

# **Lenition: A challenging issue in Sangir orthography**

by  
Pamela Day

2010

**Sulang Language Data and Working Papers:  
Phonology and Orthography Descriptions, no. 1**



Sulawesi Language Alliance  
<http://sulang.org/>

## LANGUAGES

Subject language : Sangir  
Language of materials : English

## ABSTRACT

When a Sangir word or prefix ends in a vowel, a following *b*, *d*, *l* or *g* is weakened to respectively *w*, *r*, *ʎ* (retroflexed *l*) and *gh* (velar fricative). This lenition occurs across word boundaries and across morpheme boundaries, but not within roots. The orthography recommendation is to write weakened forms across morpheme boundaries and strong forms across word boundaries.

## TABLE OF CONTENTS

Description of lenition in Sangir; Word rule versus phrase rule; Implications for orthography; Recommendations for testing; Conclusion thus far; References.

## VERSION HISTORY

Version 1 [16 September 2010] This paper initially prepared in August 2007 as part of a course assignment; revised for publication September 2010.

© 2007–2010 by Pamela Day and SIL International  
All Rights Reserved

# Lenition: A challenging issue in Sangir orthography

by  
Pamela Day

The Sangir people originate from a chain of islands in North Sulawesi, Indonesia directly south of Mindanao, Philippines. Their total population is approximately 260,000. Although they have had a written language for over 100 years, there are many inconsistencies relating to a lenition process, word breaks and use of diacritics. This paper focuses on how lenition affects speech and orthography in the Sangir language. It would be good for the Sangir people to agree on how to represent their language consistently in writing and publish a writing manual that can serve as a guide for them. As far as I know, no such manual exists. I want them to be able to create and publish their own literature for generations to come. At this time, there is little printed material in their language. As an outsider, I would like to be better informed from both linguists and from the Sangir themselves. Ultimately, the decision is up to the Sangir on how they want to write their language.

## Description of lenition in Sangir

The rule for lenition in Sangir<sup>1</sup> is that when a word or prefix ends in any vowel except the high central vowel *ɛ̃*, the following consonants *b*, *d*, *l* and *g* are weakened to *w*, *r*, *ʎ* (retroflexed *l*) and *gh* (fricative) respectively. The weakening of *l* to *ʎ* only occurs after the back vowels, *a*, *o* and *u*.

This weakening occurs across word boundaries and across morpheme boundaries, but not within roots.

## Word rule versus phrase rule

The first step toward an orthography recommendation is to ask: is the Sangir lenition process due to a word rule or a phrase rule? Following Snyder (2005), the following table summarizes key differences between word rules and phrase rules. Each distinctive is followed by a statement concerning the lenition process in Sangir (shaded cells) and whether the Sangir data supports a word rule analysis, a phrase rule analysis, or is inconclusive.

---

<sup>1</sup> For a description of Sangir phonology, see Maryott (1961).

<b>Distinctive</b>	<b>Word Rule (Lexical)</b>	<b>Phrase Rule (Post-Lexical)</b>
Exceptions	May or may not have lexical exceptions; if there are exceptions, it is definitely a word rule.	Rules never have exceptions.
No exceptions in Sangir. (inconclusive)		
Phonetic motivation	Might or might not have phonetic motivation; if NO phonetic motivation, definitely a word rule.	Rules are always phonetically motivated.
Always phonetically motivated in Sangir. (inconclusive)		
“Phoneme”-ness of the resulting sound	Output of the rule must be a phoneme.	Output of the rule may or may not be a phoneme. If NOT a phoneme, definitely a phrase rule.
The output is a phoneme. (inconclusive)		
Reference to morpheme boundaries	If the rule only applies across a morpheme boundary, it is definitely a word rule.	Rule never applies <i>only</i> across a morpheme boundary.
The rule applies across both word and morpheme boundaries. (phrase rule)		
Restrictions	If the rule applies only to certain classes of words or morphemes, it is definitely a word rule.	Phrase rules apply to all types of words.
The rule applies to all types of words. (phrase rule)		
Word boundaries	Word rules never apply across word boundaries.	If a rule applies across a word boundary, it <b>MUST</b> be a phrase rule.
The rule applies across both word and morpheme boundaries. (phrase rule)		
Reference to phrase structure	Never needs to make reference to the beginning or end of phrases.	May or may not make reference to the beginning or end of phrases.
No reference to beginning or end of phrases. (inconclusive)		
Mother-tongue speaker awareness	MT speaker is aware of the change resulting from the rule.	MT speaker is usually unaware of the change resulting from the rule.
MT speakers may or may not be aware of the change. (inconclusive?)		

This analysis points in favor of a phrase rule. As a result, words should be spelled with the strong form (*b*, *d*, *l*, and *g*) rather than the weakened form (*w*, *r*, *l* and *gh*) wherever the environment is found (see rule above).

If we always use the strong form in the environment following a vowel, then we come up with the following:

1. *i sie dimadingihě* ‘he heard’ rather than *i sie rimaringihě*
2. *su baļe* ‘in the house’ rather than *su waļe*
3. *Mawu Duata* ‘Lord God’ rather than *Mawu Ruata*
4. *i mama dingangi papaku* ‘my mother and father’ rather than *i mama ringangu papaku*
5. *su daļungu* ‘inside of’ rather than *su raļungu*
6. *nigěli* ‘given’ rather than *nighěli*
7. *nibohe* ‘written’ rather than *niwohe*

In the two years I’ve worked with the Sangir, I’ve learned to use the weakened form across both word boundaries and morpheme boundaries in the environment following a vowel. As a result, the use of the strong form looks strange to me. The ones that look especially strange to me are those found across morpheme boundaries such as in *nigěli* and *nibohe* in (6) and (7) above.

Could it be that there is a word rule that applies across morpheme boundaries and a phrase rule that applies across word boundaries? If so, then we would use the weakened forms (*w*, *r*, *l* and *gh*) across morpheme boundaries and the strong forms (*b*, *d*, *l*, and *g*) across word boundaries.

Another curious observation is that both the strong and weakened form may occur following vowels within roots. Consider the following root words:

Strong forms:

1. *mapulu* ‘want, like,’ not *mapuļu*
2. *abe* ‘don’t!’ not *awe*
3. *meda* ‘table,’ not *mera*
4. *kakěndagě* ‘love’

Weak forms:

1. *mapuļo* ‘ten’
2. *Mawu* ‘Lord’
3. *arau* ‘or’
4. *taghuang* ‘owner’

If we state that this lenition is due to a phrase rule, then it should also apply within roots. If it is only a word rule, it may or may not apply to roots.

So, let’s look at the chart again in two different ways. First let’s consider only what is happening across morpheme boundaries.

## ACROSS MORPHEME BOUNDARIES

<b>Distinctives</b>	<b>Word Rule (Lexical)</b>	<b>Phrase Rule (Post-Lexical)</b>
Exceptions	May or may not have lexical exceptions; if there are exceptions, it is definitely a word rule.	Rules never have exceptions
Exceptions within Sangir roots, but not across morpheme boundaries. (word rule)		
Phonetic motivation	Might or might not have phonetic motivation; if NO phonetic motivation, definitely a word rule.	Rules are always phonetically motivated.
Phonetically motivated in Sangir. (inconclusive)		
“Phoneme”-ness of the resulting sound	Output of the rule must be a phoneme.	Output of the rule may or may not be a phoneme. If NOT a phoneme, definitely a phrase rule.
The output is a phoneme. (inconclusive)		
Reference to morpheme boundaries	If the rule only applies across a morpheme boundary, it is definitely a word rule.	Rule never applies only across a morpheme boundary.
We are only considering cases where lenition occurs across a morpheme boundary within words. (not applicable)		
Restrictions	If the rule applies only to certain classes of words or morphemes, it is definitely a word rule.	Phrase rules apply to all types of words.
Rule applies to all types of words. (inconclusive)		
Word boundaries	Word rules never apply across word boundaries.	If a rule applies across a word boundary, it MUST be a phrase rule.
We are only considering cases where lenition occurs across a morpheme boundary within words. (not applicable)		
Reference to phrase structure	Never needs to make reference to the beginning or end of phrases.	May or may not make reference to the beginning or end of phrases.
No reference to beginning or end of phrases. (inconclusive)		
Mother-tongue speaker awareness	MT speaker is aware of the change resulting from the rule.	MT speaker is usually unaware of the change resulting from the rule.
MT speaker is aware of the change resulting from the rule. (word rule)		

This analysis points in favor of a word rule. As a result, words should be spelled with the weakened form (*w*, *r*, *l* and *gh*) rather than the strong form (*b*, *d*, *l* and *g*) wherever the environment is found (see rule above) across morpheme boundaries.

In some recent emails from a 28-year-old Sangir friend, Deice, in Indonesia, I have the following examples of using the weakened form across morpheme boundaries.<sup>2</sup> This supports the idea of a word rule for this environment.

1. *ni weliang u* ‘bought by’
2. *ni wohe* ‘written’
3. *nigheli u* ‘sent by’

Now let’s consider what’s happening across word boundaries.

#### ACROSS WORD BOUNDARIES

Distinctives	Word Rule (Lexical)	Phrase Rule (Post-Lexical)
Exceptions	May or may not have lexical exceptions; if there are exceptions, it is definitely a word rule.	Rules never have exceptions
No exceptions when the lenition occurs across word boundaries. (inconclusive)		
Phonetic motivation	Might or might not have phonetic motivation; if NO phonetic motivation, definitely a word rule.	Rules are always phonetically motivated.
Phonetically motivated in Sangir. (inconclusive)		
“Phoneme”-ness of the resulting sound	Output of the rule must be a phoneme.	Output of the rule may or may not be a phoneme. If NOT a phoneme, definitely a phrase rule.
The output of the rule is a phoneme. (inconclusive)		
Reference to morpheme boundaries	If the rule only applies across a morpheme boundary, it is definitely a word rule.	Rule never applies only across a morpheme boundary.
Rule applies across word boundaries. (phrase rule)		

---

<sup>2</sup> Please note: word break are another issue in Sangir orthography. Sometimes Deice wrote the verbs as one word and sometimes as two words. They are one word and thus the context is across morpheme, not word, boundaries, Note also that diacritics are omitted in this data set because Deice, like other Sangir, doesn’t use diacritics in emails and text messages.

Restrictions	If the rule applies only to certain classes of words or morphemes, it is definitely a word rule.	Phrase rules apply to all types of words.
The word applies to all types of words. (inconclusive)		
Word boundaries	Word rules never apply across word boundaries.	If a rule applies across a word boundary, it MUST be a phrase rule.
Rule applies across word boundaries. (phrase rule)		
Reference to phrase structure	Never needs to make reference to the beginning or end of phrases.	May or may not make reference to the beginning or end of phrases.
No reference to beginning or end of phrase. (inconclusive)		
Mother-tongue speaker awareness	MT speaker is aware of the change resulting from the rule.	MT speaker is usually unaware of the change resulting from the rule.
MT speaker is confused about when to write the weakened form or not; thus, unaware. (phrase rule)		

The orthography recommendation is to experiment with writing the strong form in this environment (across word boundaries) and never the weakened form.

In her emails, Deice usually uses the strong form across word boundaries, but sometimes the weakened form. Her seeming confusion and inconsistency reveal that she is unaware of the change reflected in spoken speech. This supports the idea of a phrase rule for this environment.

1. *ia mengendung lai bou Mawu* (Deice email, 2/8/2007)
  - *bou* not *wou*
2. *mengimang si sie su dalungu pe bawiahe* (Deice email, 2/8/2007)
  - *su dalungu* not *su ralungu*
  - *pe bawiahe* not *pe wawiahe*
3. *Ia mehale su waleng* (Deice email, 5/21/2007)
  - *su waleng* (weakened form written here)
4. *si kau dingangu dumaringihe habaru*
  - *si kau dingangu dumaringihe* not *ringangu rumaringihe*

## Implications for orthography

After applying the chart individually to what's happening across morpheme boundaries and then to what's happening across word boundaries, it seems that there are two rules. There is a word rule that applies across morpheme boundaries and a phrase rule that applies across word boundaries. As a result, the orthography recommendation is to use the weakened forms (w, r, l and gh) across morpheme boundaries and the strong forms (b, d, l, and g) across word boundaries. An example of this application is found in the following sentence:

*I sie dimaringihë tingih'u Mawu* 'He heard the voice of the Lord.'

In this sentence, the strong form *d* is used word-initially and the weakened form *r* is used word-medially, following the phrase rule and word rule respectively.

This is the linguistic recommendation, but life is not purely linguistic! ☺ There are also sociolinguistic factors at play. The Sangir have used the weakened forms for over a hundred years in both contexts. The Sangir Bible done by Lembaga Alkitab Indonesia (Indonesian Bible Society), printed as recently as 2003 and 2009, also writes weakened forms across word boundaries. I can inform the Sangir of the challenges, options and rules, but they themselves must decide what works best for them.

## Recommendations for testing

It would also be good to do some testing. Is it easier for Sangir to read and write with the weakened form or the strong form across word boundaries? We could also test the strong or weakened form across morpheme boundaries, but this is not as great an issue since there is much consistency in using only the weakened form in this context and this agrees with the word rule.

What kind of tests can be done?

### 1. Reading Test

Select a story in Sangir of 400–500 words. If it takes ten minutes to read the story, give the reader only five minutes to read the story and then mark how far he got. Then let him read the rest of the story. No comprehension questions are needed.

Have two groups of Sangir speakers, about 50 in each group. Have one group test the strong form version of the story and the other test the weakened form version of the story.

Enter the results into an MS Excel sheet and use a T-test to calculate whether the difference between the two groups is significant.

## 2. Writing Test

For a writing test, record a Sangir speaker reading a story sentence by sentence. Then use this recording for dictation, asking the test participants to write down what they hear. Do they use the strong or weakened form across word boundaries?

In choosing test participants, it's good to have a mix of both literate and non-literate. The Sangir high school students at the Indonesian Consulate school in Davao would be a good choice for test participants. They speak Sangir, but have not yet been taught to read and write in Sangir. They may know how to read and write in Indonesian or Cebuano or English or Tagalog.

## Conclusion thus far

I want to say a special thanks to both Keith Snider and David Weber for their linguistic guidance and insight into these orthography issues. I also want to say thank you to Elke Karan and Diana Weber who encouraged me to dig deeper into this topic. Thanks also to David Mead who pointed out my typographical errors and improved the linguistic reasoning of the paper. And I'm thankful for the linguistic research of Ken Maryott who worked with the Sangir for about 40 years. His papers have been very helpful.

Initially, the word rule and phrase rule issue was very confusing to me. Once I applied it to Sangir, I began to understand it, with much help from Keith and David. Keith also gave me some ideas for testing the application of the phrase rule to Sangir orthography and how to do a T-test. (I am very intimidated by testing and statistics!!)

It would be good to discuss lenition and other orthography topics with some Sangir educational leaders in Sangir. There's so much more I have to learn. This is truly on-going research. I can share with the Sangir my analysis, results and observations, but they will be the ones to choose how to write their language.

## References

Maryott, Kenneth R. 1961. The phonology and morphonemics of Tabukang Sangir. *Philippine Social Sciences and Humanities Review* 26:111–126.

Synder, Keith. 2005. You can have your cake and eat it, too: An orthography that meets the needs of both mature and immature readers. Unpublished lecture notes, 9 pp. SIL-University of North Dakota.