, it retains the periods but not the spaces.

We recommend that all spaces and periods be removed from the transcription of roman dates.

Regarding the periods, it is purely a matter of modernizing periods that appear with dates as argued in recommendation 1.2.1. Regarding the spaces, we argue that the number as a whole functions as a word, and as such should not be separated by internal spaces. The use of spaces in a machine environment has a consequence extending beyond the attempt to imitate the form of what we are transcribing. Spaces indicate a break, the marking off of a lexical unit, and are the basis for separating the units (words) in indexing. Our cataloging practice already implicitly combines what has been separated and separates what has been combined. For the former, see Example 5, in which a single word occupying

four lines without hyphens are silently combined into one: **Or | tho | gra | phia** is transcribed as **Orthographia**. Conversely, the barely perceptible quarter-space between **'t** and **vernuff** in Example 20 is regularized with a normal space; should the correct form be **'tvernuff**, the cataloger would silently close up the space.¹³ We recommend extending the same treatment to roman numerals.

- 1.2.2.1. Apostrophus.¹⁴ This is the use of an alternate form of an apostrophe, a curved line looking much like an inverted **C**, to represent large numbers. Said to originate from the ancient Roman practice of representing 1000 with **∞**, developed from the Chalcidic phi which is shown as a plain circle with a vertical diameter (not protruding beyond the circle). The sign for 500 is thought to be (the right hand) half of this; the cross bar was added to distinguish it from the letter **D**. It was only after Roman times that the letter **M**, as an abbreviation of mille/milia, came to stand for 1000. The earliest use discovered of an apostrophus in an imprint is 1566¹⁵; apparently its use by printers reveals deliberate attempts to achieve an archaic look. An example of the year MDLXXXVII in apostrophic form is **CD ID LXXXVII**.¹⁶

Regardless of the different origins of **M** and **D** and their equivalent apostrophic forms used in imprints, we nevertheless argue in favor of their transcription as **M** and **D**. The MARC character set does not accommodate an upside-down **C**, and the use of parentheses is an unsatisfactory substitute. A note may be made that the Roman date contains apostrophic forms.

- 1.3. Dashes. The use of a double-stroke hyphen was discussed above in 1.1. We concern ourselves here with the transcription of hyphens or dashes used in titles to elide letters or whole words, usually names. Typically at least the first letter of the elided word is printed followed by dashes in various configurations. Sometimes there will be one long line, sometimes a succession of distinct hyphens, standing in for one letter each. There might be attempts to make a single line by a number of dashes, which may have been

¹³ This is admittedly a weak example, but is the only one I could find from the *Examples* illustrating the point. Cf. L'hercolano / Benedetto Varchi. In *Fiorenza : Nella stamperia Filippo Giunti e Fratelli, MDLXX*. Twice on the t.p. the conjunction **e** is preceded by a comma with no space, and it is rightly transcribed with the space inserted: ... **della Toscana,e della Fiorentina** ... as ... della Toscana, e della Fiorentina ...

¹⁴ Our thanks to Brian Hillyard for bringing the relevant information in these sources to our attention: *Oxford Classical Dictionary*, s.v. "Roman Numerals"; B.L. Ullman, *Ancient Writing and Its Influence* (1963); and R.A. Sayce, "Compositorial Practices and the Localization of Printed Books, 1530-1800", *The Library*, 5th series, 21 (1966), 1-45 (re-issued Oxford Bib Soc): p. 38-41.

¹⁵ Found by Brian Hillyard in John Durkan, *Bibliography of George Buchanan* (1994); 11 years earlier than the earliest imprint found by Sayce.

¹⁶ This image of an apostrophic date is from http://www2.inetdirect.net/~charta/Roman_numerals.html

more or less successful. There has existed no guidance up until now in transcribing dashes.

The ESTC transcribes one hyphen for each distinct piece of type. Thus, a word elided with a three-inch-long single rule will be transcribed with a single hyphen, while a small word but with a dozen small hyphens close together will have twelve hyphens. This seems to us to be the most satisfactory solution. Any other solution we can think of would sacrifice too much of the content. For the messy situations, where the printer tried to make a solid line but some of the dashes do not line up exactly, we can make no recommendation but that the cataloger try to distinguish as much as possible how many distinct dashes there are, and transcribe that number.

2. Capitalization

2.1. We propose considering the adoption of the ISBD(A) prohibition against converting from lower- to uppercase. Conversion to uppercase according to modern usage is much less common than the reverse. When it is, difficulties frequently arise with it. AACR2 Appendix A indicates that such a phrase as "the Right Honourable Lady Chatsford" should be capitalized, but as often as not, such a phrase appears on a title page something like "the right, pious, vertuous, and honourable Lady Chatsford." AACR2 does not cover this situation.

2.2. "At the sign of" Another condition of uncertainty is caused by agencies transcribing addresses and faced with something like "... at the sign of the Swan." We propose establishing this convention: keep in or convert the word "sign" to lowercase, but keep in uppercase any description of the sign's image. This corresponds with a majority of how the capitalization of like phrases appears on title pages; we take it to mean that the description functions something like a proper noun. If the ISBD(A) conversion rules are adopted, then any image description in lowercase would remain that way.

3. Ligatures

We agree that the separation of ligatures is appropriate for transcription, even though the MARC character includes them. The stated goal of conveying the content although not necessarily the form is well satisfied by letter separation. Nevertheless, consideration of eliminating the exceptions of these languages leads quickly to a clash between the *principle of sufficiency and necessity* (modernizing digraphs into their component letters is sufficient for identifying, finding, and selecting entities) and the *principle of standardization* (by taking that route we would be deviating from standard LC as well as DCRB practice).

One of basic revision principles is that if a DCRB rule deviates from AACR2, it should be retained in DCRM only if there is a rare book reason for the deviation. There is no specific rare book reason to insist on ligature separation, neither is there one to delete it. Since the

exception is from an LCRI and not from AACR2, we believe it may be either retained or eliminated without violating our revision principle. ISBD(A) encourages but does not require the retention of ligatures.

We recommend ignoring the LCRI and instructing catalogers to separate the component letters in ligatures and digraphs without exception.

4. Symbols

4.1. Tironian sign. The Gothic font’s “tironian sign” should be transcribed as an ampersand, since it is the black-letter analogue for what appears as an ampersand in roman type. Both are derived from MS. contractions of Latin “et”. (This was the decision arrived at by the Bibliographic Standards Committee at the ALA annual meeting in 1999).

4.2. DCRB Appendix 2 covers other examples of letters as well as punctuation, for example, a small e used in superscript over a vowel. We propose that DCRM(B) provide a bit more guidance, possibly including the contents of the following table along with the long s, d with bent bar, 2-shaped r, tironian sign, stops used with dates and as abbreviations:

Transcribe	q̄; ꝥ	as	q[ue]
Transcribe	q̄	as	[quam]
Transcribe	ß		ss ¹⁷
	ƒz	as	ss
	ƒz		sz

5. I/J U/V W

5.1. General Recommendation

We recommend continued use of method number 4, the textual use method. Aside from quasi-facsimile transcription, it is the method simultaneously most faithful both to the individual text and to the printed language.

5.2. The complaints against textual use transcription that we know of fall into three general categories: 1) it is in fact **not** faithful to the title page; 2) it is too laborious and time-consuming; and 3) transcription is unpredictable, because not only must someone have the book in hand in order to transcribe it, s/he must also look inside the book. We will attempt to answer these complaints.

¹⁷ C.G. Allen, *A manual of European languages for librarians*, London, 1999, p. 40.

- 5.2.1. The charge of title page infidelity is argued generally by those who favor the converted textual transcription method. We believe this argument is based on a flawed understanding of the history of these graphs and their relations to each other as letters. It also reveals a failure to understand that converting to upper- to lower-case is in fact a conversion of graphs. Converting **A** to lowercase results in **a**, for instance, not in **Α**. Use of the converted graphical transcription results in records that are usually difficult to read, and contains spellings that would not actually appear in the language, most particularly in reference to medial consonantal **V**. Only very rarely is medial consonantal **V** to be seen in lowercase as **v** in the pre-modern period; this is nevertheless the most common conversion performed. Combine this with the very common practice of a title beginning in capitals and switching to a normal upper/lowercase distribution results in some strange combinations. (One librarian was heard to call it 'anathema;' we will not go that far, but will say that we find it significantly disquieting). Cf. the novvm/nouum occurrence in the sample transcriptions. For fairness' sake, we must admit that one does find the printer using different lowercase graphs for the same letter in italic and roman on the title page. This inconsistency is unavoidable for all methods except perhaps the normalization to modern use method, but we must also add that it occurs much less frequently than the graph changes in roman alone.
- 5.2.2. Textual use transcription is indeed the most taxing for the cataloger to apply. However, we argue that the correct lowercase graph of a letter as used by the printer in the text (initial vs. medial vs. final; vocalic vs. consonantal; roman vs. italic) can a substantial portion of the time be gotten by looking no further than the title page itself. This has been our personal experience, which the surveys recorded in Appendix 2 confirm. Printer inconsistency is another charge. We find that the only area in which one might find considerable inconsistency within the same work is with **II**. We deal with this difficulty later.
- 5.2.3. Before thinking through these issues methodically, one of the authors brought the charge of transcriptive unpredictability against textual use conversion herself. She thought it ridiculous that one needed to be holding the item in hand in order to even guess what its transcription might be. This is true, but she no longer thinks it ridiculous. What she has come to understand is that such is true for any transcription method of an early title page besides quasi-facsimile. Those searching an item in a catalog usually do not have a title page facsimile in front of them, nor are they generally working from a citation employing quasi-facsimile. They are generally using a regularized transcription of some sort, either from a national bibliography (such as ESTC), an online database (such as OCLC), or the citation in a scholarly work. We do not believe transcriptive unpredictability to be a serious obstacle to applying the textual use method.

5.3. Additional considerations for textual use transcription.

5.3.1. The general understanding that the cataloger must consult the internal text of the work (an understanding entirely valid given the language describing this method) would need to be amended. The language of OH should probably be revised to explicitly encourage use of title page forms when appropriate. Perhaps through a note, or perhaps through Appendix B, we might even suggest that it would be appropriate to turn to the table of last resort after a reasonably short length of time. I am unpersuaded that any time over five minutes spent at this task is worth the effort, and that with each passing minute, the cost is higher. We might even suggest that a reasonable length of time may be as soon as one discovers that the necessary lowercase use is not available on the t.p. Too, we might suggest that the final **II** use in signatures is a good and quick way to establish that use.

5.3.2. On the other side, the rule should add an awareness of the vocalic or consonantal value of the letter to be converted. This appears to add to the burden of applying this method, and in fact may do so, especially for those unfamiliar with the language of the text. However, in our experience most people can do so relatively easily in most cases. Attentiveness to vocalic or consonantal value will also serve to dispel some confusion that now exists. Novice rare book catalogers who find both medial **v** and **u** may think they are witnessing printer inconsistency, when in fact they are merely witnessing different graphs used for vowels and consonants (that is, the modern distribution).

5.3.3. We have identified greater complexity in the lowercase use of **II** than expected, especially in 16th-century continental works. More evidence is needed, but it appears that some printers use different lowercase forms of **II** depending on case. For example, one printer might appear to be using **ii** for ablative and **ij** for genitive, while another does the opposite. Or perhaps the discrepancy can be accounted for not by case, but by position; a printer might use both **ii** or **ij**, the distribution of each depending on whether it is used in the medial or final position. For the time being, we recommend amending OH to recognize this fact. And since this practice appears to be entirely or almost entirely continental, we suggest considering whether we need two tables, one for English imprints and another for continental.

5.4. **W**. We have already written about some of the difficulties attending printing and transcription of this letter and its analogues. Regarding **W**, we do not generally recommend using the printer's practice in the text, because we can almost guarantee that it will be wildly inconsistent, and that mostly because of the scarcity of the **w** graph in all the various fonts and typefaces relative to the compositor's need of it.

- 5.4.1. A **W** in a single type body itself creates no problem in transcription. We recommend transcribing a filed **VV** as **W**, and two entirely separate unaltered characters with normal kerning as **VV**.
- 5.4.2. Changes to 0H table:
Transcribe into lowercase *consonantal* **VV** as **vv**. Note: words in capitals with vocalic and consonantal **V**'s standing next to each other, such as **SAVVAGE**, would follow the general provision in the table about converting **V** to **u**, applied to each letter separately. What we are concerned with here is to make sure that a **W** analogue printed as **VV** is converted into **vv**.

5.5. Uniform titles.

- 5.5.1. DCRB Appendix 1 specifies title added entries in order to account for the various transcription possibilities of pre-modern letter use. One does not find such title added entries as often as one should in online catalogs, but that is probably because so many records for early books were retrospectively converted from cards. We want to suggest a significantly greater use of uniform titles when dealing with pre-modern graphs and letter use. Specifically, that a uniform title formulated according to AACR2 be used whenever the transcription of the title proper involves pre-modern orthographic conventions.
- 5.5.2. The first purpose for use of a uniform title in AACR2 is providing the means “for bringing together all catalogue entries for a work when various manifestations (e.g., editions, translations) of it have appeared under various titles.” In this case, it is not necessarily various titles, but various *potential* titles. We believe this is an excellent and entirely appropriate use of the collocatory function of uniform titles. Prominent display of a uniform title with modernized orthography would do much toward lessening uncertainty for the user when identifying and selecting manifestations.
- 5.5.3. Use of uniform titles implies added work for the cataloger, a concern we would like to address. Since a 246 title added entry for modern orthography of pre-modern **IJUV** conventions is already prescribed, no additional intellectual work required in formulating such a title is required. There is a greater responsibility to verify that a uniform title is correct and appropriate, but since the cataloger should be verifying the existence or lack of a uniform title for any given manifestation being cataloged, we do not believe this is an onerous task, especially compared with the benefits.
- 5.5.4. Miscellaneous note on uniform titles, which we believe have been seriously underused in rare book cataloging. Not linked specifically with the use of uniform titles for **IJUV**, the issue of uniform titles for use with titles proper

beginning with something other than the chief title is nevertheless closely associated. Early works in Latin frequently begin with the author's name in the genitive form, followed eventually by the chief title. There is no firm requirement in AACR2 or DCRB to use a uniform title in such circumstances, although that is exactly how it is consistently done in *Examples to Accompany DCRB*. We have seen, however, that uniform titles are rarely used for this case in online catalogs. We would like to urge rare book catalogers to use uniform titles in their cataloging whenever the chief title is not the first element transcribed in the title proper, or whenever pre-modern conventions of orthography are part of the title transcription.

5.5.5. Final statement regarding title transcription and added entries. The combination of title transcription, uniform title, and other title added entries for pre-modern orthographic conventions of **IJUV** together use all the transcription methods: textual use transcription assisted by contemporary tabular normalization for transcribing the title; normalization to modern usage to construct the uniform title, and converted textual transcription for additional 246's. Since all of these transcription methods have something to recommend them, and none are perfect, and given the importance of collocation especially with pre-modern spelling, we are satisfied that this solution provides the best all-around means of transcription and access.

5.5.6. Example

240 1 Oeuvres morales de Plutarque

245 14 Les oeuvres morales de Plutarque / ꝛ translatees de grec en françois, reueues et corrigees en plusieurs passages par le translateur ; comprises en deux tomes, & enrichies en ceste derniere edition de prefaces generales, d'amples sommaires au commencement d'vn chacun des traittez, & d'annotations en marge, reueuës, & de nouveau augmentees de moitié, lesquelles monstrent l'artifice & la suite des discours de l'auteur ; avec vn indice des choses memorables mentionnees esdites oeuvres.

246 3 Oeuvres morales de Plvtarqve

260 [Geneva] : ꝛ De l'imprimerie de Iacob Stoer, ꝛ MDCXXI [1621]

Appendix 1

Current transcription rules

AACR2

1.0E. Language and script of the description

1.0E1.

In the following areas, give information transcribed from the item itself in the language and script (wherever practicable) in which it appears there:

Title and statement of responsibility

Edition

Publication, distribution, etc.

Series

1.1B. Title proper

1.1B1 Transcribe the title proper exactly as to wording, order, and spelling, but not necessarily as to punctuation and capitalization. Give accentuation and other diacritical marks that are present in the chief source of information (see also 1.0G). Capitalize according to Appendix 1.

2.14E. Transcription of certain letters

2.14E1. Transcribe capitals that are to be converted to lowercase according to the usage of the text. If the usage of the text is in doubt or if it is inconsistent, transcribe

I as i

J as j

U as u (but as v when it is the first letter of the word)

V as u (but as v when it is the first letter of the word)

VV as uu (but as vv when it is the first letter of the word)

Transcribe gothic capitals in the form of J and U as I and V.

N.B. there is no instruction for conversion from lower- to uppercase, which implies an assumption of coherence to ISBD(A)'s prohibition of same.

LCRI

1.0E. Language and script of the description

Pre-Modern Forms of Letters

In general, transcribe letters as they appear in the source. However, convert earlier forms of letters and earlier forms of diacritical marks into their modern form, as specified herein. Separate ligatures that are occasional stylistic usages (Ædipus, alumnæ,

etc.) rather than standard usages in the modern orthography of the language, e.g., œ in French (as in œuvre) or æ in Danish (as in særtryk). If there is any doubt as to the correct conversion of elements to modern forms, transcribe them from the source as exactly as possible. (See also the section on Special Letters, Diacritical Marks, and Punctuation Marks.)

The following represent a special case: u/v, uu or vv/w. When these letters are used in Latin and some other languages without regard to their vocalic or consonantal value, so that "u" is used for a "v," etc., the transcription should be regularized. This means that for the bibliographic description of items published after 1800,

- 1) use v for consonants, e.g., vox, Victoria;
- 2) use u for vowels, e.g., uva, Ursa Major;
- 3) use w for consonantal uu or vv, e.g., Windelia.

Follow this guide also for publications of any date when the case is not one of bibliographic description, e.g., headings or citations from reference works. The letters i/j should be handled differently. For the bibliographic descriptions of items published after 1800, transcribe "i" and "j" as they appear; do not attempt any regularization. Follow this stipulation also for uniform titles for series. For any other case of headings, citations from reference sources, etc.,

- 1) use j for consonants, e.g., jus, Julius;
- 2) use i for vowels, e.g., iter, Ilias.

N.B. For the transcription of any of these letters in bibliographic description for pre-1801 publications, apply Descriptive Cataloging of Rare Books (DCRB). For the use of uniform titles, so that DCRB titles file properly (i.e., together with the titles of post-1800 publications), see LCRI 25.2A.

Special Letters, Diacritical Marks, and Punctuation Marks

Use the double underscore (__) as the conventional means of signaling special letters (including superscript and subscript letters), diacritical marks, and punctuation marks for which there is no exact representation in the character set. ...

Exception 1: Do not use the double underscore convention in the following cases; use instead the equivalent indicated:

- Old German small "e" = umlaut (Fürsten)
- "Scharfes s" or "ess-zet" (ß) written as ligature = ss (Auslegung)
- "Scharfes s" or "ess-zet" (<j) written as two letters = sz (Auslegung)

ISBD(A)

0.6 Language and script of the description

Elements in areas 1, 2, 4 and 6 are normally transcribed from the publication and are, therefore, wherever practicable, in the language(s) and/or script(s) in which they appear there. Interpolations in these areas are enclosed in square brackets and are given in the language and/or script of the context of that part of the description ... The spelling of words taken from the publication is preserved, but ligatures and other contemporary forms of letters and diacritics may be transcribed in their current forms when the contemporary form is not available to the cataloguing agency. No account is taken of the differing forms of letters when no orthographic difference is made in current usage. Accents and other diacritic marks not present on the source are not added. For recording of upper and lower case letters, see 0.8. Interpolations by the cataloguing agency should follow modern practice as to spelling. Latin interpolations should follow the practice of the item in hand

0.7.6 Abbreviations found in the publication

When contractions and abbreviations in continuance of the manuscript tradition of abbreviating words which were long and/or frequently encountered are found, these may be left as they stand or may be expanded wherever possible. Any such expansion must be indicated by italicizing, underlining, enclosure in brackets or a note in area 7.³

e.g. Boetij viri celeberrimi de [con]solatio[n]e
phylosophie liber : cu[m] optimo [com]me[n]to beati Thome

Dialogus beati Gregorij Pape : eiusq[ue] diaconi
Petri in quattuor libros diuisus : de vita [et]
miraculis patru[m] italicor[um] : [et] de
eternitate a[n]imarum

When the meaning of an abbreviation or contraction is conjectural, a question mark is given following the conjectural expansion, e.g. amico[rum?] etc. When the meaning of an abbreviation or contraction cannot be determined, give a question mark, enclosed in square brackets, for each indeterminable abbreviation or contraction e.g. amico[?], [?]s, or leave the abbreviation or contraction as it stands.

0.8 Capitalization

In general, the first letter of the first word of each area should be a capital; the first letter of the first word of some elements (e.g. general material designation, parallel title, alternative title, section title) should also be a capital. Other capitalization should follow the appropriate usage for the language(s) and/or script(s) used in the description (see 0.6). When more than one language and/or script appears in the description, each should be

capitalized in accordance with the usage of that language and/or script even when this produces an inconsistent pattern of capitalization for the description as a whole.

However, lower case letters are never transcribed into capitals. In converting capitals to lower case, the usage (including that of diacritics) in the publication being described should be followed. The following usage is recommended for converting I, J, U, V and VV where practice is not consistent:

I or J as i (but final IJ as ij and final larger capital I as I);
U and V as u (but initial U or V as v);
VV as uu (but vv initially).

Black letter capitals in the form J or U are transcribed as I or V. Letters of numerical value in a chronogram in the title or imprint, or in a chronistic in the text of a poem are given in capitals. Interpolations used by the cataloguing agency should follow modern practice.

DCRB

0G. Misprints, etc.

In an area where transcription from the publication is required, transcribe a misprint as it appears in the publication. Follow such an inaccuracy either by "[sic]" or by the abbreviation "i.e." and the correction within square brackets. Supply missing letters in square brackets.

An hnuble [sic] address
The notted [i.e. noted] history of Mother Grim
One day's d[u]ty

Do not correct words spelled according to older or non-standard orthographic conventions, e.g., "françoise" for "française," or "antient" for "ancient."

When the printer has left blank space for an initial letter, give the letter without square brackets, regardless of whether a guide letter is present or the letter has been filled in by hand. Make a note to show the copy's actual state in this respect.

Historiarum libri XXXV
Note: Space for initial letter of first word of title left blank by printer
Note: LC copy: Initial letter supplied in red and green ink

0H. Forms of diacritical marks and letters (including capitalization)

In general do not add accents and other diacritical marks that are not present in the source.

In general transcribe letters as they appear. Convert earlier forms of letters and diacritical marks, however, to their modern form (cf. Appendix B). In most languages, including Latin, transcribe a ligature by giving its component letters separately. Do not, however, separate the component letters of æ in Anglo-Saxon; œ in French; or æ and œ in ancient or modern Scandinavian languages. (For the transcription of i/j and u/v, see below.) When there is any doubt as to the correct conversion of elements to modern form, transcribe them from the source as exactly as possible.

Convert to uppercase or lowercase according to the rules for capitalization in AACR 2, Appendix A. ... When the rules for capitalization require converting i/j or u/v to uppercase or lowercase, adhere to the pattern of uppercase/lowercase employed by the particular printer. Only when a pattern cannot be determined should the following table for conversion be applied, for it represents a solution of last resort.

Transcribe into lowercase:

- I or J as i
- II as ii
- IJ as ij
- U or V as u (but U or V in initial position as v)
- VV as uu (or vv in initial position)

Transcribe into uppercase:

- i as I
- j as J
- u or v as V
- uu or vv as VV (i.e., two capital V's)

Do not convert to lowercase a final capital I when the preceding letters of the word are printed either in lowercase or in smaller capitals.

Treat gothic capitals in the forms J and U as I and V. (In "modern" gothic where lowercase i and j are distinguished, transcribe the gothic capitals according to the lowercase usage.)

Capital letters occurring apparently at random or in a particular sequence on a title page or in a colophon may represent a chronogram. Where there is good reason to assume that a chronogram is being used, do not convert letters considered part of the chronogram from uppercase to lowercase, or from lowercase to uppercase. See also 4D2.

0J. Abbreviations

0J1.

In an area where transcription from the publication is required, do not abbreviate any word except as permitted by 2B1.

0J2.

When special marks of contraction have been used by the printer in continuance of the manuscript tradition, expand affected words to their full form and enclose supplied letters in square brackets. When an abbreviation standing for an entire word appears in the source, record instead the word itself, and enclose it in square brackets. If the Tironian sign [cannot reproduce sign] cannot be reproduced, treat it as an abbreviation and substitute "[et]" for it. Transcribe an ampersand as an ampersand. Enclose each expansion or supplied word in its own set of square brackets, e.g., "... amico[rum] [et] ..."

Esopus co[n]structus moralizat[us] [et] hystoriatus ad vtilitate[m] discipulo[rum]

When the meaning of an abbreviation or contraction is conjectural, use the question mark after the supplied letters or word within the same set of brackets, e.g., "... amico[rum?] [et] ..." When the meaning of an abbreviation or contraction cannot be determined, substitute a question mark within brackets for each element in question, e.g., "... amico[?] [?] ..."

Appendix 2
Survey of information sources for converting IJUV

This appendix contains the results of several surveys of early books and their textual patterns, using the *Examples to Accompany DCRB* and the Folger collection. The survey grew organically, rendering these tables, particularly what data were collected, a little disorganized and inconsistent. However, we would be very interested in whether others performing surveys on their own early books support or belie our results, and encourage those interested to do so! One apparent problem is the lack of an initial consonantal lowercase **u** represented in the sample. However, even though it is a common (but probably minority) continental use, the problem may be lesser than thought if such a construction rarely involves title page transcription.

It may theoretically be possible for black letter titles to involve conversion, but since black letter almost always uses a normal upper- lowercase distribution, and since there are specific rules for converting “gothic” initials, I found that the designation of “black letter” enough to indicate that no conversion was necessary.

The column “graphs to be converted” indicates first the uppercase graph needing transcription conversion, and then the source of the lowercase use in the same typeface. “T.p.” means that the appropriate lowercase use is evident on the t.p.; “text” indicates that one need consult the interior text for lowercase use.

Examples to Accompany DCRB

No.	Date	Language	City	Graphs to be converted
1	1480	Latin	Strassburg	black letter
2	1494	Greek		not applicable
3	1501	Latin	Venice	black letter
4	1501	Latin	Venice	black letter
5	151-?	Latin	Venice	black letter
6	1511	Latin	Venice	black letter
7	1523	Latin	Basil	medial vocalic V : t.p.
8	1527	Latin	Paris	medial vocalic V : t.p. final II : text medial II : text Note: probable transcription error in book
9	1539	English	London	black letter
10	1541	Hebrew		not applicable
11	1553-1554	Italian	Venice	initial cons. V : text medial vocalic V : text medial cons. V : text

No.	Date	Language	City	Graphs to be converted
12	1561	Latin	Lyon	medial vocalic V : t.p. final II : text
13	1578	French	Paris	medial vocalic V : t.p.
14	1584	Latin	Lyon	final II : t.p. medial vocalic V : t.p.
15	1585	Italian	Palermo	medial cons. V : t.p.
16	1592	Latin	Lyon	medial cons. V : t.p. medial vocalic V : t.p. Note: incorrect 246
17	1592	German	Frankfurt	black letter
18	1602-1648			not applicable
19	1621	Spanish	Barcelona	medial vocalic V : t.p.
20	1636	Dutch	Middelburgh	medial vocalic V : t.p.
21	1643	Latin	Louvain	medial vocalic V : t.p. final II : text
22	1649	French	Amsterdam	medial vocalic V : t.p. initial I : text

Rest of examples are modern use

Analysis. Pre-modern use of **ijuv** is found up to and including Example 22, 1649. Of the 22 works under consideration, seven are black letter and therefore not problematic. Three are inapplicable for other reasons. The remaining 12 require some conversion of case. For six, use of the lowercase letters in the same typeface was found in the t.p. itself; the cataloger need not resort to the text. Five more were partly solved by the t.p. and partly by resorting to the text. There was only one in which the t.p. was no help at all, example 11.

Sample of 10 Folger continental books from 1570-1571 with **IJUV** conversions

	Date	Language	City	Graphs to be converted
1	1570	Latin	Basel	medial cons V : t.p. medial vocalic V : t.p. final II : t.p.
2	1570	Latin	Ingolstadt	medial vocalic V : t.p. medial cons V : t.p.
3	1571	French	Paris	medial vocalic V : t.p.
4	1571	Latin	Wittenberg	medial cons V : t.p. medial vocalic V : t.p.
5	1571	Latin	Venice	medial vocalic V : t.p. medial cons V : text final II : text
6	1571	Italian	Naples	medial vocalic V : t.p.

	Date	Language	City	Graphs to be converted
7	1571	Latin	Rome	final II : text Note: no definitive answer possible; t.p. in uppercase Roman, text entirely in Italic
8	1571	Latin	Lyon	medial vocalic V : t.p.
9	1570	Italian	Florence	medial vocalic V : t.p. medial cons V : t.p.
10	1571	French	Paris	medial vocalic V : t.p.

Analysis. A total of 13 books were examined, of which three had no need of conversion. Of the ten t.p.'s, five conversions were resolved from the t.p., three required partial recourse to the text, and two could be solved from the text only. Example 7 is a not uncommon occurrence—a t.p. entirely in uppercase Roman, but text entirely in Italic. It is therefore impossible to find the pattern used in the text for that typeface. Another particular problem with continental books in Latin is the predominance of titles beginning with the author's name in genitive (M. Tullij Ciceronis) and/or containing a place name in ablative (Venetijs). Ron Bogdan, the Folger continental cataloger, tells me that he sometimes finds a difference in text usage between II used in the genitive in contrast to the ablative. This may be solved by understanding that medial and final II must each be supported separately. He also says that finding II use in the text is easily the most time-consuming part of transcription.

In order to ascertain whether there is a discernible difference between continental and English books, I decided to find ten English books from the same time period (1570-1571) needing conversion, which I could not do, either because they were in black letter or the title was already in lower-text. I therefore browsed the shelves in the Folger STC vault, whose results are included below. Almost without exception, anything before 1580 required no conversion.

DFo STC books

	STC no.	Date	Language	City	Graphs to be converted
1	STC 19498	1593	English	London	medial vocalic V : t.p.
2	STC 19500	1606	English	London	medial vocalic V : t.p.
3	STC 19528	1603	English	Edinburgh	medial vocalic V : t.p.
4	STC 19531	1594	English	London	medial vocalic V : t.p.
5	STC 19560	1612	Latin	London	medial vocalic V : t.p. medial cons V : text
6	STC 19565	1610	English	London	medial vocalic V : t.p. medial cons V : text initial vocalic V : text
7	STC 19566	1631	English	London	medial vocalic V : t.p.
8	STC 19567	1607	English	London	medial cons V : text
9	STC 19688	1593	English	London	medial cons V : text
10	STC 20309	1573	Latin	London	medial vocalic V : t.p.

Analysis. Of the ten t.p.'s needing conversion, six are resolvable by the t.p. alone, two require partial resort to the text, and for two the t.p. is no help at all.

I examined all of the books requiring recourse to the text, determined what the correct DCRB transcription would be, and then compared that to use of the table of last resort. The penultimate column shows what the actual conversion graph is according to lowercase textual use, and the final column what graph would be assigned according to the OH table.

DFo STC books

No.	date	Graphs to be converted	textual use	tabular conversion
5	STC 19560	medial cons V : text	u	u
6	STC 19565	medial cons V : text initial vocalic V : text	u v	u v
8	STC 19567	medial cons V : text	u	u
9	STC 19688	medial cons V : text	v	u

The last example provides the only conflict between the table of last resort and textual usage. The title reads: A DIRECTION FOR THE GOVERN-ment of the *tongue* accor-*ding* to God's Word.

DCRM transcription following textual pattern: A direction for the government of the tongue according to God's Word

Transcription using the t.p. and table only: A direction for the government of the tongue according to God's Word

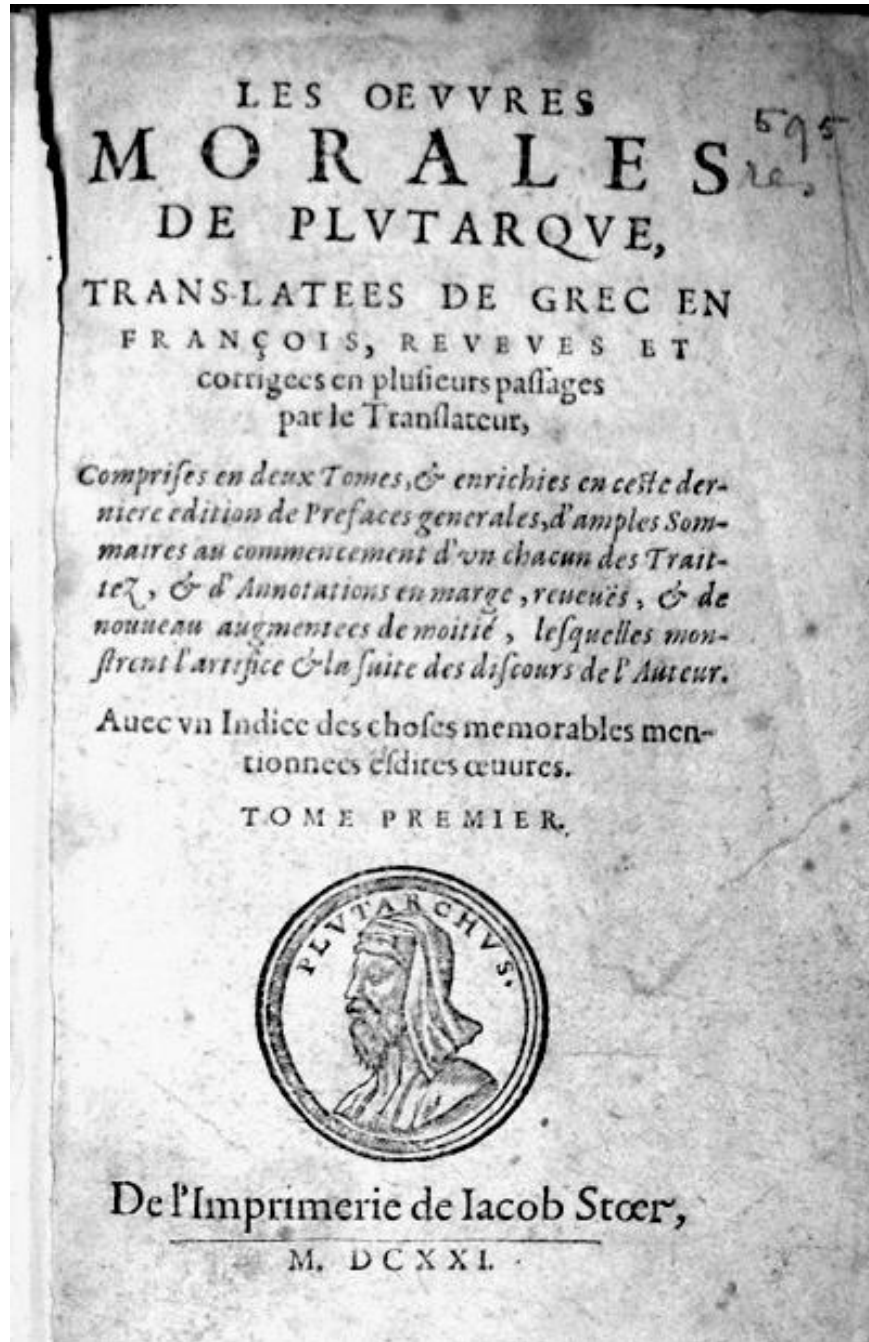
In identifying textual use, I am assuming the transcription provided with the example is correct, with the one exception of the probable transcription error which involving a final II.

Examples to Accompany DCRB

No.	date	Graphs to be converted	textual use	tabular conversion
7	1523	medial vocalic V : t.p.	u	u
8	1527	medial vocalic V : t.p. final II : text medial II : text Note: probable transcription error in book	u ij? ii?	u ii ii
11	1553- 1554	initial cons. V : text medial vocalic V : text medial cons. V : text	v u u	v u u
12	1561	medial vocalic V : t.p. text final II : text	u ij	u ii
13	1578	medial vocalic V : t.p.	u	u

No.	date	Graphs to be converted	textual use	tabular conversion
14	1592	final II : t.p. medial vocalic V : t.p.	ij u	ii u
15	1585	medial cons. V : t.p.	u	u
16	1592	medial cons. V : t.p. medial vocalic V : t.p. Note: incorrect 246	u u	u u
19	1621	medial vocalic V : t.p.	u	u
20	1636	medial vocalic V : t.p.	u	u
21	1643	medial vocalic V : t.p. final II : text	u ii	u ii
22	1649	medial vocalic V : t.p. initial cons. V : t.p. initial I : text final II : text	u v i ii	u v i ii

Appendix 3
Sample title page illustrated with different transcription methods
(B. Hillyard & DJ Leslie)



LES OEUVRES
MORALES
DE PLVTARQVE,
TRANSLATEES DE GREC EN

FRANÇOIS, REVEVES ET
corrigees en plusieurs passages
par le Translateur,
*Comprises en deux Tomes, & enrichies en ceste der-
niere edition de Prefaces generales, d'amples Som-
maires au commencement d'vn chacun des Trait-
tez, & d'Annotations en marge, reueuës, & de
nouueau augmentees de moitié, lesquelles mon-
strent l'artifice & la suite des discours de l'Auteur.*
Auec vn Indice des choses memorables men-
tionnees esdites œuures.
TOME PREMIER.
De l'Imprimerie de Iacob Stoer,
M.DCXXI.

Method 1: Normalized modern transcription

Les oeuvres morales de Plutarque, translatees de grec en françois, reueues et corrigees en plusieurs passages par le translateur, comprises en deux tomes, & enrichies en ceste derniere edition de prefaces generales, d'amples sommaires au commencement d'un chacun des traittez, & d'annotations en marge, reueuës, & de nouveau augmentees de moitié, lesquelles monstrent l'artifice & la suite des discours de l'auteur. Avec un indice des choses memorables mentionnees esdites oeuvres. De l'imprimerie de Iacob Stoer, M.DCXXI.

Method 2: Normalized tabular transcription

Les oeuvres morales de Plutarque, translatees de grec en françois, reueues et corrigees en plusieurs passages par le translateur, comprises en deux tomes, & enrichies en ceste derniere edition de prefaces generales, d'amples sommaires au commencement d'vn chacun des traittez, & d'annotations en marge, reueuës, & de nouveau augmentees de moitié, lesquelles monstrent l'artifice & la suite des discours de l'auteur. Avec vn indice des choses memorables mentionnees esdites oeuvres. De l'Imprimerie de Iacob Stoer, M.DCXXI.

Note: result same as textual use transcription, the current rules.

Method 3: Converted graphical transcription

Les oeuvres morales de Plutarque, translatees de grec en françois, reueues et corrigees en plusieurs passages par le translateur, comprises en deux tomes, & enrichies en ceste derniere edition de prefaces generales, d'amples sommaires au commencement d'vn chacun des traittez, & d'annotations en marge, reueuës, & de nouveau augmentees de moitié, lesquelles monstrent l'artifice & la suite des discours de l'auteur. Avec vn indice des choses memorables mentionnees esdites oeuvres. tome premier. De l'Imprimerie de Iacob Stoer, M.DCXXI.

Note particularly reueves alongside reueuës and oeuvres alongside oeuvres

Method 4: Textual use transcription (current rules)

Les oeuvres morales de Plutarque, translatees de grec en françois, reueues et corrigees en plusieurs passages par le translateur, comprises en deux tomes, & enrichies en ceste derniere edition de orefaces generales, d'amples sommaires au commencement d'un chacun des traittez, & d'annotations en marge, reueuës, & de nouveau augmentees de moitié, lesquelles monstrent l'artifice & la suite des discours de l'auteur. Avec vn indice des choses memorables mentionnees esdites oeuvres. De l'Imprimerie de Iacob Stoer, M.DCXXI.