

Doc Type: Complete proposal, documentation and code charts in separate files

Title: Proposal to include Duployan Shorthands and Chinook script in UCS

Source: Van Anderson

Status: For submission to UTC.

Action: For review by UTC and community. Comments to [Van Anderson](#)

Date: 2010-01-31

Discussion list: [Chinook in the UCS](#)

L2/10-026

Historical Overview of the Duployéan and adaptations

The Duployéan shorthands and Chinook script are used as a secondary shorthand for writing French, English, German, Spanish, Rumanian, and as an alternative primary script for Chinook Jargon and several first nations' languages of interior British Columbia including Okanagan, Lillooet, Shushwap, and North Thompson. It was invented by Emile Duployé, published in 1860, as a stenographic shorthand for French. It is historically one of the two most commonly used shorthand systems in France, being more popular in southern France and adjacent French speaking areas of other countries. Adaptations of Duployéan are known to have been developed for the representation of English, German, Spanish, and Rumanian. The basic inventory of consonant and vowel signs - all in the first two columns of the allocation - have been augmented over the years to provide more efficient shorthands for these languages and to adapt it to the phonologies of these languages and the languages using Chinook writing.

The Chinook script was an adaptation and augmentation of the Duployéan shorthand by fr. Jean Marie Raphael LeJeune, used for writing the Chinook Jargon and other languages of 19th c. interior British Columbia. Its original use and greatest surviving attestation is from the run of the *Kamloops Wawa*, a (mostly) Chinook Jargon newsletter of the Catholic diocese of Kamloops, British Columbia, published 1891-1923. At the time, the Chinook Jargon trade language was spoken in an area from SE Alaska to northern California, from the Pacific to the Rockies, and sporadic communities in other nearby areas. Although the Chinook Jargon was the lingua franca in many communities in this geographic area, it was generally a spoken, rather than written language. Most attempts at documentation used the Latin script to approximate Jargon phonology, and indeed, dictionaries of the Chinook Jargon are still readily available in these Latinate orthographies. In contrast, the archives of the *Kamloops Wawa*, written in Chinook, includes a considerable dictionary, but also constitutes an unparalleled 3+ decade corpus of Chinook Jargon usage during the height of its spread and utility. There currently exists no formal encoding, in any context, for the representation of the Chinook writing, and the only informal representation is transliteration by means of the Latin orthographies used in writing the Chinook Jargon. Indeed, the submission of the Duployan Shorthands and Chinook script to the [Unicode Consortium](#) has necessitated the creation, from scratch, of the first Duployéan/Chinook font, such an effort currently underway with glyph images available for review.

In 1984, the "Students' Practical Encyclopedia" (*Enciclopedia practică a copiilor*) was published in Romania, containing the "Curs de Stenografie" by Margaret Sfințescu. This shorthand was an adaptation of the Duployéan for Romanian, using a few of the Chinook and Duployan shorthand compound letters as basic letterforms, and several basic vowel forms with diacritics. It also makes use of a "doubling mark" to indicate a general duplication of a word or phonemic form.

The Pernin shorthand was originally published by Helen M. Pernin as "Pernin's Universal Phonography" no later than 1882. There is an alternate version of the Pernin shorthand published as "Pernin's Practical Reporter", that differs in the affixes used. In that same year of 1882, John Mathew Sloan published the Sloan-Duployan method. In 1918, Denis R. Perrault published the Perrault-Duployan system, all of the above being English adaptations of Duployan. All three systems share many characters with Chinook and each other. The most significant anomalies of these systems are the invariant vowel signs in Pernin, the quarter-circle combined consonants, shared between the systems with differing values, the extensive use of vowel diacritics in Sloan, and heavy shading of letters - like voiced consonants in Pitman-based systems - to indicate "r" flavored letters.

Unsupported shorthands. Currently, materials are unavailable to attempt including Carl Brandt's English Duployéan adaptation or George Galloway's extension of the Sloan-Duployan in the current encoding. Similarly, documentation of the adaptations of Duployéan to German and Spanish are unavailable, so complete support for these orthographies is probably not offered in the current allocation. Allocation space has been set aside to reasonably accomodate extensions for some of these extensions of the Duployéan script.

Typology

Duployéan is, at its core, an alphabetic (consonant & vowel) stenographic (line & curve) writing system (cf. the abjad stenography, Pitman). It classifies under the geometric shorthands, in that the model letterforms are generally based on circles and lines (cf. Gregg shorthand's ellipses). In general, there is a visual and functional distinction between consonants, which are based either on lines or large semi-circles and have invariable orientation, ie consonants do not rotate to match with surrounding letters; and the vowels, which are generally based on circles, quarter arcs, and small semi-circles, and almost always reorient based on adjacent characters (exceptions lie in the Pernin vowels). It is an LTR script, proceeding down the page in lines like most modern Western scripts, although individual letters may have strokes right-to-left and rising.

Script Structure

The core repertoire of the Duployéan writing contains several classes of letters, differentiated primarily by visual form and stroke direction, and nominally by phonetic value. Letter classes include the line and arc consonants (9 varieties total), circle vowels (A and O vowels), nasal vowels, and orienting vowels (U/Eu, I/E). In addition, the Chinook writing contains the spacing letters H and X, compound consonants, W-vowels, and one known logograph. The extended Duployéan shorthand includes four other letter classes, the complex letters (multisyllabic symbols, but just consonant forms), and high, low, and connecting terminals for common word endings. The Romanian stenography, Pernin, Perrault, and Sloan orthographies add a few letters or letter forms, and several combined letters. Since the Duployéan was originally developed as a shorthand system, strings of letters are joined together cursorily into words in Duployéan, Romanian, Pernin, Perrault, and Sloan, or nominally syllabic units in Chinook - usually with a single circle vowel for each unit. Most "core" letters have related variant forms, including the addition of ancillary dots and crosses, large size variants, and the compounding of vowels. The original Duployéan and its offshoots all encourage overlapping for initialisms and abbreviations and many prescribe overlaps and alternate text flow for some morphemes or phonemes.

Character and Block List

Supplemental Punctuation 2E00-2E7F

2E00..2E31;Existing Allocation
 2E32;STENOGRAPHIC PERIOD;Po;0;ON;;;;;N;;;;;

Duployan Shorthands and Chinook 16C00-16C9F

16C00..16C9F; DUPLOYAN SHORTHANDS AND CHINOOK
 16C00;DUPLOYAN LETTER H;Lo;0;L;;;;;N;;;;;
 16C01;DUPLOYAN LETTER P;Lo;0;L;;;;;N;;;;;
 16C02;DUPLOYAN LETTER T;Lo;0;L;;;;;N;;;;;
 16C03;DUPLOYAN LETTER F;Lo;0;L;;;;;N;;;;;
 16C04;DUPLOYAN LETTER K;Lo;0;L;;;;;N;;;;;
 16C05;DUPLOYAN LETTER L;Lo;0;L;;;;;N;;;;;
 16C06;DUPLOYAN LETTER M;Lo;0;L;;;;;N;;;;;
 16C07;DUPLOYAN LETTER N;Lo;0;L;;;;;N;;;;;
 16C08;DUPLOYAN LETTER J;Lo;0;L;;;;;N;;;;;
 16C09;DUPLOYAN LETTER S;Lo;0;L;;;;;N;;;;;
 16C0A;DUPLOYAN LETTER O;Lo;0;L;;;;;N;;;;;
 16C0B;DUPLOYAN LETTER A;Lo;0;L;;;;;N;;;;;
 16C0C;DUPLOYAN LETTER I;Lo;0;L;;;;;N;;;;;
 16C0D;DUPLOYAN LETTER U;Lo;0;L;;;;;N;;;;;
 16C0E;DUPLOYAN LETTER OU;Lo;0;L;;;;;N;;;;;
 16C0F;DUPLOYAN LETTER OW;Lo;0;L;;;;;N;;;;;
 16C10;DUPLOYAN LETTER X;Lo;0;L;;;;;N;;;;;
 16C11;DUPLOYAN LETTER B;Lo;0;L;;;;;N;;;;;
 16C12;DUPLOYAN LETTER D;Lo;0;L;;;;;N;;;;;
 16C13;DUPLOYAN LETTER V;Lo;0;L;;;;;N;;;;;
 16C14;DUPLOYAN LETTER G;Lo;0;L;;;;;N;;;;;
 16C15;DUPLOYAN LETTER R;Lo;0;L;;;;;N;;;;;
 16C16;DUPLOYAN LETTER VOCALIC M;Lo;0;L;;;;;N;;;;;
 16C18;DUPLOYAN LETTER NASAL I;Lo;0;L;;;;;N;;;;;
 16C19;DUPLOYAN LETTER NASAL U;Lo;0;L;;;;;N;;;;;
 16C1A;DUPLOYAN LETTER NASAL O;Lo;0;L;;;;;N;;;;;
 16C1B;DUPLOYAN LETTER NASAL A;Lo;0;L;;;;;N;;;;;
 16C1C;DUPLOYAN LETTER E;Lo;0;L;;;;;N;;;;;
 16C1D;DUPLOYAN LETTER EU;Lo;0;L;;;;;N;;;;;
 16C1E;DUPLOYAN LETTER ROMANIAN I;Lo;0;L;;;;;N;;;;;
 16C1F;DUPLOYAN LETTER ROMANIAN U;Lo;0;L;;;;;N;;;;;
 16C20;DUPLOYAN LETTER U N;Lo;0;L;;;;;N;;;;;
 16C21;DUPLOYAN LETTER P N;Lo;0;L;;;;;N;;;;;
 16C22;DUPLOYAN LETTER D S;Lo;0;L;;;;;N;;;;;
 16C23;DUPLOYAN LETTER F N;Lo;0;L;;;;;N;;;;;
 16C24;DUPLOYAN LETTER K M;Lo;0;L;;;;;N;;;;;
 16C25;DUPLOYAN LETTER R S;Lo;0;L;;;;;N;;;;;
 16C26;DUPLOYAN LETTER M S;Lo;0;L;;;;;N;;;;;
 16C27;DUPLOYAN LETTER N S;Lo;0;L;;;;;N;;;;;
 16C28;DUPLOYAN LETTER J S;Lo;0;L;;;;;N;;;;;
 16C29;DUPLOYAN LETTER S S;Lo;0;L;;;;;N;;;;;
 16C2A;DUPLOYAN AFFIX HIGH ACUTE ARC;Lo;0;L;;;;;N;;;;;
 16C2B;DUPLOYAN AFFIX HIGH GRAVE ARC;Lo;0;L;;;;;N;;;;;
 16C2C;DUPLOYAN AFFIX HIGH DOT;Lo;0;L;;;;;N;;;;;
 16C2D;DUPLOYAN AFFIX HIGH CIRCLE;Lo;0;L;;;;;N;;;;;
 16C2E;DUPLOYAN AFFIX HIGH LINE;Lo;0;L;;;;;N;;;;;
 16C2F;DUPLOYAN AFFIX HIGH WAVE;Lo;0;L;;;;;N;;;;;
 16C30;DUPLOYAN LETTER J N;Lo;0;L;;;;;N;;;;;
 16C31;DUPLOYAN LETTER J N S;Lo;0;L;;;;;N;;;;;
 16C32;DUPLOYAN LETTER M N;Lo;0;L;;;;;N;;;;;
 16C33;DUPLOYAN LETTER N M;Lo;0;L;;;;;N;;;;;
 16C34;DUPLOYAN LETTER J M;Lo;0;L;;;;;N;;;;;
 16C35;DUPLOYAN LETTER S J;Lo;0;L;;;;;N;;;;;
 16C36;DUPLOYAN LETTER M N S;Lo;0;L;;;;;N;;;;;
 16C37;DUPLOYAN LETTER N M S;Lo;0;L;;;;;N;;;;;
 16C38;DUPLOYAN LETTER J M S;Lo;0;L;;;;;N;;;;;
 16C39;DUPLOYAN LETTER S J S;Lo;0;L;;;;;N;;;;;
 16C3A;DUPLOYAN AFFIX LOW ACUTE ARC;Lo;0;L;;;;;N;;;;;
 16C3B;DUPLOYAN AFFIX LOW GRAVE ARC;Lo;0;L;;;;;N;;;;;
 16C3C;DUPLOYAN AFFIX LOW DOT;Lo;0;L;;;;;N;;;;;
 16C3D;DUPLOYAN AFFIX LOW CIRCLE;Lo;0;L;;;;;N;;;;;
 16C3E;DUPLOYAN AFFIX LOW LINE;Lo;0;L;;;;;N;;;;;
 16C3F;DUPLOYAN AFFIX LOW WAVE;Lo;0;L;;;;;N;;;;;
 16C40;DUPLOYAN AFFIX ATTACHED SECANT;Lo;0;L;;;;;N;;;;;
 16C41;DUPLOYAN AFFIX ATTACHED TANGENT;Lo;0;L;;;;;N;;;;;
 16C42;DUPLOYAN AFFIX ATTACHED TAIL;Lo;0;L;;;;;N;;;;;
 16C43;DUPLOYAN AFFIX ATTACHED E HOOK;Lo;0;L;;;;;N;;;;;
 16C44;DUPLOYAN AFFIX ATTACHED I HOOK;Lo;0;L;;;;;N;;;;;

16C46;DUPLOYAN LETTER AOU;Lo;0;L;;;;;N;;;;;
 16C47;DUPLOYAN LETTER OA;Lo;0;L;;;;;N;;;;;
 16C48;DUPLOYAN LETTER J S WITH DOT;Lo;0;L;;;;;N;;;;;
 16C49;DUPLOYAN LETTER S WITH DOT BELOW;Lo;0;L;;;;;N;;;;;
 16C4A;DUPLOYAN LETTER SHORT I;Lo;0;L;;;;;N;;;;;
 16C4B;DUPLOYAN LETTER EE;Lo;0;L;;;;;N;;;;;
 16C4C;DUPLOYAN LETTER IE;Lo;0;L;;;;;N;;;;;
 16C4D;DUPLOYAN LETTER UI;Lo;0;L;;;;;N;;;;;
 16C4E;DUPLOYAN LETTER YE;Lo;0;L;;;;;N;;;;;
 16C4F;DUPLOYAN DOUBLE MARK;Mn;1;L;;;;;N;;;;;
 16C50;DUPLOYAN AFFIX LOW ARROW;Lo;0;L;;;;;N;;;;;
 16C51;DUPLOYAN AFFIX ATTACHED TANGENT HOOK;Lo;0;L;;;;;N;;;;;
 16C55;DUPLOYAN LETTER J WITH DOTS INSIDE AND ABOVE;Lo;0;L;;;;;N;;;;;
 16C56;DUPLOYAN LETTER M WITH DOT;Lo;0;L;;;;;N;;;;;
 16C57;DUPLOYAN LETTER N WITH DOT;Lo;0;L;;;;;N;;;;;
 16C58;DUPLOYAN LETTER J WITH DOT;Lo;0;L;;;;;N;;;;;
 16C59;DUPLOYAN LETTER S WITH DOT;Lo;0;L;;;;;N;;;;;
 16C5A;DUPLOYAN LETTER WO;Lo;0;L;;;;;N;;;;;
 16C5B;DUPLOYAN LETTER WA;Lo;0;L;;;;;N;;;;;
 16C5C;DUPLOYAN LETTER WI;Lo;0;L;;;;;N;;;;;
 16C5D;DUPLOYAN LETTER WEI;Lo;0;L;;;;;N;;;;;
 16C5F;DUPLOYAN LETTER WOW;Lo;0;L;;;;;N;;;;;
 16C60;DUPLOYAN LETTER XW;Lo;0;L;;;;;N;;;;;
 16C61;DUPLOYAN LETTER TH;Lo;0;L;;;;;N;;;;;
 16C62;DUPLOYAN LETTER DH;Lo;0;L;;;;;N;;;;;
 16C63;DUPLOYAN LETTER SLOAN DH;Lo;0;L;;;;;N;;;;;
 16C66;DUPLOYAN LETTER SLOAN J;Lo;0;L;;;;;N;;;;;
 16C67;DUPLOYAN LETTER KK;Lo;0;L;;;;;N;;;;;
 16C68;DUPLOYAN LETTER HL;Lo;0;L;;;;;N;;;;;
 16C69;DUPLOYAN LETTER LH;Lo;0;L;;;;;N;;;;;
 16C6A;DUPLOYAN LETTER RH;Lo;0;L;;;;;N;;;;;
 16C6E;DUPLOYAN SIGN O WITH CROSS;So;0;L;;;;;N;;;;;
 16C6F;DUPLOYAN PUNCTUATION CHINOOK FULL STOP;Po;0;L;;;;;N;;;;;
 16C70;DUPLOYAN LETTER W;Lo;0;L;;;;;N;;;;;
 16C71;DUPLOYAN LETTER LONG U;Lo;0;L;;;;;N;;;;;
 16C72;DUPLOYAN LETTER UH;Lo;0;L;;;;;N;;;;;
 16C73;DUPLOYAN LETTER OOH;Lo;0;L;;;;;N;;;;;
 16C74;DUPLOYAN LETTER SLOAN U;Lo;0;L;;;;;N;;;;;
 16C75;DUPLOYAN LETTER SLOAN OW;Lo;0;L;;;;;N;;;;;
 16C76;DUPLOYAN LETTER SLOAN EH;Lo;0;L;;;;;N;;;;;
 16C77;DUPLOYAN LETTER SLOAN EE;Lo;0;L;;;;;N;;;;;
 16C78;DUPLOYAN LETTER LONG I;Lo;0;L;;;;;N;;;;;
 16C7A;DUPLOYAN LETTER PERNIN AN;Lo;0;L;;;;;N;;;;;
 16C7B;DUPLOYAN LETTER PERNIN AM;Lo;0;L;;;;;N;;;;;
 16C7C;DUPLOYAN LETTER SLOAN AN;Lo;0;L;;;;;N;;;;;
 16C7D;DUPLOYAN LETTER SLAON EN;Lo;0;L;;;;;N;;;;;
 16C7E;DUPLOYAN LETTER SLOAN ON;Lo;0;L;;;;;N;;;;;
 16C7F;DUPLOYAN COMBINING R;Lm;0;L;;;;;N;;;;;
 16C80;DUPLOYAN AFFIX LOW VERTICAL SECANT;Lo;0;L;;;;;N;;;;;
 16C81;DUPLOYAN AFFIX MID VERTICAL SECANT;Lo;0;L;;;;;N;;;;;
 16C82;DUPLOYAN AFFIX HIGH VERTICAL SECANT;Lo;0;L;;;;;N;;;;;
 16C83;DUPLOYAN AFFIX HIGH LONG GRAVE;Lo;0;L;;;;;N;;;;;
 16C88;DUPLOYAN LETTER S T;Lo;0;L;;;;;N;;;;;
 16C89;DUPLOYAN LETTER S T R;Lo;0;L;;;;;N;;;;;
 16C8A;DUPLOYAN LETTER S P;Lo;0;L;;;;;N;;;;;
 16C8B;DUPLOYAN LETTER S P R;Lo;0;L;;;;;N;;;;;
 16C8C;DUPLOYAN LETTER T S;Lo;0;L;;;;;N;;;;;
 16C8D;DUPLOYAN LETTER T R S;Lo;0;L;;;;;N;;;;;
 16C8E;DUPLOYAN LETTER WH;Lo;0;L;;;;;N;;;;;
 16C8F;DUPLOYAN LETTER W R;Lo;0;L;;;;;N;;;;;
 16C90;DUPLOYAN AFFIX LEFT HORIZONTAL SECANT;Lo;0;L;;;;;N;;;;;
 16C91;DUPLOYAN AFFIX MID HORIZONTAL SECANT;Lo;0;L;;;;;N;;;;;
 16C92;DUPLOYAN AFFIX RIGHT HORIZONTAL SECANT;Lo;0;L;;;;;N;;;;;
 16C93;DUPLOYAN AFFIX LOW LONG GRAVE;Lo;0;L;;;;;N;;;;;
 16C9A;DUPLOYAN LETTER S N;Lo;0;L;;;;;N;;;;;
 16C9B;DUPLOYAN LETTER S M;Lo;0;L;;;;;N;;;;;
 16C9C;DUPLOYAN LETTER K R S;Lo;0;L;;;;;N;;;;;
 16C9D;DUPLOYAN LETTER G R S;Lo;0;L;;;;;N;;;;;
 16C9E;DUPLOYAN LETTER S K;Lo;0;L;;;;;N;;;;;
 16C9F;DUPLOYAN LETTER S K R;Lo;0;L;;;;;N;;;;;

Shorthand Controls 16CF0-16CFF

16CF0..16CFF; SHORTHAND CONTROLS
 16CF0;SHORTHAND CONTROL LETTER OVERLAP;Cf;0;BN;;;;;N;;;;;
 16CF1;SHORTHAND CONTROL WORD OVERLAP;Cf;0;BN;;;;;N;;;;;
 16CF2;SHORTHAND CONTROL DOWN;Cf;0;BN;;;;;N;;;;;



Character names. For naming purposes, the Duployan Shorthands and Chinook script have two functionally distinct sets of characters. The first set consists of letters and letter based signs that generally interact cursorily with each other, with the exception of a few spacing characters. These characters are mostly a letter with a general, orthography-independent phonetic value. The second set consists of affix signs that can be attached/overlapping or sit above or below the adjacent characters at the beginning and end of words, and the word signs. For the most part these characters have unrelated values and positions in the different Duployan orthographies - appearing as a prefix with one value in one orthography, a suffix in another, and as an unrelated suffix and prefix in another. The third set of characters are the dotted arc consonants. The first set have character names that indicate their primary phonetic value, while the second set are described graphically.

Support, Funding, and Thanks. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities to the Universal Scripts Project (as part of the [Script Encoding Initiative, UC Berkeley](#)). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

This proposal has also been materially supported by the facilities and resources of [Michael Everson's Everttype](#), Justin Cassidy, [the United States' Library of Congress](#), and [the Timberland Regional Library](#), and the views, findings, conclusions and recommendations expressed herein do not necessarily reflect those of said organizations.

Special thanks to Laurențiu Iancu at Microsoft, Eric Muller at Adobe, Dave Robertson at University of Victoria (BC), and Michael Everson, for tips, information, documentation, and other intellectual support in this project. Thanks to Ken Whistler (Sybase), Rick McGowan (Unicode), Deborah Anderson (UC Berkeley), Justin Cassidy, Micah Ferrell, William Poser, and Asmus Freytag for feedback and logistical support. This proposal would not be nearly as complete or as well written without the involvement of every one of these people.



Character Ordering and Roadmap to the Duployan Shorthand and Chinook character block

Ordering of the characters in the Duployean-based scripts is generally undefined - most cite in Latin alphabetical order - so allocation order in the Duployan Shorthands and Chinook character block is revisable up to inclusion in the standard. The currently proposed allocation ordering and its basis is as follows: Columns 0 and 1 are occupied mostly by characters that make up the core inventory of the different Duployan shorthands and the Chinook script. Most Duployan orthographies will use almost all of the characters in these two columns, and optimization algorithms may be able to take advantage of the fact that these characters constitute the vast majority found in texts written in any of the Duployan orthographies. Columns 2 and 3 contain the French Duployean compound letters and affixes. Several of these characters are also in the core and supplementary inventory of the other orthographies. Column 4 is a mixture of diphthongs, affixes, and letters for several orthographies. Columns 5 and 6 contain contain the Chinook compound letters, and similarly constructed letters and signs from the Romanian shorthand and Sloan systems, the Chinook Full Stop and Likalisti signs, and a couple Romanian affixes. Column 7 contains vowels for the English Duployan systems, ending with the Sloan "R", which is a combining character that functions as a format. Columns 8 and 9 have two parts, each beginning with Pernin affixes, and containing the quarter-circle arcs for the English orthographies. This allocation provides for all characters needed for French Duployean in columns 0-4, Romanian shorthand in 0-5, Chinook in 0-6, with Pernin, Sloan, and Perrault using all columns through 9. Twenty code points have been left unallocated for any additions needed for the Brandt and Galloway systems or the Spanish and German adaptations of Duployean, as no documentation on the workings of these systems has been located.

Collation

Information on collation of Duployan scripts is generally ambiguous and contradictory. Many dictionaries and primers simply cite in that orthography's Latin alphabetical order with no attempt made at native collation. [Other sources](#) group words by novel alphabetization, no more or less canonical than any other. The Romanian "Curs de Stenografie" does make an effort at native collation, starting vowels, and then in the general order of the consonants in this allocation. The collation algorithm prescribed herein is based on principles derived independently from the Romanian, but results in a similar order.

The most logical collation, given the structure of the script, is to collate by general shape, which places primacy on the consonants which, being invariant, tend to determine the shape of a word. The only exception is word-initial vowels, which should be collated as if they begin with a null initial consonant, along with the non-joining letters. The order of the collation is consonants (H,P,T,F,K,L,M,N,J,S,combined consonants), then vowels (O,A,I,U,Ou,Ow,Nasals), Affixes (attached, high, low), and finally signs. All variants should be collated directly after their base letters, with voiced consonants and their variants after the last unvoiced variants, eg initial vowels, H, X, XW ... K, Kh, Kk, G, Sloan J, KM, L, Lh, hL, R, Rh, RS, M, MN, MS, MNS, M w/ dot, N, NM ... A, Wa, Oa, I, E, Wi, Wei, short I, Ee, Ie, Ui, Ye ... low wave affix, low arrow, low long grave, O w/ cross. This collation order corresponds significantly with the order of words in the Romanian "Curs de Stenografie", except that F/V comes after K/G and A comes before O.



Principles of the Duployan Shorthands and Chinook scripts

Rendering Duployan Characters. Duployan characters, like characters in most shorthand scripts, can cursively connect, combine, and change shape depending on their context. Its appearance is affected by its ordering with respect to other characters, the font used to render the character, and the application or system environment. These variables can cause the appearance of Duployan and Chinook characters to differ from their nominal glyphs (used in the code charts). Duployan and Chinook characters are classified as default joining, except for the high and low affixes and where otherwise noted in the chart. Characters marked as non-joining, and any characters from other blocks, except the Zero Width Joiner (U+200D) and the Shorthand Controls (U+xF0-U+xF2), are non-joining to Duployan and Chinook characters by default.

Invariant letters. The majority of characters in the Duployan shorthands and Chinook scripts are invariant. They have a static shape, orientation, and stroke direction, and the set of invariant characters is almost completely contiguous with the consonants. Each invariant has a size (small, medium, large), a shape (lines, half and quarter arcs), a static orientation, an inherent stroke direction, and many have derived and compound variants with markings (crosses or dots). They will usually cursively connect (end of preceding character's stroke to beginning of next) to the preceding and following characters, but can also overlap with a following character through the use of shorthand controls. A few invariant letters and all of the high and low affixes are classified as *non-joining characters* that interact typographically with adjacent characters like a word or text break, and do not have a stroke direction.

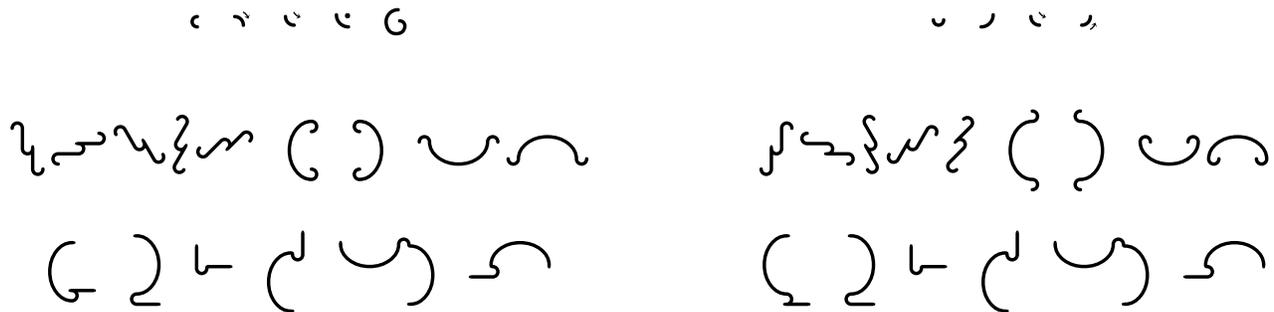
It can be assumed in the following that the letters D, D-S, TH, and DH have the same cursive, overlapping, and other connecting properties as the character on which it is based, ie T. Likewise, variations of N - N-S, N-M, N-M-S, and Ng - connect like an N, and so on. The invariants can be generally classified as *P-type* (N-S stroke direction), *T-type* (W-E), *F-type*(NW-SE), *K-type*(NE-SW), *L-type*(SW-NE), *M-type*(N-S, crescent shaped), *N-type*(N-S, crescent), *J-type*(W-E, crescent), and *S-types*(W-E, crescent), and combined consonants (see code chart). Furthermore, the P,T,F,K, and L-collectively constitute the *Line consonants*, and the M,N,J, and S-types, as well as the combined consonants, are *arc consonants*.

Orienting characters. Many vowel letters have a consistent shape, but rotate to align with the preceding character and mirror to allow the following character to attach without crossing the vowel or preceding character. When adjacent a non-joining character, or a joining character that allows either orientation, these orienting vowels will rotate to align with the adjacent (or preceding) joining character, and mirror right/up or left/down based on their identity as a *primary orienting* or *secondary orienting* vowel. Likewise, when adjacent two similar type characters, they will align with the preceding character and mirror according to their orientation. Primary orientation indicates an affinity for proceeding towards the right, and down when ambiguous. Conversely, secondary orientation is left/up. Many orienting vowels come in pairs, with opposite orientations but the same shape.

Table 1: Comparison of Primary and Secondary Orienting Vowels

Primary (right/up) Orienting Vowels-

Secondary (left/down) Orienting Vowels -



Related to the orienting vowels and invariant letters are the attached affixes. Many of these, noted in the charts with "dots [to] show position on and relative orientation to base glyph", act as spacing and non-spacing marks that do not effect joining of adjacent characters. Some, noted in the charts with "dots [to] show position on base glyph", are non-spacing invariant marks.

Circle vowels. The most commonly encountered vowel letters are classified as circle vowels. These vowels connect to preceding and following characters, with the adjacent characters entering the circle vowel at a tangent, and most (except Ou U+x0E) exiting the vowel shape at a tangent. The circle vowels will take partial contextual forms, allowing the adjacent characters to cross, thus implicitly completing the circular form of the vowel.

Circle vowels not preceded by a joining letter have a clockwise stroke direction before line consonants and will lie inside the arc of an arc consonant. Circle vowels not followed by a joining character will again sit inside the arc of an arc consonant, as if followed by a T-type if following a line consonant, and above the end of a T-type consonant. Circle vowels adjacent two line consonants will lie outside the angle created by the intersection of the two lines. When adjacent same type line consonants, they will again lie as if followed by a T-type. When adjacent an arc consonant and another invariant, the circle vowel will follow the angle rule as given above, and when the adjacent characters do not present an angle, the circle vowel will lie in the same position as if not followed by a joining character.

The Duployan Letter Sloan Ow (U+x75) and, in the Pernin orthography, a discretionary ligature form of a circle vowel followed by L (U+x05) are classified as *reverse circle vowels*. These reverse circle vowels are opposite a regular circle vowel, ie they have a withershins stroke direction, will lie outside of arc vowels, inside the angle of two line consonants, &c.

Table 2: Circle Vowels and Reverse Circle Vowels

Circle Vowels

Reverse Circle Vowels



Nasal vowels are the most diverse characters in terms of their visual representation. A fully implemented typeface will allow for three different orthography-specific renderings of the four basic nasal vowels (U+x18-U+x1B). When adjacent two consonants, the nasal vowels will render as a diacritic outside the angle of the adjacent characters, following the rules for circle vowels adjacent two characters, above. When followed by ZWJ (U+200D) + a joining character or any non-joining character, the nasal vowel will render as a primary or secondary orienting vowel, connected to the preceding joining character. Likewise, when following a joining character + ZWJ (U+200D), the nasal vowel will render as a primary or secondary orienting vowel, connected to the following joining character. The Duployan Letter Vocalic M (U+x16) is always a primary orienting vowel. When bracketed by Zero Width Joiners, nasal vowels will render as combining invariant characters as per the nominal glyph images. ZWNJ (U+200C) can be used when the orienting or invariant nasal vowel is not to be connected to an adjacent joining character. The Pernin and Sloan nasal vowels (U+x7A-U+x7E) are always invariant. The orthography of the Romanian stenography uses the two U arc vowels (U+x0D, U+x1D) as nasals, however the Romanian stenography treats nasals as orienting vowels, and no marking is needed for proper rendering.

Table 3: Nasal Vowels

\ + ʀ + -	→	ʀ	F + An + T
\ + ʀ + ZWJ + -	→	ʀ	F + An_j + T
\ + ʀ + ZWNJ + -	→	ʀ-	F + An_{nj} + T
\ + ʁ + ZWNJ + -	→	ʁ-	F + On_{nj} + T
- + ZWJ + ʀ + -	→	-ʀ	T + _jAn + T
- + ZWJ + ʁ + -	→	-ʁ	T + _jOn + T
\ + ZWJ + ʀ + ZWJ + -	→	ʀ	F + An_i + T

P.S. The logic behind the prescribed use of ZWJ/ZWNJ is that it deprives the surrounding context from the nasal vowel. They could be replaced by any non-joining character and result in the same rendering of the nasal. The Joiner/Non-Joiner choice merely defines whether the adjacent characters will cursively attach or not.

Compound vowels. The default rendering of compound vowel sequences depends on the nature of the vowels involved. Most orthographies prefer a new character to the compounding of circle vowels - Duployan letters AOU, OA, and the Compound W vowels, though see ligatures, below for exceptions in Romanian. However, compounding that represents each member is regularly encountered in sequences involving orienting vowels combined with a circle vowel or other orienting vowels. As a rule, circle vowels act as if an adjacent orienting vowel were a line consonant whose orientation is determined by the adjacent joining character. The entire sequence should be rendered as if it were an orienting vowel, although advantage can be taken of an initial or final circle vowel sharing the adjacent joiner. These vowel sequences have primary or secondary orientation determined usually by the first character of the sequence, but the last character when not preceded by a joining character. When the compound sequence is not adjacent any joining characters, default rendering is along a horizontal mid-line.

Table 4: Compound vowels

o + e	→	oe	A + I
o + u	→	ou	A + E

o + c + -	→	Ɱ	A + I + T
o + v + -	→	Ɱ	A + E + T
c + o + -	→	Ɱ	I + A + T
l + o + c	→	Ɱ	P + A + I
l + o + c + -	→	Ɱ	P + A + I + T
l + o + c + (→	Ɱ	P + A + I + M
v + c	→	Ɱ	E + I
c + v	→	Ɱ	I + E
o + c	→	Ɱ	O + I
o + G	→	Ɱ	A + Rom U
c + o + G	→	Ɱ	I + A + Rom U

Ligatures, Allographs, and Alternates. Ligaturing behaviour is fairly limited in the Duployan orthographies, especially in comparison with other cursive scripts like Arabic and Devanagari. All ligatures found are discretionary, but can be requested in plain text by the use of Zero Width Joiner (U+200D). Unless indicated otherwise, the default is to not render the ligated forms of character sequences.

The first discretionary ligature is found in the Pernin orthography whenever a circle vowel precedes the Pernin R (Duployan letter L, U+x05), and is NOT followed by another circle vowel. The ligature form is an identically sized reverse circle vowel (see Circle Vowels, above).

Pernin also makes use of ligatures for repeated consonants, formed by reducing the second consonant to a small small blot (in writing, caused by increasing pen or pencil pressure) at the end of the first character's stroke. This applies to both identical and similar consonants, with the first consonant represented by its full form, eg. T+T = T+D = T+Th.

Pernin prescribes a "slight upward tick inclining to the left" for an L (U+x05, Pernin R) following R (U+x15, Pernin L), and one "to the right" for an R after L. This upward tick can also sometimes be found, only at word end, following other consonants. These ticks should be treated as an optional contextual variant of the Duployan Letter L and the Duployan Letter R.

Most orthographies have some means of indicating the junction of two same type line consonants. Usually, this comes in the form of a slight (\leq line width) jog at the intersection, a short cross-tick, or even an angle change. Default rendering is the jog.

The Duployan Letter W (U+x70) is the most variable letter in the Duployan scripts. In the Sloan and Perrault orthographies, it is a full quarter arc, written NE-SW, 12 o'clock to 9 o'clock. On the other hand, in Pernin, it is closer to a one-sixth arc, starting closer to the 11 o'clock position, though still roughly the same length arc (larger diameter) than the Sloan/Perrault variety. Following K and G (U+x04, U+x14), the Duployan Letter W can take the form of a hook - Perrault tending a bit more wave-like than Pernin. Sloan prescribes combined consonants instead of these sequences.

The Romanian orthography prescribes contextual forms for the Romanian U character (U+x1F). The nominal form given in the code charts is for isolated, initial, and final contexts. When medial, it takes the form of Duployan Letter Ow (U+x0F). Romanian also prescribes a ligature for non-medial forms of the sequence O + Romanian U (U+x0A + U+x1F), in the form of an elongated, oval shaped, plain circle vowel. Medially, these two will sit on opposite sides of the intersection of the two adjacent joining characters - Romanian U again appearing like Ow. Likewise, in sequences with other vowels, Romanian U can appear in diminished form (see Compound Vowels, above).

Lastly, the Sloan Combining R (U+x7F) does not have a visual form of its own, but should cause the previous character to be rendered as a thick variant. The Duployan Letter R (U+x15) can not substitute a ligature behavior for the Sloan Combining R. This is mandatory default rendering behaviour.

Table 5: Ligatures, Allographs, & Alternates

Discretionary features

+	→		B + P-type
/ + /	→	/	R + L
/ + /	→	/	L + R
- + /	→	⌋	T + L
⌒ 	→	⌒	W variant

Default and obligatory features

/ + ⌒	→	⌒	K + W
+ G + -	→	⌒	P + Rom U + T
+ O + G	→	⌒	B + O + Rom U
+ O + G + -	→	⌒	B + O + Rom U + D
- + -	→	⌒	T + D
\ + r + °	→	⌒	F_R + A

Joined text. The most common form of character interaction is that of the cursive connection. The termination of a character stroke leads directly into the beginning of the next character. Vowel signs are the same, joining at a tangent, but follow the dynamic shaping discussed above. Non-joining characters - any character from other scripts, and those found in Duployan - have a small intervening space.

Unjoined text. The Duployan script has a cursive conjoining property similar to Arabic (without the alternate forms) that is effected by the use of the Zero Width Non-Joiner (ZWNJ, U+200C). ZWNJ encodes a break within a word, turning an otherwise joining character into a non-joining character. This break is usually only found at nominally syllabic boundaries in Chinook texts, and in the Duployan shorthands where a separated letter or letters indicates a common prefix or suffix. This break is smaller than a word space, in some instances involving negative kerning, and is not a word break. ZWNJ and Zero Width Joiner (U+200C) will also change the behaviour of the nasal vowels (see nasal vowels, above). Except for Romanian U and its compounds, there are no medial forms that require ZWJ to display in isolation.

Overlapping text. The use of overlapping letters to indicate abbreviations and initialisms is well-attested in many systems of shorthand. As such, the current proposal includes allocations for a separate block containing shorthand control characters, placed between any characters with non-default text flow. Included are two overlap control characters: the first (U+xF0) indicating a single *letter overlap*, with the text continuing to flow as if that overlapping character did not exist, and the second (U+xF1) indicating a *word overlap* where the text flow proceeds from the overlapping character. This behaviour seems to be limited to consonants being overlapped by consonants, circle vowels, or orienting vowels.

The overlapping behavior in Duployan shorthands and Chinook is fairly straightforward: for two line consonants, two arc consonants, or a vowel overlapping any consonant, the two characters overlap at approximately 3/5 along the stroke of the first consonant and 2/5 along the stroke of a second character. For arc consonants and line consonants overlapping, the arc consonant is split into the first and second half of the arc, an arc overlapping a line taking place in the first half of the arc, line over arc on the second half. The line consonant will overlap the arc at a perpendicular angle, or as close as possible, never beyond the middle of the arc, nor past an end.

It is unknown if or how M-type and N-type or J-type and S-type would overlap until such a time as examples of this occurrence are documented. Default rendering should indicate the overlap, either displaying control characters or offsetting the second consonant below and to the right. Same type line consonants similarly will not overlap, the default rendering again with visible controls, or with a low-right offset; this applies to L-type and K-type consonants, as well.

As indicated above, the flow of text continues either with the first character in the case of U+xF0, or with the second in the case of U+xF1. An overlapping letter can also take another overlapping letter before returning the text flow back to the original text flow. Also, in the Romanian shorthand, long line consonants (U+x11-U+x15) can take two overlapping characters, indicated by two Letter-Overlap control characters (U+xF0 + U+xF0) followed by the two overlapping characters. With double overlaps, the first overlapping character overlaps at approximately 1/3 of the stroke length of the base character, the second at ~ 2/3.

Under text. The Romanian shorthand prescribes that a certain set of word endings be indicated by letters following not in the default direction of text flow - to the right, but below the word. As such, a shorthand control has been defined (U+xF2) that indicates the following character should be rendered below the previous character, with any subsequent joined characters proceeding relative to the lowered glyph. The lowered text does not proceed beyond characters joined to the first lowered character.

Table 6: Text flow

Joined Text

| + ◡ + • → |◡• **PJH**

— + / + † → — / † **DKX**

Unjoined Text

| + ZW
NJ + ◡ + ZW
NJ + • → |◡• **P.J.H**

— + ZW
NJ + / + ZW
NJ + † → — / † **D.K.X**

Letter Overlaps

|-\// + [X] + C → ~~X~~C Line_XS

C + [X] + |-\// → C~~X~~ S_XLine

| + [X] + / → |X B_XR

- + [X] + / → -X D_XG

\ + [X] + - → \X V_XD

/ + [X] + | → /X G_XB

/ + [X] + \ → X R_XV

(+ [X] + (→ (X M_XM

(+ [X] + C → (X M_XS

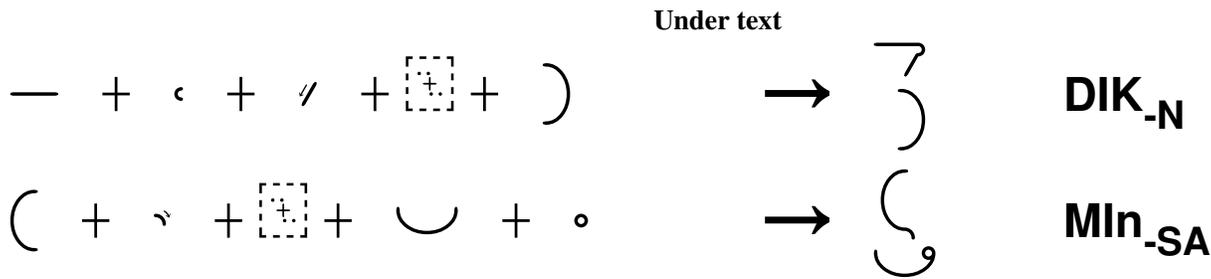
(+ [X] + C → (X M_XJ

Word & Double Overlaps

\ + • + - + [X] → ~~X~~ KAT_{X+}KAT
 + \ + • + -

C + [X] + | + [X] + C → C(X) S_XB_XJ

- + [X] + [X] + • + / + C → -X(A + K) + Un



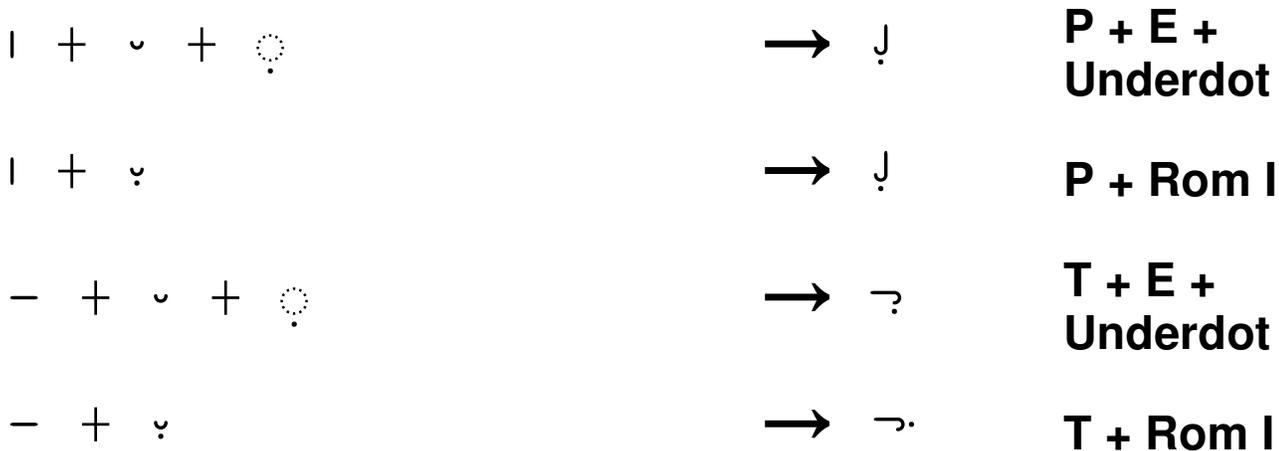
Combining diacritical marks on vowels. Duployéan orthographies use combining diacritical marks on vowel characters. These diacritics include acute, grave, breve, macron, under macron, over dot, under dot, diaeresis, under diaeresis, &c. They can appear on orienting vowels, circle vowels, and nasal vowels (On, and An). Although there are several precomposed vowel letters with diacritics included in the allocation, these are not decomposable as a combining sequence, as the diacritic marks change position along with their "base" orienting vowel. Combining diacritics indicate vowels with diacritics that consistently appear above or below the base character, no matter the context.

Affixes. Except for Chinook, every Duployan orthography makes extensive use of a set of marks and letters to symbolize lexical affixes. First, the high and low Duployan affixes (U+x2a-2f, U+x3a-3e, U+x50, U+x83, U+x93) act much like spacing characters - the marks are written above or below the beginning or final letter of a word, but have an advance width so successive affixes can be agglutinated.

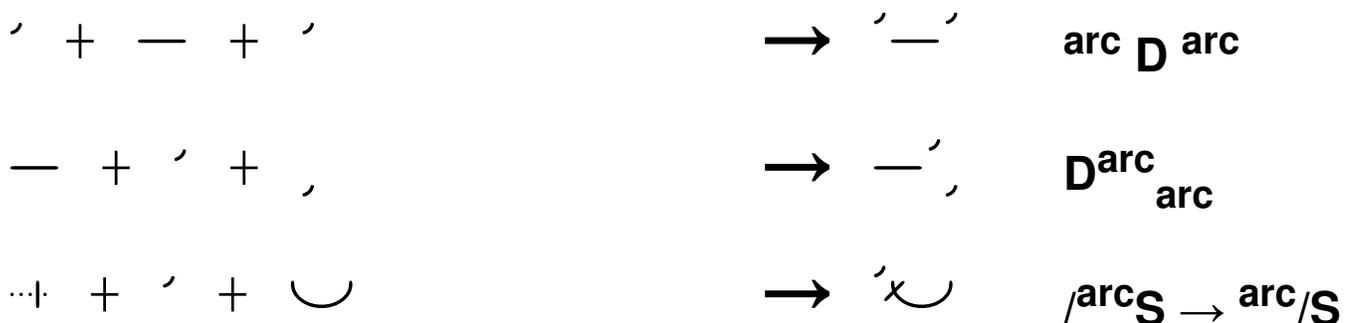
Second, the attached affixes (U+x40, x41, x44, etc.) are the affixes that touch or cross the first or last letter of a word (again for prefixes or suffixes), with the location of crossing (and touching if not evident) symbolized by a dotted line in the charts. The character names list specifies if the character rotates to complement the angle of the base letter, or is consistent. An attached affix will always attach to a letter, never to a high or low affix. Since affixes are encoded logically, the displayed order of attached and unattached affixes may be different from the encoded order.

Third, some orthographies use letters or sequences of letters to indicate affixes, some of which appear similar to the high or low affix signs. As a rule, signs that are similar to a letter, but unmotivated - that is, they don't symbolize a sound of the affix - or if a high and low pair is found in the orthography, are symbolized by the affix signs. Signs that are motivated and aren't paired should be represented by a letter, often separated by ZWNJ (U+200C) from the root, whether the affix usually appears lower or higher than the adjacent character or not.

Table 7: Diacritics and affixes
Diacritics and Precomposed Vowels



Affix Signs and Letters



— + ○ + // + ZW
NJ + (→ ϕ(**DOK.M**

// + - + ZW
NJ + / → L/ **KT.R**

Numbers. Each shorthand system seems to have a distinct means of expressing numbers. Some number systems must utilize markup to represent all aspects of the number system, and as of this time, there is no expectation that a full transcription of all number forms should be representable in plain text. The Chinook number system uses Duployan characters and markup to indicate numbers. The Romanian shorthand and French Duployéan use regular European/Arabic numerals in conjunction with Duployan characters, combining marks, and markup to indicate magnitude and aspect. Sloan and Pernin use markup and non-Duployan characters in conjunction with regular European/Arabic numerals. It is important to note that while Duployan characters do not join to characters from other scripts, they do have relative positions to the European numbers found in the Basic Latin block

Chinook numbers. The Chinook number characters are 1-P, 2-T, 3-F, 4-K, 5-R, 6-M, 7-N, 8-J, 9-S, 0-O, 10-A, 100-Wa, and 1000-enclosing circle handled with markup. The numbers can be indicated Hanzi-style with {P-S} combining with O, A, or Wa to indicate value, although an O, A, or Wa must be preceded by a P to indicate a single hundred or ten, unlike Hanzi numerals. {P-S} connect to O, A, and Wa the same as in text, and may or may not follow the graphic breaking on "syllables" pervasive in Chinook writing. O is used unconnected to indicate a zero or connected for the tens as a following digit zero, while A is used when connecting the tens to a ones digit. The enclosing circle for thousands surrounds the entire group of up to five characters {P-S} Wa {P-S} A/O {P-S}, and can nest inside itself to indicate millions - a separate circle surrounding a following thousands group. Chinook numbers can also be indicated Indian/Arabic style, with the digits 0-9 (O-S) having place value. This is especially common when writing years or when numbering items, as opposed to enumerating them. The digits generally connect cursorily, the same as in Hanzi-style numbering. For most Chinook writers, the numeral forms of M, N, J, and S are about half-size normal, and a fully featured Chinook font will have these as a stylistic alternate. Variation sequences may be defined in future for these small arc consonant-number forms.

Romanian numbers. The Romanian number system uses the European/Arabic numerals to indicate numbers 0-99, with marks to indicate further powers of ten: an overdot (U+0307) for hundreds, a preceding Middle Dot (U+00B7) for thousands, a dot below (U+0323) for millions, and a following Middle Dot for thousand millions. As with most systems using marks to indicate magnitude, these marks can be used in conjunction, e.g. a dot above and dot below for hundred millions. Multiplicative forms (with the prefix *în-*) use the character An (U+x1B) before a number, percentages with Combining Ring Above (U+030A), and grade with the Duployan Affix High Circle (U+x2D). Ordinals are symbolized by a following T (U+x02), while fractions are written numerator over denominator, with no solidus or line. This representation of fractions constitutes a presentation form of already encoded fraction signs or can be explicitly expressed using markup, never with the Shorthand Control Down (U+xF2). Shorthand controls (U+xF0-U+xF2) should never be used to define a relationship between non-shorthand characters.

Pernin numbers. The Pernin number system uses the European/Arabic numerals to write numbers, although periods (U+002E) can be used instead of zeros. An underline (by markup) indicates ordinals (first, second...), while an overline (again) indicates the numerical adverbs (once, twice...). The Pernin system suggests, however, that "when large numbers are to be written ... it is better to indicate ... us[ing] a corresponding shorthand contraction for thousand, million, etc.", such contractions left to the individual.

Sloan numbers. The Sloan number system uses the European/Arabic numerals to write numbers, and can be used for ordinals, iteratives, &c. e.g. 2: two, twice, second, secondly. The shorthand aspect in the Sloan system is the use of an overline, strikethrough, and underline (all represented with markup) for magnitude as follows: Overline: hundreds; Strikethrough: thousands; Underline: millions. Again, these can be used in conjunction with each other to indicate, for example hundred millions with an overline and underline.

French Duployéan numbers. The French Duployéan number system, like the Romanian, uses the European/Arabic numerals with Duployan letters and affixes indicating magnitude and aspect. Magnitude is indicated as follows: Hundreds with an S (U+x09) below the number - Zero Width Non Joiner (U+200D) should be used; Thousands with the Duployan Affix High Dot (U+x2C) following the number; Millions with the Duployan Affix Low Grave Arc (U+x3B) following; and Thousand Millions (Milliards) with an R (U+x15) like a large solidus, again using ZWNJ. As above, these indicators of magnitude can be combined, e.g. a ZWNJ + S and high dot indicating hundred thousands. For ordinals, the Duployan Affix Low Dot (U+x3C) is used following any indications of magnitude; Adverbs with the Duployan High Acute Arc (U+x2A); Approximates (*dizaine, douzaine, &c*) with the Duployan High Grave Arc (U+x2B); Adverbials with the Duployan High Circle (U+x2D); Percents with the Duployan Low Circle (U+x3D) and doubled for Permill. Manuscripts will indicate the numbers 4 and 6 with an underline to distinguish these number forms from the words "quittance" and "mot" to which the regular number forms show affinity; This distinction should be handled with markup or by typeface choice.

References

[Archives of the Kamloops Wawa 1891-1900](#) (subscription required)

Dictionary of the Chinook Jargon, by George Gibbs, Echo Library ISBN 1-40680-924-1

Chinook:.... A History and Dictionary, by Edward Harper Thomas, 1935, Metropolitan Press, Portland, OR

Cours de Sténographie Duployé Fondamentale, by A. Hautefeuille and C. Ramaude

Pernin's Practical Reporter, compiled and published by H. M. Pernin, 1882, O. S. Gulley Printing House, Detroit, MI

Pernin's Universal Phonography, 16th ed, by H. M. Pernin, 1902, Detroit, MI

Curs De Stenografie, publicat de Margareta Sfințescu în Enciclopedia practică a copiilor, Editura Ion Creangă, 1984

Stenographie Integrale, http://www.stenographie.ch/stenographie_integrale.pdf

Modern Shorthand. the Sloan-Duployan Phonographic Instructor, 11th ed, by J. M. Sloan; 1st ed. 1882; Ramsgate, England; St. John's, NL; Brisbane, QLD

Modern Shorthand: the Sloan-Duployan system. Reporters' Rules, by John Mathew Sloan, 1892, London.

The Wawa Shorthand Instructor, first edition, by the Editor of the Kamloops "Wawa", 1896, Kamloops, BC

Perrault-Duployan Complete Elementary Course of Stenography in Six Lessons, Sixth Edition, by Denis R. Perrault, 1918, Montreal.

nouveau site duployé, <http://cf.geocities.com/barouder396/>



ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

Form number: N3702-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)

A. Administrative

1. Title:	<i>Proposal to include Duployan Shorthands and Chinook script in Unicode / ISO-10646.</i>
2. Requester's name:	<i>Van Anderson vanisaac@boil.afraid.org</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>
4. Submission date:	<i>2010-01-16</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>X</i>
(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):	<i>Yes</i>
Proposed name of script:	<i>1) Duployan Shorthands and Chinook 2) Shorthand Controls</i>
b. The proposal is for addition of character(s) to an existing block:	<i>Yes</i>
Name of the existing block:	<i>Supplemental Punctuation</i>
2. Number of characters in proposal:	<i>142 - (1 in Supplemental Punctuation, 3 in Shorthand controls, 138 in Duployan Shorthands and Chinook)</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary	<i>X</i> B.1-Specialized (small collection)
C-Major extinct	E-Minor extinct
F-Archaic Hieroglyphic or Ideographic	G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided?	<i>Yes</i>
a. If YES, are the names in accordance with the "character naming guidelines"	<i>Yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Van Anderson vanisaac@boil.afraid.org</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Van Anderson https://boil.afraid.org/Chinook/DuployanProp.ttf</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes, for some of repertoire</i>
7. Special encoding issue	
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)?	<i>Yes</i>
	<i>Information is included in this document, above.</i>

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	_____	<i>No</i>
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.)? If YES, available relevant documents:	_____	<i>Yes</i>
	<i>Online forums: Forum du petit sténographe (http://forumsteno.vosforums.com/), Chinook Language List (http://listserv.linguistlist.org/archives/chinook.html)</i>	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	_____	<i>Yes</i>
	<i>Script will be used primarily by small community of hobbyists and linguistic/historical scholars, with expected minor utility to legal and government historians, due to extensive usage of Duployan shorthands in Canada.</i>	
4. The context of use for the proposed characters type of use; common or rare? Reference:	_____	<i>rare</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	_____	<i>Yes</i>
	<i>Still in use by small hobbyist community, mostly in France. Scholarly and historical/cultural preservation use.</i>	
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:	_____	<i>No.</i>
	<i>Except for addition to Supplemental Punctuation, characters should be allocated in SMP (Plane 1) as per Roadmap.</i>	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	_____	<i>Yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If Yes, reference:	_____	<i>No</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If Yes, reference:	_____	<i>No</i>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If Yes, reference:	_____	<i>Yes.</i>
	<i>Any similarities in appearance are coincidental.</i>	<i>Yes</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If Yes, reference:	_____	<i>Yes</i>
	<i>Several orthographies use optional combining accents to distinguish similar vowel sounds. Further justification is contained in document, above.</i>	<i>Yes</i>
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If Yes, reference:	_____	<i>No.</i>
	<i>All undocumented composite sequences have trivial rendering.</i>	
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	_____	<i>Yes</i>
	<i>The 3 Shorthand Control characters and Duployan Combining R are discussed in detail, above.</i>	
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If Yes, reference:	_____	<i>No</i>