# A second-last position clitic in Sm'algyax: a solution

HENRY DAVIS
*University of British Columbia*

COLIN BROWN
*University of California, Los Angeles*

## 1   Recapitulation

In Brown and Davis (this volume) (henceforth, B&D), we introduced the *wh*-clitic *=du* in Sm'algyax and showed the following:

(a)  The syntactic position of *=du* is high in a root clause (taking CP as its complement).

(b)  Phonologically, *=du* is an enclitic: it is phonologically integrated with a phrase to its left.

(c)  Its linear position falls into three distributional patterns, which we have characterized as *wh-placement*, *predicate placement*, and *argument placement*.

We propose here a unified explanation for these generalizations. The core of our proposal is the following:

(1)   *Morphologically, =du is a **pro**clitic.*

   Section 2 outlines how (1) can account for the three linear positions outlined in (c) (for details, see B&D), Section 3 addresses the clause-final *wh*-particle *=da*, and Section 4 concludes and outlines a number of implications of our analysis.

## 2   Toward an explanation

Let us consider how the idea in (1) plays out in the three attested positions where *=du* occurs. We begin with *argument placement*, wherein the *wh-*

clitic encliticizes to the transitive subject, linearizing to the left of the object (2).[1] Argument placement is schematized in (3) below.

(2)   Dzindeł        dmt      dzapdit       Meelidu
        dzindeh=ł    dm=t    dzap-t=t     Meeli=**du**=a
        IRR.when=IRR.CN  PROSP=3.I  make/fix-3.II=PN  Mary=Q=CN
          ts'ikts'ik?
          ts'ikts'ik
          car
      'When will Mary fix the car?'


(3)   [ WH V DP$_A$=**du** DP$_O$]                 *Argument placement*

Assuming that =*du* is base-generated at the right periphery of the root clause (B&D Section 2) we suggest that the underlying syntactic structure for (3) can be represented as in (4):

(4)   [ [ WH V DP$_A$ DP$_O$] =**du** ]

Since its phonological requirements are met here, the question immediately arises as to why =*du* cannot remain in its base-generated position. The answer is that as a *morphological proclitic*, =*du* must precede a constituent to its *right*: hence, assuming that it will choose the most local possible morphological host, it will linearize inside the (O) DP immediately to its left, as in (3). Since it will still have a phonological host to its left, its phonological requirements will also be met, and the structure will be licit.

    Next, we turn to predicate placement, shown in (5) and schematized in (6):

---

[1]  1 = first person, 3 = third person, AX = agent extraction morpheme, CN = common noun connective, I = series I clitic, II = series II suffix, IRR = irrealis, PN = proper noun connective, PROSP = prospective, Q = question particle, REAS = reason subordinator, REL = relative, SG = singular, SX = subject extraction morpheme, T = "Big T" verbal morpheme, TR = transitive.

(5)  Goł          gan  dawłdut        Dzon?
     goo=ł        gan  dawł-t=**du**=t   Dzon
     what=IRR.CN  REAS  leave-3.II=**Q**=PN  John
     'Why did John leave?'


(6)  [ WH V=**du** DP$_{S/A/O}$]                    *Predicate placement*

It should be apparent that exactly the same analysis will account for this configuration, beginning with the base structure in (7):

(7)  [ [ WH V DP$_{S/A/O}$ ]  =**du** ]

This structure is identical to (3) except that here there is only one postverbal DP rather than two. Once again, as a morphological proclitic, =*du* will be forced to linearize inside the DP to its immediate left, this time phonologically encliticizing to the verb, and deriving the correct surface form.
     Now, let us turn to *wh-placement*, where =*du* attaches to the *wh*-phrase itself.

(8)  [ WH=**du** [ …] ]                             *Wh-placement*

As long as the *wh*-phrase is followed only by the predicate, the account we have already given will extend naturally to *wh*-placement, since =*du* will morphologically procliticize to the predicate to its right and phonologically encliticize to the preceding *wh*-phrase, as in (9) and (10):


(9)  Naayu   ksüüt?                (10)  Naayu   sibaasu?
     naa=**du**  ksüü-it                  naa=**du**  sibaas-u
     who=**Q**   go.out-SX                who=**Q**   scare-1SG.II
     'Who left?'                         'Who did I scare?'

However, there are plenty of cases where =*du* attaches to a *wh*-phrase even when multiple constituents follow it, as for example in (11), repeated from B&D (3):

121

(11)   Naayu   int      gaba            ts'ik'aaws?
       naa=**du**  in=t    gap[-t]=a ___ ts'ik'aaws
       who=**Q**   AX=3.I  eat[-3.II]=CN  split.salmon
       'Who eats split dried salmon?'

Here, we expect predicate placement; and in fact, it turns out that in such cases *=du* freely alternates between attaching to the *wh*-phrase, as in (12a) below, and attaching to the predicate, as in (12b):

(12)   a.   Naadu   int      yoyksa          nooł?
            naa=**du**  in=t    yoyks[-t]=a    nooł
            who=**Q**   AX=3.I  wash[-3.II]=CN  dish

       b.   Naał            int      yoyksdu           nooł
            naa=ł           in=t    yoyks-t=**du**=a    nooł
            who=IRR.CN      AX=3.I  wash-3.II=**Q**=CN  dish
            'Who washed the dishes?'

We suggest that the key to extending the morphological proclitic analysis to cases such as (12a) is to treat the entire string following the *wh*-phrase as *a single DP* whose internal structure is opaque to *=du*. In that case, the *wh*-clitic will be morphologically proclitic to the DP, and will phonologically encliticize to the preceding *wh*-phrase. In contrast, in cases such as (12b), the constituent following the *wh*-phrase will be CP, and *=du* will attach to the predicate, as expected.

However, in order to avoid circularity, the claim that the constituent following a *wh*-phrase + *=du* sequence is a DP rather than a CP needs to be independently motivated. Fortunately, there is a test. Recall from B&D Section 2 that *wh*-questions in Tsimshianic can either be derived by direct or indirect movement. Direct movement parallels *wh*-movement in English: the *wh*-phrase moves to a position on the left periphery of CP, leaving a clausal remnant. Indirect movement, on the other hand, involves a base-generated *wh*-predicate on the left periphery, followed by a DP argument, which usually takes the form of a headless relative clause, as exemplified in (13) and schematized in (14), repeated from B&D (7)–(8).

(13)  Godu    gu  yoyksis        Meeli?
      goo=du [**gu**  yoyks-i[-t]=s    Meeli ___]
      what=Q  **REL**  wash-TR-3.II=PN  Mary
      'What did Mary wash?'

(14)  [ **Goo** [   ___  [   ___  [ **gu** [    yoyksis Meeli ___ ] ] ] ] ]
      [IP **WH** [DP *pro* [CP O̶r̶e̶l̶ [  C [IP  yoyksis Meeli O̶r̶e̶l̶ ] ] ] ] ]

Since the argument of the *wh*-predicate in the indirect movement structure
in (13) is a DP, this is exactly the configuration where we predict =*du* will
encliticize to the *wh*-phrase rather than the predicate (as indeed, it does in
(13)).

    Furthermore, since the indirect movement structure can be readily
identified by the optional presence of the *wh*-relative pronoun *gu* (B&D
Section 2), we can formulate the following prediction:

(15)  *If a wh-question contains gu, =du will always attach to the wh-*
      *phrase*

This prediction is borne out. The examples in (16) show that only *wh-*
*placement* is available for =*du* in questions containing *gu* (compare (12b)
above):

(16)  a.  Naadu  gu   int     yoyksa        nooł?
          naa=**du gu**  in=t   yoyks[-t]=a    nooł
          who=**Q**  **REL**  AX=3.I  wash[-3.II]=CN  dish
          'Who washed the dishes?'

      b.  *naa=ł       **gu**  in=t   yoyks-t=**du**=a  nooł
          who=IRR.CN  **REL**  AX=3.I  wash-3.II=**Q**=CN  dish

The object questions in (17) make the same point: without *gu*, =*du* can
either attach to the *wh*-phrase (17a) or in penultimate position (17b), re-
flecting ambiguity between direct and indirect movement. With overt *gu*,
however, only indirect movement is possible, and therefore =*du* must at-
tach to the *wh*-phrase (17c); attempts to attach it to the predicate in penul-
timate position are ungrammatical, as shown in (17d).

(17)  a.  Godu     yoyksis        Meeli?
          goo=**du** yoyks-i[-t]=s   Meeli
          what=**Q**  wash-TR-3.II=PN  Mary

     b.  Goł           yoyksadut       Meeli?
          goo=ł          yoyks-i-t=**du**=t    Meeli
          what=IRR.CN  wash-TR-3.II=**Q**=PN  Mary

     c.  Godu     gu  yoyksis        Meeli?
          goo=**du** gu  yoyks-i[-t]=s   Meeli
          what=**Q**  REL  wash-TR-3.II=PN  Mary
          'Who washed the dishes?'

     d.  *goo=ł          **gu**  yoyks-i-t=**du**=t    Meeli
          what=IRR.CN  REL  wash-TR-3.II=**Q**=PN  Mary

It still remains to be explained why the DP (relative clause) con-
stituent in (17c) is impenetrable to =*du*, as evidenced by the ungram-
maticality of (17d). Here we appeal to the notion of a phase (Chomsky
2001 and much subsequent work). One of the leading ideas behind this
notion is that phases act as "chunks" for the purposes of spell-out, and
once spelled out, will be opaque to further operations — in this case, to
procliticization by =*du* in the morphological component. It is commonly
assumed that DPs are phases, and we adopt this assumption here.

By the same token, the CP complement of =*du is* penetrable to cliti-
cization: this means either that CP is not a phase or that =*du* is part of
the same phase as its CP complement. Evidence for the latter comes from
long distance extraction, where procliticizaton of =*du* takes place in the
matrix rather than the subordinate clause, as shown in (18)–(19), repeated
from B&D (40)–(41).

(18)  Ndeł           małdidut        Betty gooys       Meeli?
       ndeh=ł          mał-t-i-t=**du**=t  Betty [goo-i[-t]=s   Meeli ___]
       where=IRR.CN  say-T-TR=**Q**=PN  Betty  go-TR-3.II=PN  Mary
       'Where did Betty say Mary went?'

(19)    Goł          ha'ligoodut     Bettyt     giindit

          goo=ł       ha'ligoot-t=**du**=t Betty  [=t   giin-t=t

          what=IRR.CN  think-3.II=Q=PN   Betty   =3.I give-3.II=PN

             Michaelt    Henry?

             Michael=t   Henry &underline;    ]

             Michael=PN  Henry

        'What does Betty think Michael gave Henry?'

In long-distance questions, =*du* always appears in the matrix rather than an embedded CP (see B&D Section 2.2). The inaccessibility of subordinate CPs follows if, like DPs, they constitute phases. At the point of clitic linearization, embedded CPs have already been spelled out, whereas the matrix CP has not, and is therefore accessible to =*du* placement.

    Finally, as we saw in B&D Section 4, linearization of =*du* only pays attention to the predicate and its DP arguments. If we treat =*du* as a phrasal proclitic, as seems necessary to account for its positioning with respect to DPs, this distribution appears odd, since the predicate is a head. However, the only cases where =*du* apparently procliticizes to a predicate are precisely those where we have just shown that the *wh*-phrase is itself a predicate taking a DP argument (i.e., cases of indirect movement, such as in (16a) and (17c) above). In other words, here =*du* is a morphological proclitic to DP, just as in cases of argument placement. This means that we can now eliminate the predicate from the set of possible proclitic hosts, leaving us with a starkly simple generalization:

(20)    *Only DPs count for the linearization of =du*

We take it as a virtue of the current analysis that what appears at first sight to be a very complex distributional pattern is reducible to the interaction of simple constraints on linearization such as (20), together with independently motivated structural properties of the language (the distinction between direct and indirect Ā-movement) and widely accepted conditions on the interface (spell-out by phase).

## 3   *=Da*

We have seen how the dual status of =*du* as a phonological enclitic and a morphological proclitic accounts for its "second last position" behaviour:

it will always end up sandwiched between a phonological host to its left and a morphological host (a DP) to its right. But what happens when there is only a single constituent to attach to?

It turns out that in these cases, the *wh*-question clitic takes a separate form, *=da*, which is uniformly *en*clitic.[2] This form is obligatory in reduced questions consisting of just a *wh*-phrase:

(21)  Naaya?/Naada?  
    naa=**da**  
    who=Q  
    'Who (is it)?'

(22)  Goya?/Goda?  
    goo=**da**  
    what=Q  
    'What (is it)?'

*=Da* is also optional instead of *=du* in some non-reduced questions, where it always surfaces in final position.[3]

(23)  Goł      gabida?  
    goo=ł      gap-i-t=**da**  
    what=IRR.CN  eat-TR-3.II=Q  
    'What did s/he eat?'

(24)  Naał      int  
    naa=ł      in=t  
    who=IRR.CN  AX=3.I  
         gapda?  
         gap-t=**da**  
         eat-3.II=Q  
    'Who ate it?'

The distribution of *=da* provides further indirect support for our analysis of *=du*, since it surfaces exactly where we expect *=du* to be impossible.[4]

---

[2] The morphophonology of *=da* is consistent with that of *=du* as described in B&D §3 (for example, *=da* also optionally exhibits free variation between [da] and [ja] when immediately following a *wh*-word), suggesting that the two are allomorphs. For reasons of space, we do not explore this possibility further here.

[3] Final position is typical for question clitics across Tsimshianic: in Sgüüxs (Southern Tsimshian), both *wh*-Qs and yes-no Qs are marked by a final enclitic *=i*, and in Interior Tsimshianic, yes-no Qs are marked by a final *=aa*, with no marking for *wh*-Qs.

[4] There is more to say about the distribution of *=da* in cases where *=du* is also possible. To be specific, *=da* is available as a (preferred) alternative to *=du* in cases of A and O extraction with a third person and no following DP, such as

## 4   Conclusion and Further Implications

The main points of the analysis are summarized in (i)–(vi) below.

(i)   *=du* is base-generated in the syntax in a high MoodP on the right periphery which encodes illocutionary force and takes a root CP as its complement (B&D Section 2)

(ii)   *=du* is a phonological enclitic which attaches to a prosodic host to its left (B&D Section 3)

(iii)   *=du* is a morphological proclitic which must precede a phrasal host to its right (B&D Section 4, Section 2 of this paper)

(iv)   as phases, DP and CP are opaque to cliticization once spelled out (Section 2 of this paper)

(v)   only a DP may serve as a proclitic host for *=du* (Section 2 of this paper)

(vi)   where the dual requirements of *=du* as a morphological proclitic and a phonological enclitic cannot be met, *=da* (which is both a morphological and a phonological enclitic) is inserted instead (Section 3 of this paper)

In this final section, we briefly explore some of the implications of our account and the model of grammar which it entails. While for reasons of space we cannot give an explicit formal analysis, we will point to the kind of grammatical architecture which we think will be necessary to handle the Sm'algyax facts.

We begin with a significant theoretical claim which we think is almost unavoidable, given the facts we have presented.

(25)   *The linearization of clitics is not reducible to either their syntax, their phonology, or any combination of the two.*

---

those in (23) and (24); otherwise, only *=du* is permitted. We set aside this extra complication here for the sake of space.

Indeed, insofar as our account is successful, it serves as a proof-of-concept of the existence of a separate morphological component of the grammar responsible for the linearization of clitics.

Second, we observe that this component must occupy a very specific position in the architecture of the grammar: it takes the syntax as its input, and the phonology as its output. Clitic linearization, in other words, takes place at the interface between syntax and phonology.

Third, our account supports a two-step model of lexicalization, in which features governing linearization are activated at the first step, and morphophonological features (e.g., those governing contextual allomorphy) come into play in the phonological component only after linearization has taken place.[5]

Fourth, we have outlined some of the parameters of the linearization operation itself. The following factors appear to be relevant:

(a) The direction of (morphological) cliticization (left for enclitics, right for proclitics).[6]

(b) The nature of the (morphological) host. There are two factors to consider here:

   (i) Whether the host is a phrase (leading to "phrasal affixation") or a head.

   (ii) The categorial signature of the host.

For =*du*, the direction is rightward, and the host is a DP.

In its appeal to morphology, the model we have briefly outlined here owes an intellectual debt to previous accounts of cliticization such as those of Billings (2002), Anderson (2005), and particularly Klavans (1985). In fact, a significant empirical contribution of this paper is to vindicate one of the predictions made by Klavans' parametrization of possible clitic

---

[5] A model of clitic linearization with exactly these properties is laid out in Huijsmans (2023) on the basis of evidence from second position clitics in Salish: see also Davis and Huijsmans (2021).

[6] Given the existence of what Mulder and Sellers (2010) refer to as "flexiclitics" in Sm'algyax (that is, clitics which indiscriminately attach either to the left or right), it is possible that this parameter can remain unspecified for some clitics.

positions: the existence of penultimate or "second-last" position clitics (Type 5 in her typology: see Klavans 1985: 103).[7]

However, the architecture which we employ and the division of labour between the narrow syntax and its interfaces is very much in the derivational tradition of minimalism, as is our use of the machinery of spell-out by phase. Overall, we hope to have shown here that a separate operation of clitic linearization in the morphology allows for an elegant account of a very complex pattern of cliticization in Sm'algyax, with broader implications for the treatment of clitics cross-linguistically and the architecture of the syntax-phonology interface.

## References

Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford University Press.

Billings, Loren A. 2002. Phrasal clitics. *Journal of Slavic Linguistics* 10:53–104.

Brown, Colin, and Henry Davis. This volume. A second-last position clitic in Sm'algyax: a puzzle.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, 1–52. MIT Press.

Davis, Henry, and Marianne Huijsmans. 2021. Clitics and coordination in two Salish languages. Poster presented at WCCFL 39, University of Arizona.

Huijsmans, Marianne. 2023. Second-position clitics, from morphosyntax to semantics: The ʔayʔaǰuθəm (Comox-Sliammon) perspective. Doctoral Dissertation, University of British Columbia.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61:95–120.

Mulder, Jean, and Holly Sellers. 2010. Classifying clitics in Sm'algyax: Approaching theory from the field. In *Fieldwork and linguistic documentation in indigenous languages of the Americas*, ed. An-

---

[7] See Billings (2002) for criticism of Klavan's original evidence for this claim, based on data from the Australian indigenous language Nganhcara.

drea L. Berez, Jean Mulder, and Daisy Rosenblum, 2, 33–56. URL http://nflrc.hawaii.edu/ldc/, Language Documentation and Conservation Special Publication.