Non-concatenative morphology and interlinear translations: A Lillooet example

Jan P. van Eijk
First Nations University of Canada

A common way of presenting bilingual Salish-English texts, employed in a number of publications, is the interlinear format, with morpheme-by-morpheme or word-by-word glosses given below the Salish originals. While this format ideally provides an immediate read-off between the Salish forms and their English equivalents, it is not without certain problems, especially where non-concatenative morphology is involved. However, once these problems are recognized, they can be easily resolved with sufficient and consistent attention to grammatical detail.

1 Introduction

Over the years, the Salishanist community (comprising both speakers and linguists, including linguists who have acquired a measure of fluency in the language of their study) has produced a large number of bilingual Salish-English texts. From a rich crop we select the following sample: Davis and Saunders 1980 on Bella Coola (Nuxalk), Kuipers 1967 and 1969 on Squamish (Skwxwu7mesh), Bierwert 1996, and Hess 1995, 1998 and 2006 on Lushootseed, Kuipers 1974 and 1989, and Dixon and Palmantier 1982 on Shuswap (Secwepemctsin), Matthewson 2005 and 2008, and Van Eijk and Williams 1981 on Lillooet (St’át’imcets), Mattina 1985, Mattina and DeSautel 2002, and Lindley and Lyon 2012 on Okanagan, and Vogt 1940 on Kalispel.

Unfortunately, not all Salish languages are so well represented, and for some, especially extinct or near-extinct languages like Pentlatch, Tillamook-Siletz or Quinault, the textual resources are skimpy. Another problem is that of linguistic reliability. While the collections above were all taken down by trained linguists and delivered in either the mostly standardized Amerindianist
transcription or in a phonemic orthography based on the 26-letter Latin stock (or, in the case of Matthewson 2005, in both types of transcription), in other cases we are not so lucky. For example, around 1900 Charles Hill-Tout published a large number of texts in about a dozen Salish languages and dialects, but in a woeful transcription that reflects his lack of proper phonetic training. (The English versions of these texts are available in Maud 1978. Davis 2001, in a remarkable display of linguistic sleuthing, resurrects a Lillooet text collected by Hill-Tout by retranscribing it in the Amerindianist symbols, together with a detailed grammatical analysis and comments on the narrative structure.)

Finally, there is a certain variety in the ways in which these texts, and those in other collections, are presented. While, for example, Vogt 1940, and Van Eijk and Williams 1981, present the texts in both the original languages and their English translations, on opposing pages or on the top vs. bottom of the same page, but without interlinear translations, many if not most sources provide an interlinear morpheme-by-morpheme or word-by-word translation. Although this technique ideally provides an immediate read-off between Salish morphemes or words and their English translations, it is not without its own challenges, which we address in the next two sections.

2 Problems in interlinear translations

Providing a morpheme-by-morpheme translation of the targeted text poses no problem when the morphemes are entered serially from left to right, so that there is no conflict between locus and ordering. A Dutch example will suffice here (PL = plural):

(1) arbeid-er-s
    labour (verb)-person-PL
    ‘Labourers’

However, three issues may complicate this picture. In the first place, morphophonemic changes may obscure the morphological make-up of a word to the point that the underlying morphemes are not, or only barely, recognizable. This problem can be addressed by having the surface forms and
the underlying forms on separate lines, right above each other. Taking a Thompson (Nlaka’pamux) example from Thompson and Thompson 1992 and putting it in the following four-line format we get:

(2)  
\[
\begin{array}{l}
\text{pūnmcip} \\
pūn-min-t-sey-ep \\
\text{find-RLT-TR-1SG/O-2PL/S} \\
\text{‘You folks find me’}
\end{array}
\]

\(\text{RLT = relational (transitive) maker, TR = transitivizer, 1 = 1^{st}}\)
\(\text{person, SG = singular, O = object, 2 = 2^{nd} person, S = subject).}\)

In the second place, locus and order of morphemes may get out of synch with each order, in which case we have conflicting IC (= Immediate Constituent) structures. An English example would be ‘unfriendly’ with ‘un-’ added to the adjective ‘friendly,’ vs. ‘unkindly,’ with ‘-ly’ added to adjective ‘unkind.’ This problem can be easily addressed by using square brackets to indicate the different IC structures, viz.:

(3)  
\[
\begin{array}{l}
\text{un-friend-ly} \\
\text{[NEG-[friend-ADJ]]} \\
\text{‘Unfriendly’}
\end{array}
\]

(4)  
\[
\begin{array}{l}
\text{un-kind-ly} \\
\text{[[NEG-kind]-ADV]} \\
\text{‘Unkindly’}
\end{array}
\]

In the third place, non-concatenative morphology (such as ablaut, infixation, or reduplication) may interfere with a linear morpheme-by-morpheme read-off in that in these cases a morphological operation takes place inside a morpheme, rather than at the periphery of a morpheme. English examples are ‘ran’ and ‘took,’ which are marked as past tenses by ablaut of the vowel of the base stems ‘run’ and ‘take.’ (Hockett 1954 discusses this particular problem in a classic study to which I refer the interested reader.)
Lillooet, with a morphology that is far more complex than English, poses its own problems with regard to concatenative and non-concatenative operations. It is to this issue that we turn next.

3 Lillooet morphological operations

Concatenative operations in Lillooet include first of all prefixation, marked with a period after the prefix when part of a word, but with a following hyphen in isolated quotations, as in n.s.ʔiɬən ‘my food’ (n· 1SG/POSS = 1st singular possessive, ʔiɬən ‘to eat,’ s· NOM = nominalizer). Another productive operation is suffixation, marked with a hyphen preceding the suffix, as in cʔas·miɲ·cih·as ‘s/he is coming for you’ (cʔas ‘to come,’ -miɲ RLT, -cih 2SG/O, -as 3SG/S).

Where concatenative morphemes are applied serially from right to left (prefixes) or left to right (suffixes), as in n.s.ʔiɬən and cʔas·miɲ·cih·as above, the IC structure does not need to be indicated with the use of square brackets or other devices. However, we do have contrastive IC structures here as well, as in -l-akaʔ ‘instrument’ (+ connective, -akaʔ ‘hand’), which is added in its entirety to kʷzús-əm ‘to work’ (kʷzus ‘work,’ -əm INTR = intransitivizer) to form kʷzús-əm[-l-akaʔ] ‘tool.’ (In order to avoid notational clutter, from here on I omit square brackets that enclasp an entire form.)

As for non-concatenative morphology, Lillooet has cases like pálaʔ ‘one’ > pá<k>p>laʔ ‘one person,’ where we have the consonant before the stressed vowel reduplicated and then placed after that vowel, and pálaʔ > pá<k>p>laʔ ‘one animal,’ with the above reduplication (which generally marks the diminutive) and a change á > á. Another example is √nuqʷ ‘warm (atmosphere)’ > nu(<ʔ)qʷ ‘to get warm,’ with infixation of ? (after the root vowel) as the inchoative marker. (As the examples show, I use angular brackets to mark the diminutive reduplication, and swing brackets to mark the inchoative interior reduplication.)

To complicate matters further, we may have non-concatenative morphology combined with contrastive IC structures and morphophonemic adjustments, as in √kaw ‘away’ > ká(?a)ʔw ‘to go far away (on land),’ with inchoative interior glottalization which, as in this case, may trigger ə-eapenthesis and glottalization of a following resonant, ká(?a)ʔw > ká<k>ʔa)ʔw
'to go a bit further' (with diminutive reduplication, change of á to á, and regular change of a > a in the final syllable). On the other hand, we have √̄̄ī̄w 'raw' > ̄̄ī̄(?)̄̄ā̄w 'raw (but something that should be cooked),' with interior glottalization following the diminutive reduplication in terms of rule ordering, in contrast to k̄̄ā̄k̄̄<k̄̄ā̄>̄̄w, where we have the opposite ordering.

In order to deal with problems like these, Van Eijk 2004 suggests a three-line format, with underlying form on one line, the English translations of the morphemes and processes in the underlying form on the next line, and the surface form and its English translation on the third line. Applying that solution to k̄̄ā̄k̄̄<k̄̄ā̄>̄̄w and ̄̄ī̄(?)̄̄ā̄w, but with the surface form on the first line and its English translation on the third line, and some other adjustments as explained below, we get:

(5) k̄̄ā̄k̄̄<k̄̄ā̄>̄̄w
   away{INCH}<DIM/a>
   ‘To go far away (on land)’

(6) ̄̄ī̄(?)̄̄ā̄w
   raw<DIM>{INCH}
   ‘Raw (but should be cooked)’

Van Eijk 2004 uses angular brackets for both diminutive reduplication and interior glottalization when marked inside the morpheme, but swing brackets for both operations in the underlying form, a practice that is abandoned here in favour of standardized marking throughout. The location of the translated morphemes in the second line with regard to the root shows which one is applied first and which one second (with, of course, the result that they show up in the reverse order in the surface form.) The code <DIM/a> indicates that the vowel that precedes the consonant that results from reduplication is underlyingly a, not a.

In addition to diminutive reduplication and interior glottalization, Lillooet also employs augmentative (AUG) reduplication, which repeats the first CVC of the root and places this copy to the left side of the root, with stress falling on either the copy or the root or a suffix, following a set of largely regular rules. The CVC copy is marked with a colon following it, as in s.yap

240
‘tree’ > sɨ%pɨyɒp ‘trees,’ s.q”am ‘mountain’ > s.q”ám:q”am ‘mountain range,’ mác-xál ‘to write’ > mák:mac-xál ‘to write a lot.’ (See below for the derivational ordering of the prefix s- and augmentative reduplication.) A less productive operation is CV reduplication which serves various functions, reduplicates the first C of the root, and either copies the first V of the root as well or has a prespecified vowel. It is also marked with a following colon, as in pálɑ? ‘one’ > pìpálɑ? ‘one at a time.’ Finally there is continuative reduplication, which signals a continuative (CONT) aspect, with an out-of-control function as well, and reduplicates the stressed vowel of the targeted word, plus the following consonant. The VC copy is marked with a preceding equal sign when at the periphery of a targeted morpheme, but with a combination of angular brackets and the equal sign when inside a morpheme, as in √pú4 ‘get boiled’ > pú4=ǝɬ ‘to be boiling,’ pálɑ? ‘one’ > pál=ǝɬa? ‘to come together.’

Of course, the picture gets more complicated where we have concatenative and non-concatenative morphology applied in the same word, as in sɨ%pɨyɒp above, where s- is applied before the augmentative reduplication, but still appears farther from the root, in contrast to n.s.7ɨɬǝn, where locus and order of the morphemes are not out of synch with each other. As is mentioned above, in n.s.7ɨɬǝn we do not need to indicate the IC structure, since we have the unmarked order of operations here, but the IC structure of sɨ%pɨyɒp is captured as s.[ɨp]:yɒp. A more complex case is čaʔ:ca(?i:x-ú<4>4 ‘always, shy, really shy,’ with √cax ‘shy, ashamed,’ interior/inchoative glottalization, augmentative reduplication, -ú4 (inherently stressed) ‘always,’ and diminutive reduplication, all to be presented as:

(7) čaʔ:ca(?i)x-ú<4>
    [AUG:[shy[INCH]]-always]<DIM>
‘Always shy, really shy’

There is not enough evidence to decide whether the augmentative reduplication or the interior glottalization are applied first, which is why they are presented here as a combined operation, but since the form čaʔ:ca(?i)x-ú4 (without diminutive reduplication) was also recorded, it is clear that diminutive
Reduplication comes last in the derivation of this form. (Note also that the \text{?} that results from interior glottalization is read off by the CVC augment, an issue explored in greater depth in \cite{VanEijk1993}.)

In addition to reduplication, infixation and affixation, Lillooet also employs compounding, with the interfix \text{+aɬ+} linking the two members of the compound as in \text{ḻp+ aɬ+ kʷúna?} ‘buried (\text{√ḻp}) salmon roe (kʷúna?)’ (the traditional way of curing and preserving salmon roe, sometimes referred to as “Indian Limburger cheese”).

Clitics are indicated with the underloop (\text{ˬ}) which follows proclitics and precedes enclitics, as in \text{t蛊 tmíxˬa} ‘the land (tmíx”)’ (t蛊 ‘present/known/singular’ article, which always requires the reinforcing element \text{ˬa}).

4 A Lillooet text: The two coyotes

The following text, which was recorded from Bill Edwards in 1973, will serve as an illustration of the points made in the preceding sections. The English version of the text is also available as \cite{VanEijk2008} and as part of a set of Lillooet tales in \cite{Elliott1931}. Both the Lillooet and English versions are also available as pp. 3-4 of \cite{VanEijkWilliams1981}. As per the discussion above, the first line gives the surface forms, the second line the underlying form with the translations or codes of the morphemes or processes, plus the IC structure where this is not strictly linear, and the third line gives the running translation. As Lillooet morphophonemics are relatively simple, in contrast to, for example, Nlaka’pamux, a separate line for the forms after morphophonemic adjustments is not necessary. Abbreviations that are not explained in the preceding sections are explained in section 5. An underlined polymorphemic form is translated in its entirety in the next line, but analyzed in section 5.

\begin{Verbatim}
(8) \text{ṯá̱k kʷu? ká.tí? \text{?i.ṉkyá̱p.a}, \text{?á̱k?a}÷\text{ñwas}.}
\end{Verbatim}

\begin{flushleft}
\text{go REP around DEM ART/PL coyote REIN, two<DIM>}
\end{flushleft}

‘Two coyotes were going there.’
Then one of them said: "I am a coyote. Everybody knows that I am a coyote, but you are not a coyote, you are 'another one.'"

"Not at all, I am also a coyote," he said.

"I am going to go across this garden here, you listen to the people!"
(13)  ꧘ʔáq.kʷuʔ  ṭʔay_/ ṭʔáq.kʷuʔ ṭʔá.tiʔ,
      cross.REP  next,  cross.REP  thither.DEM,
      ṭʔačx-ɬ-ʔom.kʷuʔ  ṭʔa.kliʔuxʷ“almíxʷ”a
      seen-TR-PASS.REP  by.ART/PL.person.REIN.
      ‘Well, he went across, he went across that way, he was seen by
      the people.’

(14)  “tay,  ṭʔak  kəntʔú  ta  n.ʔyáp-a,  n.ʔyap
      “hey,  go  around.DEM  ART.coyote.REIN,  coyote
      ká.tiʔ  ta  ṭʔák-a.”
      around.DEM  ART.go.REIN.”
      “‘Hey, there’s a coyote going there, it is a coyote that is going
      there.’”

(15)  ṭʔák.kʷuʔ,  ka.ʔim̓-a.kʷúʔ  tuʔ.
      go.REP,  RES.disappear.REINF.REP_FIN.
      ‘He carried on, and he went out of sight.’

(16)  qʷačćəc  4əl.kʷʔú  niʔ-na, nůkʷ”a,
      take off  from.DEM  DEM.ART.other.REIN,
      ka.əxʷ-a.kʷúʔ  ṭʔuʔ,
      RES.appear.REIN.REP_so,  notice-RLT-3PL/S_REP.
      ‘Then the other one took off, he suddenly appeared, and they
      noticed him.’

(17)  “ⱇʔak  mútaʔ?  ta  pá<paʔ>p?, a,  pá>p>lə?
      “go  again  ART_one<DIM/a>_REIN,  one<DIM/a>
      ká.tiʔ  ta  ṭʔاك-a  mútaʔ.”
      around.DEM  ART.go.REIN  again.”
      “‘There goes another one, it’s another one that’s going there.’”
(18) ḳákʷ_u?,  cíxʷ_u?  ?ayʔ,  pzán-as_kʷ_u?
go REP,  get there REP  next,  meet-3SG/S REP
na_s.núkʷ_aʔ-s_a.
ART_[NOM.friend]-3SG/POSS_REIN.
‘He went on, and he got to the other spot, where he met his
friend.’

(19) “ʔáʔhan-cu,”  cún-əm_kʷ_u?,  “ʔáʔhan-cu,
“see?-2SG/POSS(?)”  tell-PASS REP,  “see?-2SG/POSS(?)”
qańim-ən-s-wít-kaxʷ_ha,”  n.kyáqפי-kan,
hear-TR-TR-3PL/O_INT,”  coyote-1SG/S
pə<p>laʔ-4káxʷ  s.núwa.
one<DIM/a>-2SG/S  NOM.2SG.
“See?,” he was told, “See?  Did you hear them?  I am a coyote,
but you are ‘another one.’”

5 Grammatical comments

Abbreviations used in the text and not explained previously are: REP
= reportative, used to indicate that the speaker did not witness the described
events him/herself; it is used with great frequency in legends, like the present
one. DEM = demonstrative pronoun; Lillooet distinguishes 12 of these,
marking three degrees of distance, visible vs. non-visible, and singular vs.
plural; tiʔ marks ‘that (middle distance, singular, visible);’ káʔtiʔ ‘around there’
expresses ‘at all’ when in combination with ḡwʔaz ‘not.’ ART = article;
Lillooet distinguishes eleven of these, including ili, ‘present, known, plural,’ taₐ,
(tí in the Mount Currie dialect) ‘present, known, singular,’ naₐ (niₐ in the
Mount Currie dialect) ‘absent, known, singular,’ all three of which are used in
the text; after proclitic prepositions, ili is kiₐ, as in sentence (13). REIN =
reinforcing enclitic required by a number of articles and by RES = the
resultative marker kaₐ, as in sentences (15) and (16). KAT = kataphoric
pronoun “it is the one who...;” the kataphoric pronoun niₐ is often combined
with the discourse particle ḡuʔ (broadly translated as ‘so’) into a construction
that generally translates as ‘and so, and then’ and requires the following verbal
construction to be in the factual (FACT) form (marked with the nominalizer s-), as in s.cut-s in sentence (9). ADH = adhortative, used mostly in commands, but also in non-imperative constructions to put a certain emphasis on the utterance. PASS = passive marker. FIN = final state marker, used mainly to indicate that something is over and done with. POSS = possessive; in ʔáʔhan-cu (19), the suffix -cu is probably an allomorph of -su 2SG/POSS, although this cannot be proven with complete certainty.

As for the underlined forms, translated as a whole, n.ältap ‘coyote’ (sentence 8) contains the root ħyap ‘coyote,’ and the (here semantically empty) prefix n- ‘referent of root is located in a larger setting.’ The suffix combination -c-al (9) contains -c 1SG/O and the connective -al, required when -c is followed by -it-as which itself consists of -it 3PL and -as 3SG/S. The form ħa.same (9) is here interpreted as containing ka ‘apparently’ and the adhortative enclitic same, but Kuipers 2002:41 reconstructs a monomorphemic Proto-Salish root *kí/amal for this form. The form 4k"ún-s.a ‘now’ (11) contains 4k"un ‘now,’ which is often extended with -s 3SG/POSS and the reinforcing enclitic same. In n.lap:xál-ton ‘garden’ (12) we have the locative prefix n- here forming an ambifix (circumfix) with -ton ‘instrument, implement’ to express ‘setting,’ the root lap ‘to plant, and the intransitivizer -xal; the IC structure of this word is thus n.[ləp:xál]-ton.

The forms ʔákʔa>nwas (8) and pá<k>laʔ (passim) show the diminutive reduplication that typically marks the category ‘animals’ in numerals. The form pá<k>laʔ is here translated as ‘another one’ rather than ‘one animal’ to better capture the meaning of the Lillooet form in the context of the story.

Further details on the morphological operations that are employed in this text are provided in Van Eijk 1997.

6 Conclusions

As should be clear from the above, non-concatenative morphology, IC structure and morphophonemic changes may conspire to complicate a one-on-one read-off between (in this case) Salish morphemes or processes and their English glosses. This has of course been recognized in a number of text editions. For example, Thompson and Thompson 1992:226 (sentence 309) code /k[ʔ]éw ‘to go far’ (with the root kew ‘far,’ indicated by the slash /, and
with the inchoative infix Ɂ, indicated by the square brackets [..] as ‘/far[INC].’ Also, on p. 217 (sentence 199) they code /qâ[Ɂ[Ɂ]-mâ ‘to go hunting’ (with diminutive reduplication that patterns like that in Lillooet, on the root /qâ ‘to shoot’) as ‘/shoot[•DIM]-MDL’ (MDL = middle voice).

The problem of matching non-concatenative morphology with morpheme-by-morpheme interlinear translations is of course not limited to Salish. For example, Pustet 2002:386 codes Lakota na-ma-ya-x’u ‘you hear me,’ where ma ‘me’ and ya ‘you (subject)’ are infixed in the stem na-x’u, as ‘S-1SG.P-2SGA-hear’ (P = patient, A = agent), with S (= stem) indicating that the first part of the inflected form (na) is part of a larger stem.

An additional problem is that where the over-all meaning of a form cannot be easily predicted from the constituent parts of that form, at least not from an English point of view (or that of any language into which the target language is translated). A case in point is n.l ǝ-xál-tǝn ‘garden,’ the constituent parts of which are analyzed above. Interestingly, Matthewson 2005:99 (sentence 250) codes this form as ‘LOC-plant-ACT-thing’ (ACT = active intransitivizer). On the other hand, Matthewson 2005:71 (sentence 112) lists n.x’xà-wúcin ‘four people’ without a morphemic break-down and codes it ‘four(human),’ while I would list this form as n.x ’xà-wúcin and code it as ‘[LOC.CV:]four,’ with the square brackets indicating that n- and CV: operate in tandem to indicate the category ‘people,’ and with the translation ‘four people’ appearing in the running translation. Thus, it is also up to the individual editor of the texts to decide what should be presented in terms of constituent morphemes or as a global translation.

In summary, there is, or there should be, a great deal of freedom as to how one wishes to present bilingual texts, as long as consistency in coding is maintained, and enough grammatical information is provided, either through an introduction or through footnotes, to allow the reader to gain an in-depth understanding of the texts. For example, Mattina and DeSautel 2002 present a number of Okanagan texts in a four-line format (surface phonetics, segmentation of the line into stem and inflectional affixes, morpheme gloss line, and free translation line), while Lindley and Lyon 2012 present their Okanagan texts in a three-line format (phonemic transcription with morphemic segmentation, morpheme glosses, and free translation). It is up to the individual readers to decide what they prefer, but both sets of texts are edited
and presented in what I consider an exemplary manner. (Both Mattina and DeSautel and Lindley and Lyon also add an unbroken English translation, while Lindley and Lyon also add an unbroken transcription of the Okanagan texts.)

Finally, one may argue that an interlinear presentation is not always necessary when a large enough grammatical commentary is provided. For example, Kuipers 1974 contains eight texts in their original Shuswap form followed by their unbroken English translations, and with the first two analyzed word by word in a section that follows the Shuswap original but without interlinear glossing. Even so, Kuiper’s detailed analysis provides the reader with enough information to understand the Shuswap words and sentences in the remaining six texts, and I would therefore strongly argue for a *vive la différence* approach when it comes to editing and presenting Salish texts, as long as the grammatical information provided is detailed enough to gain an in-depth insight into these texts.

7 Acknowledgements

The collection, translation, analysis and presentation of the St’át’imcets text in this article would not have been possible without the invaluable assistance of Bill Edwards, a man who is to be remembered as much for his courage, strength of character and moral rectitude as for his widely recognized talents as a St’át’imc speaker and story teller. My Salishanist colleagues, who over many years have made me the benefit of their insights, are hereby thanked collectively. As always, the intellectual, financial and logistic support of the First Nations University of Canada is gratefully acknowledged.

References


Davis, Henry. 2001. ‘Kayám: An Early St’át’imcets Text.’ *Anthropological*


Working Papers in Linguistics Vol. 32.
Van Eijk, Jan [P.], and Lorna Williams (eds.). 1981. *Lillooet Legends and


jvaneijk@fnuniv.ca