- **Doc Type:** Working Group Document
- **Title:** Proposal to add BREATH punctuation marks to the UCS
- **Source:** Andrés Sanhueza
- **Status:** Individual Contribution
- **Date:** 2019-02-19

Proposal

This document proposes to add to the UCS two punctuation marks for the *Supplemental Punctuation* block that are common on American comics yet have no direct equivalent currently: the opening breath mark (⸏) and the closing breath mark (⸐). These marks normally look like three dashes radiating onwards from the word or words they surround (the total width and height being like regular parenthesis and brackets) and are used to wrap non-word breath sounds like "gasp" and "whew" (Chiarello & Klein, 2004). They appear regularly on American comics since at least the mid '40s (Klein, 2010) as a sort of expansion from the art of drawing onomatopaeyas. Their usage is more prevalent on manual lettering, although many custom comic-like digital fonts also features them (Roshell, 2004) or they can sometimes be featured on regular typesetting (Groening, 2010), yet in those latter cases is most common to use parenthesis or surrounding asterisks as a fallback (Yus, 2011).

Besides being called "breath marks", the characters are also referred by various names such as "whiskers", "fireflies" and "crow's feet" (Piekos, 2009) (Klein, 2010). The name "breath mark" is proposed as they relate better to their function and their status as symbols. It can cause confusion for some with the otherwise unrelated U+1D112 𝄒 MUSICAL SYMBOL BREATH MARK, yet there's a precedent of the "Musical symbol" name motif already is used to indicate a "brace" and a "bracket" characters (U+1D114, U+1D115), which are also known regular punctuation earlier in the charts.

Sometimes, empty opening and closing breath marks can be used when the dialog show a person unresponsive or fainting (Piekos, 2009) (Edwards, 1994). On these cases, the characters often looks more like an open asterisk (Zimmer, 2013). It could be arged that that counts as a stand-alone character on its own —specially as some glyph variants aren't symetrical—, but that's outside the scope of this proposal as these characters are more related to the symbols used in comics to represent swears words (Law, 2010) and if there's the need to add those to the UCS certainly requires separate and mayor research due to being long in quantity. However, using the opening and closing mark to compose a single asterisk-like symbol is a common practice on custom digital fonts that features the characters, mapping them to other lesser used characters (Roshell, 2004) (Piekos, 2009)

Reasoning for not map to existing characters

The proposed characters have the properties of opening and closing punctuation marks respectively and their default look on a text font use roughly the same total width and height that a parenthesis and other Latin brackets. Therefore, they aren't suitable to be mapped to similar-looking existing characters.

Three lines converging right/Three lines converging left

The characters mapped at U+269E and U+269F originated on the ARIB STD-B24 standard for closed caption. These are both symbols intended to work among text independently of each one and their sample glyphs gives them both an emoji-like width.

Three rays left/Three rays right

The characters mapped at U+1F5E6 and U+1F5E7 —along the "bottom" and "top" variants at U+1F5E4 and U+1F5E5—, originated from the Webdings font and are similar in spirit to the "Rocket Ornaments" (U+1F66C to U+1F66F) from the same font in that are intended as generic symbols for decoration rather that actual punctuation, mostly noticed in that the sample glyphs have CJK-like spacing.

Character Properties

```
2E4F;LEFT BREATH MARK;Ps;0;ON;;;;;Y;;;;;
2E50;RIGHT BREATH MARK;Pe;0;ON;;;;;Y;;;;;
```

Bibliography and References

Chiarello, M., & Klein, T. (2004). *The DC Comics Guide to Coloring and Lettering Comics.* New York, USA: Watson-Guptill Publications.

Edwards, J. (1994). Archie in: "For Pete's Sake". *Archie*, 6.

Fox, G. (1945). Wildcat. *All Star Comics*, 28.

Groening, M. (2010). *Simpsons World the Ultimate Episode Guide: Seasons 1-20.* USA: HarperCollins.

Hughes, R. (1966). Popnecker the Pilgrim! *Herbie*, 12.

Klein, T. (2010, March 7). *Punctuating Comics: Breath Marks*. Retrieved from Todd's Blog: https://kleinletters.com/Blog/punctuating-comics-breath-marks/

Law, G. (2010, July 9). *Grawlixes.* Retrieved from Statoids: http://www.statoids.com/comicana/grawlist.html

Lee, S., & Ditko, S. (1962). Spider-man! *Amazing Fantasy*, 2.

Piekos, N. (2009). *Comic Book Grammar & Tradition*. Retrieved February 18, 2019, from Blambot Comic Fonts & Lettering: http://www.blambot.com/articles_grammar.shtml

Roshell, J. (2004, April 1). *Font and Character Usage.* Retrieved from Balloon Tales: http://www.balloontales.com/font-and-character-usage/

Watterson, B. (1987). *Calvin and Hobbes (Book 1).* Kansas, USA: Andrews McMeel Publishing.

Yus, F. (2011). *Cyberpragmatics: Internet-mediated Communication in Context.* John Benjamins Publishing.

Zimmer, B. (7 de February de 2013). *The cyberpragmatics of bounding asterisks*. Obtenido de Language Log: http://languagelog.ldc.upenn.edu/nll/?p=4466

Table of new characters proposed

Supplemental Punctuation

Breath marks for comics

Code	Glyph	Name	Comments
2E4F		LEFT BREATH MARK	→ ⩾ 269E three lines converging right → ⩾ 1F5E6 Three rays left
2E50		RIGHT BREATH MARK	→ ⩽ 269F three lines converging left → ⩽ 1F5E7 Three rays right

Add comments to existing characters

Code	Glyph	Name	Comment to add
269E		THREE LINES CONVERGING RIGHT	→ 2E4F left breath mark
1F5E6		THREE RAYS LEFT	
269F		THREE LINES CONVERGING LEFT	→ 2E50 right breath mark
1F5E7		THREE RAYS RIGHT	

Examples

Figure 1: Breath marks around the word "whew" at a Wildcat comic. Glyph features more dashes than usual. (Fox, 1945)

Figure 2: Breath marks around the word "sob" at the comic debut of Spider-man. (Lee & Ditko, 1962)
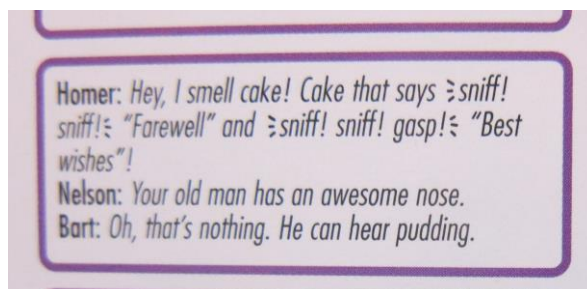

Figure 3: Various breath marks used between a sans-serif font for dialog transcriptions of episodes of *The Simpsons* (Groening, 2010)

Figure 4: The word "sigh" surrounded twice by breath marks. (Hughes, 1966)

Figure 5: Empty opening and closing breath marks at the end of a sentence. (Edwards, 1994)
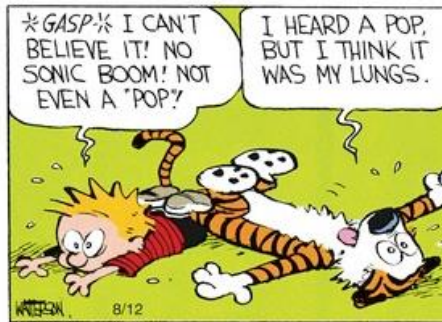

Figure 6: Characters doing the function of the breath marks around the word "gasp" at a 1986 Calvin & Hobbes strip. Although the signs resemble more an asterisk or explosion, it could be argued they are the opening and closing breath marks at both sides (Watterson, 1987)

Figure 7: Table showing how a custom comic font for DC Comics maps the brackets to breath marks (Chiarello & Klein, 2004)

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html
for guidelines and details before filling this form.
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** ***Proposal to encode breath punctuation marks to the UCS***
2. Requester's name: *Andrés Sanhueza*
3. Requester type (Member body/Liaison/Individual contribution): *Individual*
4. Submission date: *2019-02-19*
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): *No*
        Proposed name of script:
    b. The proposal is for addition of character(s) to an existing block: *Yes*
        Name of the existing block: *Supplemental Punctuation*
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?
                                The author
        b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
                                The author
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)? *No*

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: N4502-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.)?	*No*
If YES, with whom?	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	*Everybody*
Reference:	*See text*
4. The context of use for the proposed characters (type of use; common or rare)	*Common*
Reference:	*See text*
5. Are the proposed characters in current use by the user community?	
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	*Yes*
If YES, is a rationale provided?	*Yes*
If YES, reference:	*To keep it along similar characters*
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*
If YES, is a rationale for its inclusion provided?	*Yes*
If YES, reference:	*See text*
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:	