

DOI: 10.2478/linpo-2013-0014

BACKERNAGEL IS WACKERNAGEL LITE. ON THE “P-MINUS 2” CLITICS OF SANTALI

HANS HENRICH HOCK

ABSTRACT: Hans Henrich Hock. *Backernagel is Wackernagel Lite. On the “P-minus 2” Clitics of Santali*. *Lingua Posnaniensis*, vol. LV (2)/2013. The Poznań Society for the Advancement of the Arts and Sciences. PL ISSN 0079-4740, ISBN 978-83-7654-274-4, pp. 67–75.

Santali presents structures with subject clitics in “P minus 2” (P-2) position, before the final verb and enclitic on the preverbal element, a position called “Backernagel” by KIDWAI (2005). P-2 is commonly considered to lack clear cross-linguistic support; moreover, while generative accounts can accommodate utterance-second position (P2) as adjunction to a left-peripheral projection, they have no ready way of accommodating P-2.

The history and synchrony of Munda “P-2” have elicited several accounts. ANDERSON (2007) considers three possibilities: Reanalysis of Proto-Munda subject proclitics as enclitic; extension of postverbal object clitics to preverbal subject function; attachment of original resumptive pronouns to the preverbal element.

I present evidence for a different hypothesis: The Santali Backernagel clitics originate as P2 or classical Wackernagel elements. A more fine-grained definition of Wackernagel in terms of different PROSODIC domains (such as utterance/theme vs. rheme) permits the hypothesis that the apparent P-2 is still a Wackernagel position, but within the rheme rather than the entire utterance, and that within the rheme, the prosodically strongest, preverbal-focus element is the most attractive clitic host.

I support my account with evidence from Santali and other Kherwarian languages (which offer traces of an original P2 position) and parallel developments in Iranian (where the different stages in the development can be traced in greater detail).

Backernagel, thus, is a subtype of Wackernagel, and there is no need to assume a typologically problematic P-2 position for Munda (or for various Iranian varieties).

KEY WORDS: clitics, P2 (Wackernagel) vs. “P minus 2” (Backernagel), prosody, language change

Hans Henrich Hock, University of Illinois at Urbana-Champaign, hhock@illinois.edu

“Wackernagel’s Law” or “P2” – the placement of clitics in clause-second position – is a well-known phenomenon and has been documented in numerous languages, although there is a great amount of disagreement concerning its historical motivation; see e.g. the various contributions to HALPERN & ZWICKY 1996. In this paper I accept the view that Wackernagel or P2 is in principle prosodically defined; see RADANOVIĆ-KOČIĆ’s (1988, 1996) convincing demonstration that Serbian-Croatian P2 can only be accounted for in terms of prosodic phrasing, and similar arguments by HOCK (1996) based on the evidence of Vedic poetic phrasing.

The logical opposite of Wackernagel or P2 would be what KIDWAI (2005) has referred to as “Backernagel”, the placement of clitics in clause-penultimate or “P-minus 2” position. It is commonly assumed that this is a position for which there is no clear evidence; see e.g. HOCK 2010. In fact, while generative accounts can easily accommodate P2 as adjunction to a left-peripheral projection (CP, IP, or the like), they have no ready way of accommodating P-2.

A possible counterexample to this view is found in Santali, which presents structures with subject clitics in what appears to be P-2 or Backernagel position, before the final verb (but enclitic to the preverbal constituent); see example (1). The historical explanation and synchronic status of this position have given rise to several different accounts.

- (1) *abu hola snema jɛl-lagit=**bun** cəlaw-len-a*
 we.INCL. yesterday cinema see-APP.=we.INC. go-PST-FIN
 ‘We had gone to see the movie yesterday.’ (from KIDWAI 2005)

In this paper I argue that Santali P-2 and similar structures in other languages, both Munda¹ and Iranian, can be historically accounted for as P2 or Wackernagel “Lite”, in the domain of the “rtheme”, rather than the “theme” (which is the more usual domain for Wackernagel).²

Let me begin with a survey of different accounts that have been proposed for Santali.

ANDERSON (2007) entertains three different historical explanations:

1. Proto-Munda verbal subject proclitics were reanalyzed as enclitic.
2. Postverbal object clitics were extended to subject function, with the reasons for their becoming preverbal being unknown.
3. Proto-Munda had resumptive pronouns which in North Munda were attached to the preverbal element.

Eventually Anderson comes to prefer the first of these scenarios. A similar account is that of CYSOUW (2004). KIDWAI’s (2005) solution is more similar to Anderson’s third scenario. The weakest of the three scenarios clearly is the second, because it fails to explain how and why postverbal subject clitics could be extended to subject function and, in addition, change to preverbal position.

A problem for both Scenarios 1 and 2 is the existence of Santali structures such as (2), with the clitic not after the preverbal constituent but after an earlier one (2a), even a clause-initial one (2b). KIDWAI (2005) cites such examples and the examples below come from her paper, but she does not explicitly address their different behavior.

- (2a) *abu hola=**bu** snema jɛl-lagi[t] cəlaw-len-a*
 we.INCL. yesterday=**we**.INCL. cinema see-APPL. go.PST-FIN
 ‘We had gone to see the movie yesterday.’
- (2b) *abu=**bu** snema jɛl-lagit cəlaw-len-a*
 we.INCL=**we**.INCL. cinema see-APPL. go.PST-FIN
 ‘We had gone to see the movie.’

Although such examples may be rare, they are not completely isolated. Consider the additional passages in (3), from other Munda languages or dialects. The general tendency is for

¹ In this paper, “Munda” is used for the Kherwarian subgroup of the larger Munda family.

² This claim is similar to, but broader than, that of GOLDBERG 1996, who argues that structures such as the one on the left side of (6b) below show Wackernagel clitic placement in the “domain of the predicate”.

the clitic to appear after the first constituent or element of the clause, with (2a) constituting a notable exception. Note that in examples (3b) and (3c), we can observe doubling of the clitic later in the clause, either after the preverbal constituent (*bāms-pahār*) or the verb (*nu-a*).

- (3a) *hōlā tikin=iñ buru dubich'-dubich' bāzār sen-ter-ā*
 yesterday noon=I mountain towards bazaar went
 ‘Yesterday at noon I went to the bazaar in the direction of the mountain.’
 (Korwa; GRIERSON 1906: 154)

- (3b) *dō=bū pahār, bāms-pahār=bū sen*
 well=we mountain, bamboo-mountain=we shall.go
 ‘Well we shall go to the mountain, the bamboo mountain.’
 (Korwa; GRIERSON 1906: 165)

- (3c) *tab se=le ina-ra? ba:d-re boḍoj²-ili nu-a=le*
 then from=we this-GEN. after.LOC rice.beer-rice.beer drink.IND=WE
 ‘Then after that we drink rice beer.’
 (Kera? Mundari; KOBAYASHI & MURMU 2008: 176)

Note further that in both Santali and other Munda languages, the dominant preverbal pattern of subject clitics varies with a postverbal one, sometimes with the clitic appearing in both positions; see GHOSH 2008: 54–55 for Santali, OSADA 2008: 120 for Mundari, and KOBAYASHI & MURMU 2008: 176 for Kera? Mundari, as well as example (3c) and the further illustration in (4), where the clitic is both pre- and postverbal. As Osada notes, the postverbal position is preferred by younger speakers. This suggests that it is a new development.

- (4) *āyub-berā=e goch'-nen-ā-e*
 evening-time=he died=he
 ‘In the evening he died.’
 (Asuri; GRIERSON 1906: 142)

Finally, note that in utterances whose only non-clitic constituent is the verb, the clitics always follow the verb, as in (5). It is possible that the innovated postverbal clitic placement in Santali and related languages results from an extension of this postverbal placement.

- (5) *nel-gɔt'-ka-t'-ko-a=ko*
 see-EMPH-COMPLASP-TR.3PL-FIN=3PL
 ‘they saw them off’
 (Santali; ANDERSON 2007: 245)

Let us, then, summarize our findings for Santali and the related languages.

- i. Subject clitics most commonly occur in preverbal position, where they are enclitic on the preceding constituent or element.
- ii. In this position near the end of the clause they may vary with postverbal placement, an apparent innovation.
- iii. Subject clitics may also occur after earlier elements, especially clause-initial ones. This pattern appears to be rare and thus may be an archaism.
- iv. In addition there is doubling.

To these findings we must add the fact that as in other SOV languages, the preverbal constituent or element is likely to be in FOCUS position, hence prosodically the strongest element in the clause, presumably reflecting an interaction between two cross-linguistic tendencies:

the tendency for the rheme to be prosodically stronger than the theme and the tendency for final verbs to have the weakest prominence in the utterance; see Figure 1. For discussion of these tendencies see KIM 1988 with HARLIG & BARDOVI-HARLIG 1988 and LADD 1980, 1996.³ The preverbal Focus position therefore is a prosodically optimal clitic host.

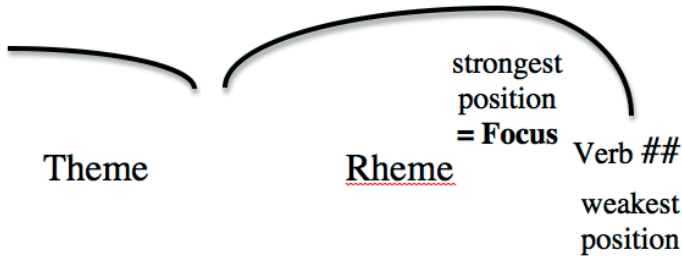


Fig. 1. Theme-Rheme and the Preverbal Focus Position

Against this background we can propose the following preliminary hypothesis on the origin and development of subject clitics (ignoring the issue of doubling).

- A. The Munda subject clitics started out as regular, clause-scope, Wackernagel or P2 elements, enclitic on the first constituent or element of the clause or utterance. This accounts for the marked pattern iii.
- B. The constituent in preverbal focus position, having the highest prosodic prominence in an SOV language, attracts the subject clitics away from P2 (i.e., from the position after the theme) into the domain of the rheme.
- C. The final step consists in the movement of clitics into the postverbal position.

Comparison with various Iranian languages, especially with Parachi, provides support for this hypothesis, as well as suggesting certain refinements.

In Old and Middle Iranian, the unmarked position of clitics is P2; see SKJÆRVØ 2009a b, YOSHIDA 2009, DURKIN-MEISTERERNST 2009, EMERICK 2009. Occasionally, however, clitics may be in postverbal position. Moreover, if the verb is the only non-clitic constituent, clitics, being enclitic, naturally follow (with complications arising in prefix-verb combinations).

By contrast, in Modern Persian and Tajik, subject clitics are always postverbal and non-subject clitics are postverbal except optionally in prefixed and compound verbs, as in (6); see WINDFUHR & PERRY 2009.

(6a) *did-am=aš*
 see.PAST-1SG.SUBJ=3SG.OBJ
 ‘I saw him.’

(6b) *bidār-aš kard-am* BESIDE *bidār kard-am-aš*
 waking=3SG.OBJ_{do}.PST-1SGSUBJ waking_{do}.PST-1SG.SUBJ=3SG.OBJ
 ‘I woke him up.’

³ LADD (1996) considers Bangla an apparent exception to verb-final prosodic weakness (with reference to HAYES & LAHIRI 1991); but Dutta & Hock show that Bangla firmly conforms to this crosslinguistic tendency.

In many other West Iranian languages, subject or agent clitics exhibit greater freedom, commonly attaching to the preverbal constituent, as in (7). (Data from HAIG 2008.)

- (7a) *tani ha=s kærd*
 he.OBL. running=3SG do.PST
 ‘He ran away.’ (Vafsi)
- (7b) *āmīn kâghâz=îm kîâst*
 I letter=1SG send.PST.3SG
 ‘I sent a letter.’ (Gorami)
- (7c) *min ewā=m bini*
 I you.PL.=1SG see.PST
 ‘I saw you.’ (Suleimani Kurdish)

Haig attributes this phenomenon to a “rightward drift” of “head attraction”, starting from the original Wackernagel or P2 placement, without offering further explanation. (The focus of his monograph is on Alignment.) But if we consider the overall evidence of West Iranian, we can detect a tendency very similar to Munda, namely placement of subject clitics either after the preverbal constituent or element or, presumably as an innovation, in postverbal position.

Similar phenomena can be observed in Balochi; see (8). (Data from JAHANI & KORN 2009.) But as (8c) shows, the clitic may also attach to an earlier constituent.

- (8a) *wāja ālī gopt=ī*
 Mr. A. said=3SG
 ‘Mr. Ali said.’
- (8b) *ālī jān=ī=ī košta*
 Ali wife=3SG=3SG killed
 ‘Ali killed his wife.’
- (8c) *cī-ā man=et weyl kot-ag*
 what-OBL. me=2SG abandon do.past-pple
 ‘Why have you abandoned me?’

Most interesting is the situation in Parachi, a Southeastern Iranian language. Here clitic pronouns typically attach to the leftmost element, as in (9a, b), and the agentive-marking agent clitic is placed after the verb if there is no other constituent (9c). Otherwise the agent clitic preferably attaches to the constituent immediately preceding the verb; but it may alternatively attach to an earlier constituent ‘for selective emphasis’ (9d); see KIEFFER 2009. Put differently, the placement of the clitic =*a* in (9d) depends on which constituent in the verb phrase or rheme is emphasized and hence prosodically most prominent in that domain.

- (9a) *ce=om qasam xor*
 that=1SG oath ate
 ‘that I swore an oath’
- (9b) *mēr=an=om te*
 kill=3PL.=1SG FUT.
 ‘they will kill me’

- (9c) *xor*=**a**
ate=2SG
'you ate'
- (9d) *tū nī waxān(=a) nāgōn(=a) ce-pen(=a) xor*
you.SG. tonight(=2SG) bread(=2SG) what-with(=2SG) ate
'With what did you eat bread tonight?'

The combined Iranian evidence examined so far, thus, provides a strong, historically well-documented parallel to the scenario hypothesized earlier for Santali and other Munda languages:

- A. The starting point is a classical Wackernagel or P2 placement of clitic pronouns.
- B. Subject (and other) pronouns get attracted by an accented element in the rheme, which presumably is prosodically more prominent than the clause-initial theme.
- C. The best candidate for this attraction is the preverbal element which presumably is in Focus Position.
- D. The final step consists in the movement of the clitics into the postverbal position.

The account can be further refined by considering additional evidence from Eastern Iranian.

Among the Iranian languages, Pashto has preserved clitic Wackernagel/P2 most faithfully; for details see TEGEY 1977. However, as ERSCHLER (2009) notes, the placement of clitics may be delayed, as in (10b) vs. the normal pattern in (10a).

- (10a) *sangin=ba war mat kr-i*
Sangin=FUT door broken.M.SG do.PST.PF-M.3SG
'Sangin will break the door.'
- (10b) *sangin war=ba mat kr-i*
Sangin door=FUT broken.M.SG. do. PST.PF-M.3SG
'Sangin will break the door.'

Erschler further notes that in Shughni, another East Iranian language, clitic placement can be delayed even further, as in (11), but that 'the acceptability of a sentence usually decreases as the enclitic is moved rightwards'.

- (11a) *Madina=ta ar ruz garḏa piʒ*
Madina=HAB every day bread cook.PST.F
'Madina bakes bread every day.'
- (11b) *Madina ar ruz=ta garḏa piʒ*
- (11c)? *Madina ar ruz garḏa=ta piʒ*

These data suggest that there may have been an important intermediate step in the "rightward drift" of the clitics – a shift from after the clause initial, theme position to after the first element in the rheme, i.e. to a Wackernagel or P2 position within that domain. An alternative might be the presence of a prosodic break after the clause-initial element, with P2 enclisis to the first element after that break, similar to what is found in Serbian-Croatian

(RADANOVIĆ-KOČIĆ 1988, 1996). Unfortunately, Erschler does not provide prosodic information for his examples; and his own account downplays the role of prosody and argues for a purely syntactic approach. That prosody plays a greater role than Erschler admits can be gauged from the fact that in combinations of “Type II clitic pronoun” + disyllabic postposition + copula, the position of P-2 clitics varies, depending on whether it is the pronoun-postposition complex or the copula which is accented; see (12) and TEGEY 1977.

(12a) *rv-sarú de wi*
 me-with MODAL PARTICLE be
 ‘Let it be WITH ME.’

(12b) *rv-sara wí de*
 ‘Let it BE with me.’

This interpretation makes it possible to argue for a further refinement of the historical account for Santali/Munda and Iranian along the following lines.

- A. The starting point is a classical Wackernagel/P2 placement of clitics within the domain of the clause (or the utterance), after the first element of the theme (Old and Middle Iranian; traces in Munda).
- B. Since the rheme crosslinguistically has greater prosodic prominence than the theme, the clitic shifts to the rheme, landing after the first accented element in that domain – “Wackernagel (or P2) Lite” (Pashto and Shughni).
- C. Within the rheme the clitic shifts further to the right, to attach to the element with highest prominence within the domain (Parachi; perhaps also Pashto and Shughni).
- D. Since in SOV languages the preverbal position is the Focus position, prototypically with the highest prominence within the rheme, the clitic shifts to the position after this element – a further reduction in prosodic domain (most of Iranian and traditional Munda).
- E. Finally, aided perhaps by the postverbal placement in utterances whose only major constituent is the verb, the clitic shifts to the postverbal position (Persian and innovative Munda).

Although this account clearly is speculative, it does suggest that the apparent Backernagel or “P-2” behavior of subject clitics in Santali and other Munda languages, and similar phenomena in various Iranian languages, can be explained as arising from an original, classical Wackernagel or P2 behavior.

Moreover, by linking “P-2” to the prosodic strength of the PREVERBAL focus position it explains the fact that “P-2” is fundamentally different from “P2”. In “P-2”, attachment is not to the final element of the clause, but to the next-to-final, focused element. Moreover, whereas the element following “P-2” always is the VERB, the element preceding “P2” may belong to any morphological category. The present account, thus, captures the intuition that there is no clear cross-linguistic evidence for “P-2”, as well as the fact that while generative accounts can easily accommodate P2 as adjunction to a left-peripheral projection (CP, IP, or the like), they have no ready way of accommodating P-2.

The account could be strengthened – or challenged – by detailed work on the prosodic behavior of the various Munda and Iranian patterns that would test my assumption that the element preceding the clitic typically is a Focus element and hence has higher prosodic prominence than other elements within its domain. I very much hope that my hypothesis will motivate further work on these languages, paying close attention, not only to syntax, but also to phrasal prosody.

REFERENCES

- ANDERSON Gregory D.S. 2007. *The Munda Verb: Typological Perspectives*. Berlin–New York: Mouton de Gruyter.
- ANDERSON Gregory D.S. (ed.). 2008. *The Munda Languages*. London–New York: Routledge.
- BUBENIK Vit, LURAGHI Silvia. 2010. *Continuum Companion to Historical Linguistics*. London–New York: Continuum Books.
- CYSOUW Michael. 2004. “The Rise of Person Inflection with Special Reference to the Munda Languages.” In: *11th International Morphology Meeting, Vienna*, <http://email.eva.mpg.de/~cysouw/pdf/cysouwMUNDA.pdf>.
- DURKIN-MEISTERERNST Desmond. 2009. “Khwarezmian.” In: WINDFUHR 2009: 336–376.
- DUTTA Indranil, HOCK Hans Henrich. 2006. “Interaction of Verb Accentuation and Utterance Finality in Bangla.” In: HOFFMANN & MIXDORFF 2006.
- EMERICK Ronald E. 2009. “Khotanese and Tumshuqese.” In: WINDFUHR 2009: 377–415.
- ERSCHLER David. 2009. “Modularity and 2P Clitics: Arguments from Digor Ossetic.” In: *Proceedings of IATL 25* [Israel Association for Theoretical Linguistics], <http://linguistics.huji.ac.il/IATL/25/Erschler.pdf>.
- GHOSH Arun. 2008. “Santali.” In: ANDERSON 2008: 11–98.
- GOLDBERG Adele. 1996. “Words by Default: Optimizing Constraints and the Persian Complex Predicate.” In: *Proceedings of the 1996 Berkeley Linguistic Society Meeting*, 132–146.
- GRIERSON George A. 1906. *Linguistic Survey of India*. Vol. 4: *Munda & Dravidian Languages*. Calcutta: The Government Printing Offices.
- HAIG Geoffrey L.J. 2008. *Alignment Change in Iranian Languages: A Construction Grammar Approach*. Berlin–New York: Mouton de Gruyter.
- HALPERN Aaron, ZWICKY Arnold (eds.) 1996. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications.
- HAMMOND M. et al. (eds.). 1988. *Studies in Syntactic Typology*. Amsterdam–Philadelphia: Benjamins.
- HARLIG Jeffrey, BARDOVI-HARLIG Kathleen. 1988. “Accentuation Typology, Word Order, and Theme-rheme Structure.” In: HAMMOND et al. 1988: 125–146.
- HAYES Bruce, LAHIRI Aditi. 1991. “Bengali Intonational Phonology.” *Natural Language and Linguistic Theory* 9, 47–96.
- HOCK Hans Henrich. 1996. “Who’s On First: Toward a Prosodic Account of P2 Clitics.” In: HALPERN & ZWICKY 1996: 199–270.
- HOCK Hans Henrich. 2010. “Typology and Universals.” In: BUBENIK & LURAGHI 2010: 56–69.
- HOFFMANN Rüdiger, MIXDORFF Hansjörg. 2006. *CD-ROM Proceedings of Speech Prosody: 3rd International Conference, Dresden, May 2-5, 2006*. (Studentexte zur Sprachkommunikation, 40.) Dresden: UD press.
- JAHANI Carina, KORN Agnes. 2009. “Balochi.” In: WINDFUHR 2009: 634–692.
- KIDWAI Ayesha. 2005. “Santali ‘Backernagel’ Clitics: Distributing Clitic Doubling.” In: SINGH 2005: 189–207.
- KIEFFER Charles M. 2009. “Parachi.” In: WINDFUHR 2009: 693–720.
- KIM A.H. 1988. “Preverbal Focusing and Type XXIII Languages.” In: HAMMOND et al. 1988: 147–169.
- KOBAYASHI Masato, MURMU Ganesh. 2008. “Kera? Mundari.” In: ANDERSON 2008: 165–194.
- LADD D. Robert. 1980. *The Structure of Intonational Meaning: Evidence from English*. Bloomington–London: Indiana University Press.
- LADD D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- OSADA Toshiki. 2008. “Mundari.” In: ANDERSON 2008: 99–164.
- RADANOVIĆ-KOČIĆ Vesna. 1988. *The Grammar of Serbo-Croatian Clitics: A Synchronic and Diachronic Perspective*. Urbana: University of Illinois Ph.D. dissertation in Linguistics.

- RADANOVIĆ-KOČIĆ Vesna. 1996. “The Placement of Serbo-Croatian Clitics: A Prosodic Approach.” In: HALPERN & ZWICKY 1996: 429–445.
- SINGH Rajendra (ed.). 2005. *The Yearbook of South Asian Languages and Linguistics (2005)*. Berlin–New York: Mouton de Gruyter.
- SKJÆRVØ Prods Oktor. 2009a. “Old Iranian.” In: WINDFUHR 2009: 43–195.
- SKJÆRVØ Prods Oktor. 2009b. “Middle West Iranian.” In: WINDFUHR 2009: 196–278.
- TEGEY Habibullah. 1977. *The Grammar of Clitics: Evidence from Pashto and Other Languages*. Urbana: University of Illinois Ph.D. dissertation in Linguistics. [Published also by the International Center for Pashto Studies, Kabul, apparently with unchanged pagination.]
- WINDFUHR Gernot (ed.). 2009. *The Iranian Languages*. London–New York: Routledge.
- WINDFUHR Gernot, PERRY John R. 2009. “Persian and Tajik.” In: WINDFUHR 2009: 416–544.
- YOSHIDA Yutaka. 2009. “Sogdian.” In: WINDFUHR 2009: 279–335.