TO: The Unicode Technical Committee and JTC1/SC2 WG2

FROM: Luanne von Schneidemesser (Senior Editor, *Dictionary of American Regional English*, University of Wisconsin-Madison), Lewis Lawyer (C. Phil., Dept. of Linguistics, UC Davis), Ken Whistler (Sybase), and Deborah Anderson (SEI, UC Berkeley)

RE: Proposal for Two Phonetic Characters

DATE: 12 July 2012

This proposal requests the following two characters, both of which appear in Pullum and Ladusaw's *Phonetic Symbol Guide* (Chicago, 1986) but are not yet in Unicode:

      (1) LATIN SMALL LETTER INVERTED ALPHA

      (2) LATIN LETTER SMALL CAPITAL OMEGA

I. PROPOSED CHARACTERS

1. ɒ LATIN SMALL LETTER INVERTED ALPHA

The character "Inverted Script A" appears on page 11 of the *Phonetic Symbol Guide* (Pullum and Ladusaw, 2nd edition, Chicago, 1986) with the following glyph, which is enlarged in order to show the shape more clearly:

The use of this character is described as follows (Pullum and Ladusaw, p. 11):

> Not in general use. Kurath (1939, 126 [figure 1b]) reports that some of the field workers for the *Linguistic Atlas of New England* (e.g., Bloch) had used Turned Script A indifferently for both rounded and unrounded low back vowels. This character was invented by analogy to unambiguously denote an unrounded low back vowel so that Turned Script A [U+0252] could be reserved for a rounded low back vowel in future work. Extremely rare.

This character, also known as *turned script a right* or *reverse turned script a,* appeared in the vowel chart of Hans Kurath's *Linguistic Atlas of New England* (see figure 1a), with an accompanying description (see figure 1b).

Although this character is not in the International Phonetic Alphabet, it has been adopted by other American linguistic atlases, such as *The Linguistic Atlas of the Upper Midwest* (Allen 1973-76, see figures 2a and 2b), *Handbook of the Linguistic Atlas of the Middle and South Atlantic States* (Kretzschmar et al. 1993), *Handbook of the Linguistic Atlas of the Gulf States* (Pederson 1986), as well as other publications, such as Kretzschmar 2003 (see figure 3).

LATIN SMALL LETTER INVERTED ALPHA also appears in the five-volume *Dictionary of American Regional English*[1] (see figure 4), which has just completed its print publication. The *DARE* project is currently working with Harvard University Press to produce a full electronic edition of the dictionary set to launch in 2013. In order to be able to represent the text in its electronic version, *DARE* requires LATIN SMALL LETTER INVERTED ALPHA. The character also appears in other materials cited in *DARE*, including *Dialect Notes*, *Publications of the American Dialect Society*, and *American Speech*, all from the American Dialect Society.

The proposed name is based on U+0251 ɑ LATIN SMALL LETTER ALPHA, since this character is an inverted version of LATIN SMALL LETTER ALPHA.

Figures

4. The vowel symbols of the Linguistic Atlas alphabet are presented in the following diagrammatic arrangement, which differs only in detail from the diagram of the IPA. For an explanation of the general principles of this arrangement the reader is referred to various works by members of the IPA (especially Daniel Jones, *An Outline of English Phonetics*,

Figure 1a. Source: *Handbook of the Linguistic Geography of New England* (Kurath 1939), page 123

14. ɒ usually denotes a weakly rounded low-back vowel; but several of the field workers use it also for the unrounded variety. A low-back vowel with full rounding is sometimes written ɒ (see **29**). In the later regional atlases of the United States and Canada the use of ɒ in this double function will be avoided by the addition of a new symbol: ɒ will be written only for the rounded variety, ɒ̑ for the unrounded variety. Weak rounding will then be

Figure 1b. Source: *Handbook of the Linguistic Geography of New England* (Kurath 1939), page 126

```
American ˜/ɪ/ˊ and /ɨ/ replaces /ɪ/. Also
used here is the American symbol for
the low back unround vowel, /ɑ/, which,
though not used in the original New Eng-
land fieldwork, has been used in subse-
sequent American research.
```

Figure 2a. Source: *The Linguistic Atlas of the Upper Midwest* (Allen 1973-76), vol. 1, p. 122.

```
sausage   37.4a. NE 304. PEAS 162; F132.

   The dominant eastern and UM pronuncia-
tion of sausage is with a rounded vowel
ranging from /ɔ/ to /ɒ/. In contrast is a
scattered variant with an unrounded vowel
ranging from /ɑ/ to /ɑ/. The eastern New
England minority form with /æ/ is not re-
corded in the UM.
```

Figure 2b. Source: *The Linguistic Atlas of the Upper Midwest* (Allen 1973-76), vol. 3, p. 259.

Frontness vector

i (y) ~ e (ø) ~ ɪ (ʏ) ~ ɛ (œ) ~ ɨ (ʉ) ~
  æ ~ ɘ ~ ɨ (ʉ) ~ a ~ ə ~ ɜ (ɞ) ~ ɐ ~ ɑ ~
  ɤ (ʊ) ~ θ ~ ʌ ~ ɯ (u) ~ o ~ ɔ ~ ɑ (ɒ)

Figure 6. Realizations of vowels in atlas phonetics
(rounded = ( ), see Kretzschmar *et al.*, 1993 for a full description of Atlas phonetics)

Figure 3: Source: Kretzschmar 2003, page 385.

|ɑ| *cot;* a lower-low-central unrounded vowel; see also
     Pronc Intro 3.II.9, 10
       [a] fronted variant
       [ɑ] retracted variant

Figure 4. Source: Pronunciation Guide from *Dictionary of American Regional English* (Cassidy and Hall 1985-2012), vol. 5.

References:

Allen, Harold B. 1973-76. *The Linguistic Atlas of the Upper Midwest*. 3 vol. Minneapolis MN: University of Minnesota Press.

Cassidy, Frederic G., and Joan Houston Hall, eds. 1985-2012. *Dictionary of American Regional English (DARE)*. 5 vols. Cambridge MA: Belknap Press of Harvard University Press.

Kretzschmar, William A., Jr., Virginia G. McDavid, Theodore K. Lerud, and Ellen Johnson. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago: University of Chicago Press.

Kretzschmar, William. 2003. "Distributional Foundations for a Theory of Language Change." *World Englishes* 22 (2003), 377-401. [with Susan Tamasi]

Kurath, Hans. 1939. *Handbook of the Linguistic Geography of New England*. Providence: Brown University, for ACLS. [2nd ed., rev., New York: AMS Press, 1973.]

Pederson, Lee. 1986. *Handbook of the Linguistic Atlas of the Gulf States*. Athens: University of Georgia Press.

2. Ω LATIN LETTER SMALL CAPITAL OMEGA

The character "Small Capital Omega" appears in *Phonetic Symbol Guide* (Pullum and Ladusaw, 2nd edition, Chicago, 1986) with the glyph:

Ω

This character is described as follows (Pullum and Ladusaw, p. 148):

> Proposed by Block and Trager (1942, 22: see also Trager 164 (p. 26)) for a mid back rounded vowel ("mean-mid" in Bloch and Trager's terms, i.e., between the height of Cardinals 6 and 7, IPA [ɔ] and [o]). The IPA transcription for such a vowel would be [o̞] or [ɔ̝]. A large capital *omega* is used by some Indo-Europeanists for an *o*-coloring laryngeal (see, e.g., Hamp 1954a (p. 123, note 3).

The symbol was an innovation by Bloch and Trager (1942), and until recently was not believed to have been adopted by other linguists. However, Lewis Lawyer, a linguistics graduate student at UC Davis, has found regular usage of LATIN LETTER SMALL CAPITAL OMEGA in Elizabeth Bright's field notes and grammar from 1952 and in the field notes of Don Ultan, dating to 1960/61. Lawyer was working on materials for the Patwin language, a severely endangered member of the Wintuan family of languages spoken in the area of Sacramento Valley of Northern California.

Lawyer is currently digitizing the field notes, and requires this character, which appears in a text corpus repeatedly as a standard character for phonetic transcription. It now appears that

some linguists trained at UC Berkeley in the late 1950's or early 60's may have learned the Bloch and Trager phonetic conventions and used them when transcribing actual linguistic material.

Figure 5 shows the character in Elizabeth Bright's field notes, which date to 1952. A description of the character is given in E. Bright's grammar in figure 7. Figure 8 shows the character in the original field notes of Don Ultan.

Figures 6 and 9 are transcriptions with GREEK LETTER CAPITAL OMEGA in place of LATIN LETTER SMALL CAPITAL OMEGA. The Greek letter is clearly not correct, because the small capital omega is notionally lowercase, as can be seen in its use with other lowercase phonetic letters in Latin transcription. The character is intended specifically as an Americanist phonetic transcription letter with particular phonetics, and hence is eligible for separate encoding.

Figures
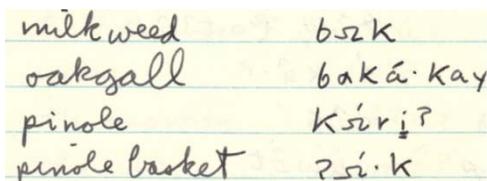
Figure 5. An excerpt from Elizabeth Bright's field notebooks in which some words are transcribed with a small-capital omega (Bright 1952b).
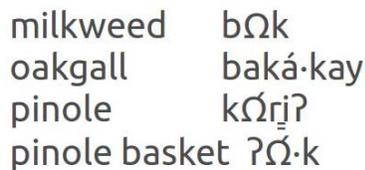

Figure 6. A transcription by Lewis Lawyer of text from figure 5, rendered in the same font as the transcription image in figure 9, using GREEK LETTER CAPITAL OMEGA in place of LATIN LETTER SMALL CAPITAL OMEGA.
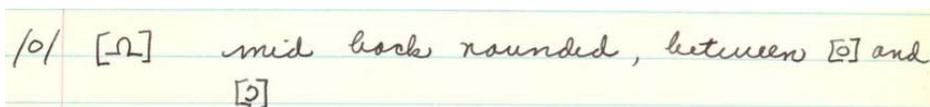

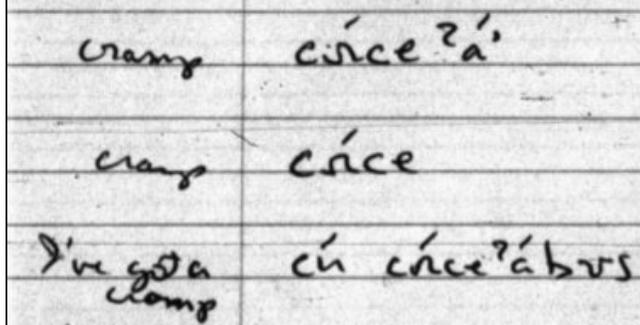*Figure 7.* A phonetic description of the Patwin /o/ from Elizabeth Bright's 1952 grammar of Patwin (Bright 1952a).

Figure 8. Original field notes from Don Ultan (Ultan 1961–1962).

Figure 9. Transcription by Lewis Lawyer with GREEK LETTER CAPITAL OMEGA in place of LATIN LETTER SMALL CAPITAL OMEGA.

References:
Bright, Elizabeth. 1952a. Patwin grammar. Two notebooks. Handwritten draft of a grammar. SCOIL: BrightE.002.004–005.
Bright, Elizabeth. 1952b. Patwin field notes. Eight notebooks. Includes texts, vocabulary, and notes on grammar. SCOIL:BrightE.002.001–003,005–010.
Ultan, D. 1961–1962. Patwin fieldnotes. Part of the Harvey Pitkin Papers: Mss.Ms.Coll.78 Series I-A, Oversize 1: Folder Series IA-1–IA-7. American Philosophical Society, Philadelphia. (Microfilm at SCOIL, UC Berkeley: UltanD.m001.)

II. Proposed Location
The proposed location for the two proposed characters is in the Latin Extended-D block, at U+A7AE and U+A7AF.

III. Character Properties
The gc character properties are:
A7AE; LATIN SMALL LETTER INVERTED ALPHA;Ll;0;L;;;;;N;;;;;
A7AF; LATIN LETTER SMALL CAPITAL OMEGA;Ll;0;L;;;;;N;;;;;

IV. Collation
LATIN SMALL LETTER INVERTED ALPHA should sort between 1D9B MODIFIER LETTER SMALL TURNED ALPHA and 0062 LATIN SMALL LETTER B.

LATIN LETTER SMALL CAPITAL OMEGA should be sorted after LATIN LETTER O WITH LONG STROKE OVERLAY but before LATIN SMALL LETTER CLOSED OMEGA

A. Administrative

1. **Title:**   *Request for two phonetic characters*

2. Requester's name: *Luanne von Schneidemesser (Senior Editor, Dictionary of American Regional English, University of Wisconsin-Madison), Lewis Lawyer (C. Phil., Dept. of Linguistics, UC Davis), Ken Whistler (Sybase), and Deborah Anderson (SEI, UC Berkeley)*

3. Requester type (Member body/Liaison/Individual contribution):   *Liaison contribution*

4. Submission date: *9 July 2012*

5. Requester's reference (if applicable):

6. Choose one of the following:
    This is a complete proposal: *yes*
    (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
    a. This proposal is for a new script (set of characters):
        Proposed name of script:
    b. The proposal is for addition of character(s) to an existing block: *yes*
        Name of the existing block: *Latin Extended-D*

2. Number of characters in proposal:

3. Proposed category (select one from below - see section 2.2 of P&P document):
    A-Contemporary   **x**   B.1-Specialized (small collection)   B.2-Specialized (large collection)
    C-Major extinct   D-Attested extinct   E-Minor extinct
    F-Archaic Hieroglyphic or Ideographic   G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?   *yes*
    a. If YES, are the names in accordance with the "character naming guidelines"
        in Annex L of P&P document?   *yes*
    b. Are the character shapes attached in a legible form suitable for review?   *yes*

5. Fonts related:
    a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

6. References:
    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?   *yes*
    b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
    of proposed characters attached?   *yes*

7. Special encoding issues:
    Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?   *yes*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

[1] Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *no*
    If YES explain

2. Has contact been made to members of the user community (for example: National Body,
    user groups of the script or characters, other experts, etc.)? *yes*
        If YES, with whom? *Linguists and editors of Dictionary of American Regional English*
        If YES, available relevant documents: *See proposal*

3. Information on the user community for the proposed characters (for example:
    size, demographics, information technology use, or publishing use) is included? *Users of DARE*

    Reference: *See proposal*

4. The context of use for the proposed characters (type of use; common or rare) *rarish*
    Reference:

5. Are the proposed characters in current use by the user community? *yes*
    If YES, where? Reference: *See DARE and other examples from the proposal*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
    in the BMP?
        If YES, is a rationale provided?
            If YES, reference:

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *yes*

8. Can any of the proposed characters be considered a presentation form of an existing
    character or character sequence? *no*
        If YES, is a rationale for its inclusion provided?
            If YES, reference:

9. Can any of the proposed characters be encoded using a composed character sequence of either
    existing characters or other proposed characters? *no*
        If YES, is a rationale for its inclusion provided?
            If YES, reference:

10. Can any of the proposed character(s) be considered to be similar (in appearance or function)
    to, or could be confused with, an existing character? *yes (omega)*
        If YES, is a rationale for its inclusion provided? *yes*
            If YES, reference: *See proposal*

11. Does the proposal include use of combining characters and/or use of composite sequences? *no*
    If YES, is a rationale for such use provided?
        If YES, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
        If YES, reference:

12. Does the proposal contain characters with any special properties such as
    control function or similar semantics? *no*
        If YES, describe in detail (include attachment if necessary)

13. Does the proposal contain any Ideographic compatibility characters? *no*
    If YES, are the equivalent corresponding unified ideographic characters identified?
        If YES, reference: