

Onset Cluster Typologies: An Appendix

Stephen Jay and Steve Parker

Dallas International University

October 26, 2023

Table of Contents

Introduction.....	1
OG only	2
Màwés Aas'è (Bambassi)	2
OG NG.....	4
Angaataha	4
OG NG LG	6
Ga'dang	6
OG NG LG GG.....	8
Shilluk.....	8
OL only	10
Tshangla.....	10
OL NL.....	12
Isirawa.....	12
OG OL	14
Southern Zazaki	14
OG NG OL	16
(Ban Lua) Bisu.....	16
OG NG LG OL	18
Phula (Phukha).....	18
OG OL NL.....	21
Bumthangkha.....	21
Dakpakha	24
OG NG OL NL	26
Khengkha.....	26
OG NG LG OL NL.....	28
Orya	28
Abbreviations.....	31
References.....	32

Introduction

The following is the appendix to a 2020 paper (https://www.researchgate.net/publication/349926320_Onset_Cluster_Typologies) highlighting the Minimum Distance to Offset (MDO) model which accounts for the distribution of onset clusters cross-linguistically. This appendix contains a representative language (or two) for each of the 11 language types that the MDO model predicts, restricted here mainly to rising clusters, not plateaus or sonority reversals.

Each entry in this appendix is organized as follows. First we begin with an overview of the language. The demographic information comes from the Glottolog (www.glottolog.org) and the Ethnologue (www.ethnologue.com), which will be referred to simply by those names, with bibliographic information provided at the end of the appendix. The overview section also provides the maximum syllable template, an inventory of all consonant phonemes, and a list of all attested syllable-initial clusters. Phonemes or clusters which appear within parentheses are marginal, rare, or only occur in loan words. Then we provide examples of the attested onset clusters in that particular language. Finally, any pertinent notes about the interpretation of segments or syllable structure are included in order to give the reader as much information as possible. The references for each language are then listed twice: once at the end of the respective language page, as well as in the full bibliography at the end of this appendix.

Exhaustive list of onset cluster types present: OG only

Màwés Aas'è (Bambassi)

Language Overview

		Source(s)
Language Name	Màwés Aas'è	Ethnologue
	Bambassi	Glottolog
Code	ISO 639-3: myf	Ethnologue
	Glottocode: bamb1262	Glottolog
Classification	Afro-Asiatic, Omotic, North, Mao	Ethnologue
	Blue Nile Mao, Bambassi	Glottolog
Number of Speakers	2,300	Ethnologue
Alternate Names	Amam, Bambassi, Bambeshi, Bambessi Mao, Didessa (Mao), Fadiro, Mao of Bambeshi, Mao of Diddesa, Mawes Aasse, Northern Mao, Siggoyo	Ethnologue and Glottolog
Location	Ethiopia (Benishangul-Gumuz region and Oromia region)	Ethnologue
	9.83 N, 34.65 E	Glottolog
Maximum Syllable Template	[CGVV(N)C] The complex coda (NC) is rare, only occurring in three monomorphemic forms (Ahland 2009, 18).	Ahland 2009, 22
Inventory of Consonant Phonemes (21)	p t k p' t' k' b d g (tʃ) ts' s ʃ h z m n ŋ l r j w	Ahland 2009, 3
Syllable-Initial Clusters (12)¹	Word-Initial (12): pj kj p'j k'j gj tw kw t'w k'w gw sw ʃw Word-Medial (4): kj p'j kw k'w	Ahland 2009, 20-21

¹ The cluster type [sw] was not included in Ahland's (2009; 2012) list of clusters, but was confirmed by Parker (2012, 140), through personal communication with Ahland.

Examples of Onset Clusters

OG initial:	[gja:je,HL] ‘many’	(Ahland 2009, 24)
	[twaŋ.gi.le,MLL] ‘elephant’	(Ahland 2009, 42)
OG medial:	[ze.p’je,LL] ‘together’	(Ahland 2009, 37)
	[ha.kjam.ba,MHM] ‘to hunt’	(Ahland 2009, 26) ²

Other Notes

Ahland (2009) discusses the ambiguity of OG clusters and concludes that these should be interpreted as CC clusters, rather than either as CVV syllables, or as single palatalized (Cⁱ) and labialized (C^w) consonants. Ahland argues that interpreting them as Cⁱ and C^w would add 11 consonants to the inventory, that non-geminate VV sequences do not occur in unambiguous monomorphemic words, and that there is otherwise no evidence for diphthongs (21). Ahland also notes that these clusters are more common word-initially than word-medially, and “consonant distribution and syllable structure suggest they must be seen as onset clusters” (21). However, while the first two arguments are clearly supported by the data in Ahland (2009, 2012), no examples are given in either source to illustrate this latter argument in monomorphemic words. Parker (2012, 141-142) adds the evidence that OG clusters co-occur with a trisegmental rhyme (such as VVC). If these clusters were interpreted as CV, it would result in a four-segment rhyme, which is otherwise unattested.

References

- Ahland, Michael. 2009. “Aspects of Northern Mao (Bambassi-Didessa) Phonology.” *Linguistic Discovery* 7: 1-40. <https://doi.org/10.1349/PS1.1537-0852.A.332>.
- Ahland, Michael Bryan. 2012. “A grammar of Northern Mao (Màwés Aas'è).” PhD diss., University of Oregon. <https://cla.csulb.edu/departments/linguistics/wp-content/uploads/2016/01/Ahland-2012.pdf>.
- Parker, Steve. 2012. “Sonority Distance vs. Sonority Dispersion: A Typological Survey.” In *The Sonority Controversy*, edited by Steve Parker, 101-165. Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110261523>.

² This example is polymorphemic: [ha-kjamb-a] ‘AFF-hunt-DECL’ (Ahland 2012: 91). The preceding example [ze.p’je,LL] ‘together’ was the only monomorphemic example of word-medial OG found in either Ahland (2009) or Ahland (2012).

Exhaustive list of onset cluster types present: OG NG

Angaataha

Language Overview

		Source(s)
Language Name	Angaataha	Ethnologue and Glottolog
Code	ISO 639-3: agm	Ethnologue
	Glottocode: anga1290	Glottolog
Classification	Trans-New Guinea, Angan, Angaatiha	Ethnologue
	Angan, Angaataha	Glottolog
Number of Speakers	2,500	Ethnologue
Alternate Names	Angaatiha, Angaataha, Angaatiya, Angataha, Langimar	Ethnologue and Glottolog
Location	Papua New Guinea (Menyamya district in Morobe province)	Ethnologue
	7.15 S, 146.35 E	Glottolog
Maximum Syllable Template	[CGVV]	Huisman, Huisman, and Lloyd 1981, 51
Inventory of Consonant Phonemes (11)	p t k ʔ tʃ m n ŋ r j w	Huisman, Huisman, and Lloyd 1981, 52
Syllable-Initial Clusters (4)	nj pw kw (ʔw) mw All four main clusters may occur word-initially or word-medially. The rare [ʔw] may occur word-medially, but it is unclear if it also occurs word-initially.	Huisman, Huisman, and Lloyd 1981, 53

Examples of Onset Clusters

- OG initial: [pwə.tʃə.pà.tə] ‘weapon’ (Huisman, Huisman, and Lloyd 1981, 53)
[kwî.pé.ʔə] ‘seeds from grass that stick’ (Huisman, Huisman, and Lloyd 1981, 53)
- OG medial: [nà.pú.pwè] ‘close it’ (Huisman, Huisman, and Lloyd 1981, 53)

	[àu.ṇ.kwò.wài.ʔò] ‘thumb’	(Huisman, Huisman, and Lloyd 1981, 53)
	[á.ʔwì.pá.tò] ‘trap type’	(Huisman, Huisman, and Lloyd 1981, 53)
NG initial:	[mwó.tʃḗ.ʔà] ‘axe’	(Huisman, Huisman, and Lloyd 1981,)
	[njà.njé.ʔà] ‘flower type’	(Huisman, Huisman, and Lloyd 1981, 53)
NG medial:	[mà.njí.njái] ‘children (objective form)’	(Huisman, Huisman, and Lloyd 1981, 54)

Other Notes

Huisman, Huisman, and Lloyd (1981, 53) provide the following evidence for a glide analysis of the high vocoids. First, the glide offsets do not carry tone, precluding a CV.V interpretation. Second, glides do occur as single consonants elsewhere, again not carrying tone, as in [já.pè.pà] ‘ground’ (Huisman, Huisman, and Lloyd 1981, 56). Because of this, a glide analysis does not add extra phonemes to the inventory. Finally, the addition of labialized and palatalized consonants would add four extra consonants to the inventory. One final piece of evidence not provided by Huisman, Huisman, and Lloyd (1981), but supported by the data therein, is that CG onsets co-occur in syllables with VV peaks, as in [kwài.pè.ʔò] ‘tree type’ (Huisman, Huisman, and Lloyd 1981, 52). If the glides were instead analyzed as vowels, there would be three vowels in the nucleus, which is rare cross-linguistically.

Consonant clusters of the combination NO also occur at the beginning of words, but those nasals are syllabic and carry their own tone (Huisman, Huisman, and Lloyd 1981, 53). For this reason, they are not considered here as onset clusters, leaving the possible types as OG and NG.

Huisman and Lloyd (1981) point out three tones in Angaataha, marking High as [á], High-Low as [â], and Low as unmarked [a]. It appears that Huisman, Huisman, and Lloyd (1981) followed the same pattern. For clarity, the low vowels that are unmarked in both of those works are marked as [à] in the examples contained here.

References

- Huisman, Ronald, Roberta Huisman, and Joy Lloyd. 1981. “Angaatiha Syllable Patterns.” In *Angan Languages are Different*, edited by Phyllis M. Healey, 51-62. Huntington Beach, CA: Summer Institute of Linguistics.
- Huisman, Ronald and Joy Lloyd. 1981. “Angaatiha Tone, Stress, and Length.” In *Angan Languages are Different*, edited by Phyllis M. Healey, 63-82. Huntington Beach, CA: Summer Institute of Linguistics.

Exhaustive list of onset cluster types present: OG NG LG

Ga'dang³

Language Overview

		Source(s)
Language Name	Ga'dang	Ethnologue and Glottolog
Code	ISO 639-3: gdg	Ethnologue
	Glottocode: gada1258	Glottolog
Classification	Austronesian, Malayo-Polynesian, Northern Luzon, Northern Cordilleran, Cagayan Valley, Ibanagic, Gaddangic	Ethnologue
	Austronesian, Malayo-Polynesian, Northern Luzon, Cagayan Valley, Ibanagic, Gaddangic, Cagayan-Baliwon Gaddang	Glottolog
Number of Speakers	6,000	Ethnologue
Alternate Names	Baliwon, Gaddang, Ginabwal	Ethnologue and Glottolog
Location	Philippines (Ifugao province, Kalinga province, and Mountain province, all within Cordillera Administrative Region)	Ethnologue
	17.22 N, 121.45 E	Glottolog
Maximum Syllable Template	[CGVC]	Troyer 1959, 99
Inventory of Consonant Phonemes (16)	p t k ʔ b d g ɸ s m n ŋ r l j w	Troyer 1959, 96-98
Syllable-Initial Clusters (17)	pj tj kj bj dj gj nj ŋj lj tw kw bw dw sw mw nw lw	Troyer 1959, 97

³ Note that there is a sister language listed in the Ethnologue and Glottolog as Gaddang (ISO 639-3: gad). The Ethnologue notes that the two languages have 80% lexical similarity, and less than 80% mutual intelligibility. Based on comments by Troyer (1959, 95, footnote 1), it appears that the language analyzed in Troyer 1959 is the one listed as Ga'dang (ISO 639-3: gdg) in the Ethnologue and Glottolog. As all of the language data here come from one source (Troyer 1959), this distinction is not crucial for our purposes.

Examples of Onset Clusters

OG initial:	['kwi] 'to'	(Troyer 1959, 97)
	['kjaw] 'bow wow (of a small puppy)'	(Troyer 1959, 97)
OG medial:	['mid.djob] 'will blow'	(Troyer 1959, 96)
	[mas.'swel.do] 'will pay wages'	(Troyer 1959, 96)
NG initial:	['ɲjaw] 'meow'	(Troyer 1959, 97)
NG medial:	['pin.ɲjə] 'pineapple'	(Troyer 1959, 97)
LG medial:	[mal.'lwag] 'boiling'	(Troyer 1959, 97)

Other Notes

Troyer (1959, 96-97) acknowledges the ambiguity of glide offsets, and provides the following evidence for these to be analyzed as CGV, rather than CVV or labialized/palatalized consonants (i.e. C^wV). First, there are no non-suspect vowel sequences in Gaddang. Second, these CGV clusters are realized as one syllable phonetically, precluding a CV.V analysis. Finally, if these clusters were analyzed as labialized/palatalized, the abundance of these offsets would result in more than double the current consonant inventory.

Although many of the examples in Troyer (1959) involve geminates, these geminates “are made up of re-articulated phonemes when pronounced syllable by syllable, but are long phonemes in fast speech” (Troyer 1959, 6, footnote 6). Parker (2012, 130) notes that the only examples that do not involve geminates in Troyer (1959) are OG clusters.

References

- Parker, Steve. 2012. “Sonority Distance vs. Sonority Dispersion: A Typological Survey.” In *The Sonority Controversy*, edited by Steve Parker, 101-165. Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110261523>.
- Troyer, Madeline. 1959. “Gaddang Phonology.” *The Philippine Journal of Science* 88 (1): 95-102. https://philjournalsci.dost.gov.ph/images/pdf_upload/pjs1959/PJS_Vol_88_No1_Mar_1959.pdf.

Exhaustive list of onset cluster types present: OG NG LG GG

Shilluk

Language Overview

		Source(s)
Language Name	Shilluk	Ethnologue and Glottolog
Code	ISO 639-3: shk	Ethnologue
	Glottocode: shil1265	Glottolog
Classification	Nilo-Saharan, Satellite-Core, Core, Eastern Sudanic, Southern (n languages), Nilotic, Western, Luo, Northern, Shilluk	Ethnologue
	Nilotic, Western Nilotic, Lwoo, Northern Lwoo	Glottolog
Number of Speakers	574,000	Ethnologue
Alternate Names	Chulla, Colo, Dhocolo, Shulla	Ethnologue and Glottolog
Location	South Sudan (Upper Nile state)	Ethnologue
	9.85 N, 31.79 E	Glottolog
Maximum Syllable Template⁴	[CGVVC]	Gilley 1992, 99
	[CGVVVC]	Remijsen, Ayoker, and Mills 2011, 111-112
Inventory of Consonant Phonemes (19)⁵	p t̪ c k b d̪ g m n̪ ŋ l r j w	Remijsen, Ayoker, and Mills 2011, 113; Gilley 1992, 23
Syllable-Initial Clusters (31)	pj t̪j tj cj kj bj d̪j dj j̪j gj mj n̪j ŋj lj rj pw t̪w tw cw kw bw d̪w dw ɣw gw mw nw n̪w ŋw lw jw	Gilley 1992, 22-25

⁴ Gilley (1992) only mentions two vowel lengths, short and long, while Remijsen, Ayoker, and Mills (2011) reference a three-way length difference for vowels. Either way, the 2-3 vowels in the nucleus are composed of one vowel quality with separate lengths, not diphthongs (Gilley 1992, 24; Remijsen, Ayoker, and Mills 2011, 118).

⁵ Remijsen, Ayoker, and Mills (2011, 113) analyze the rhotic as a trill (IPA /r/), while Gilley (1992, 22-23) analyzes the rhotic as a flap (IPA /r̩/).

Examples of Onset Clusters⁶

OG initial:	[djêl] ‘goat.SG’	(Remijsen, Ayoker, and Mills 2011, 112)
	[cwāk] ‘type of fish’	(Gilley 1992, 25)
OG medial:	[ó.tjêṃ] ‘dragonfly’	(Gilley 1992, 25)
	[ó.twēl] ‘type of fish’	(Gilley 1992, 25)
LG initial:	[lwôl] ‘open gourd.SG’	(Remijsen, Ayoker, and Mills 2011, 112)
	[ljēc] ‘elephant’	(Gilley 1992, 25)
LG medial:	[á.lwê:.dṣ] ‘crab’	(Gilley 1992, 53) ⁷
	[o.lwi.lwi] ‘type of duck’	(Heasty 1937, 77) ⁸
NG initial:	[ɲwɛ.lɔ] ‘earthworm’	(Gilley 1992, 25) ⁹
	[ɲjɛl] ‘to trundle’	(Gilley 1992, 25)
NG medial:	[ò.nwà:.ɲò] ‘black ants’	(Gilley 1992, 25)
	[à.mjél] ‘stubborn’	(Gilley 1992, 25)
GG initial:	[jwét] ‘defile.DVNA’	(Remijsen, Ayoker, and Mills 2011, 112)
	[jwòt] ‘flying termites’	(Gilley 1992, 25)

Other Notes

Gilley (1992, 24-26) addresses the interpretation issue of the high glides [j] and [w], giving two main pieces of evidence for analyzing them as CG. First, unambiguous vowel clusters (of differing vowel qualities) do not occur in Shilluk. Second, these clusters are quite productive, with only a few gaps. If these clusters were analyzed as palatalized or labialized consonants, the consonant inventory would multiply greatly.

References

- Gilley, Leoma. 1992. *An Autosegmental Approach to Shilluk Phonology*. Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics 103. Dallas: Summer Institute of Linguistics.
- Remijsen, Bert, Otto G. Ayoker, and Timothy Mills. 2011. “Shilluk.” *Journal of the International Phonetic Association* 41 (1): 111-125.
<https://doi.org/10.1017/S0025100310000289>.

⁶ Gilley (1992) and Remijsen, Ayoker, and Mills (2011) both list 10 phonemic vowels, but represent them differently. Gilley (1992, 28-29) uses /i e a ɔ o/ for [-ATR] vowels, and the underlined /i e a ɔ o/ for [+ATR] vowels, noting that ATR may not be the best phonetic distinction, instead invoking the feature ‘Expanded Pharynx.’ Remijsen, Ayoker, and Mills (2011, 116) describe the vowel inventory as [-ATR] /i e a ɔ o/ and [+ATR] /i e ʌ o u/. For consistency, the vowel inventory from Remijsen, Ayoker, and Mills (2011) is used for these examples, with the examples from Gilley (1992) converted to this format.

⁷ The tone on the second syllable of this word is represented in Gilley (1992) as a High-Mid contour.

⁸ This word is from a Shilluk-English dictionary, which uses the orthographical form, and does not mark tone. Regardless, this word is clearly the phonetic form listed here.

⁹ This word, among others, does not have a tone marking in Gilley (1992). The reason for this is unclear.

Exhaustive list of onset cluster types present: OL only

Tshangla

Language Overview

		Source(s)
Language Name	Tshangla	Ethnologue and Glottolog
Code	ISO 639-3: tsj	Ethnologue
	Glottocode: tsha1245	Glottolog
Classification	Sino-Tibetan, Tibeto-Burman, Western Tibeto-Burman, Bodish	Ethnologue
	Sino-Tibetan, Bodic, Tshanglic	Glottolog
Number of Speakers	181,200	Ethnologue
Alternate Names	(Canglo, Cangluo, Central, Motuo) + Menba, Monba, Monpa; Changla, Sangla, Sarchapkkha, Shachobiikha, Shachopkha, Sharchagpakha, Sharchhokpa, Sharchhop, Sharchokpa-lo, Tsangla, Tsanglo, Tschanglo, Tshalingpa	Ethnologue and Glottolog
Location	Bhutan (Samdrup Jongkhar and Trashigang) China (Xizang Autonomous Region) India (Arunachal Pradesh and Assam)	Ethnologue
	27.37 N, 91.77 E	Glottolog
Maximum Syllable Template	[CLVX]	Andvik 2010, 14
Inventory of Consonant Phonemes (29)	p t t̥ k p ^h t ^h t̥ ^h k ^h b d d̥ g t̥s̥ t̥s̥ ^h t̥s̥ ^h (dz) dz s ɕ z (z) m n ɲ ŋ l (l̥) r j w h	Andvik 2010, 8-12
Syllable-Initial Clusters (6)¹⁰	pr p ^h r br kr k ^h r gr	Andvik 2010, 14

¹⁰ Andvik (2010, 15) recognizes one OO cluster, [pe], which is only found in two lexical items: [pei] ‘four’, and [pei p^hule] ‘to pass gas’. It is not included here due to its rarity.

Examples of Onset Clusters

OL initial:	[p ^h raŋ.ga] ‘underneath’	(Andvik 2010, 14)
	[bra] ‘other’	(Andvik 2010, 14)
	[gre.p ^h e] ‘to crush, to crack’	(Andvik 2010, 14)
	[krem.ta.la] ‘lean, thin’	(Andvik 2010, 14)

Other Notes

Andvik’s (2010) main focus is on the Trashigang dialect (often considered the “standard” Tshangla dialect). Andvik (2010, 14) notes that the Trashigang dialect pronounces the Cr clusters as one retroflex coronal consonant, but that other dialects preserve the Cr clusters. The examples given here are referenced as examples from non-Trashigang dialects.

Tshangla also has word-medial /C+I/ clusters, but only in morphologically derived environments, such as [got.la] ‘is looking’ (Andvik 2010, 67). None of these examples contain three successive consonants, so the syllable break may occur between the two consonants. Similarly, a consonant may be followed by the cluster /ui/, but this only occurs across morpheme boundaries, as in /bu-i/ ‘take-imperative’ (15). Parker (2012, 153) confirmed with Eric Andvik that the stress in these instances falls on the /u/ rather than /i/, preserving penultimate stress, and that this sequence does not occur in monomorphemic contexts. Consequently, there are no clear cases of glide offsets in Tshangla.

Finally, Andvik (2010, 15) recognizes one OO cluster, [pɛ], which is only found in two lexical items: [pɛi] ‘four’, and [pɛi p^hule] ‘to pass gas’. It is not included here due to its rarity.

References

- Andvik, Erik. 2010. *A Grammar of Tshangla*. Boston: Brill.
- Parker, Steve. 2012. “Sonority Distance vs. Sonority Dispersion: A Typological Survey.” In *The Sonority Controversy*, edited by Steve Parker, 101-165. Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110261523>.

Exhaustive list of onset cluster types present: OL NL

Isirawa

Language Overview

		Source(s)
Language Name	Isirawa	Ethnologue and Glottolog
Code	ISO 639-3: srl	Ethnologue
	Glottocode: isir1237	Glottolog
Classification	Tor-Kwerba, Greater Kwerba, Isirawa	Ethnologue
	Greater Kwerba	Glottolog
Number of Speakers	1,800	Ethnologue
Alternate Names	Okwasar, Saberi, Saweri, Sawuri-Hablifuri	Ethnologue and Glottolog
Location	Indonesia (Papua province)	Ethnologue
	1.89 S, 138.59 E	Glottolog
Maximum Syllable Template	[CCVVVC] Onset is obligatory, except word-initially. The three-vowel nucleus only occurs word finally.	Oguri and Erickson 1975, 53-54
Inventory of Consonant Phonemes (12)	p t k tʃ f s (h) β m n r j w	Oguri and Erickson 1975, 43-48
Syllable-Initial Clusters (7)	pr tr kr fr sr βr mr	Oguri and Erickson 1975, 53

Examples of Onset Clusters

OL initial:	[pran. 'ka] 'cooking pot'	(Oguri and Erickson 1975, 40)
	['sri.fa] 'language'	(Oguri and Erickson 1975, 40)
	[tri] 'thunder'	(Oguri and Erickson 1975, 56)
	[krai krai] ~ [kʰrai kʰrai] 'bird type'	(Oguri and Erickson 1975, 40)
OL medial:	[a. 'prɛ.sa] 'lower arm'	(Oguri and Erickson 1975, 56)
	[wa.wa.fri. 'a.na] 'to wear'	(Oguri and Erickson 1975, 57)
	['naβ.sra] 'three'	(Oguri and Erickson 1975, 56)

NL initial:	[mri] ~ [m ^h ri] ‘one’	(Oguri and Erickson 1975, 40)
	[mra. 'ri.ma.na.pi] ‘bathed’	(Oguri and Erickson 1975, 63)
NL medial:	[fi.ta. 'mra] ‘corpse’	(Oguri and Erickson 1975, 56)

Other Notes

The phoneme /h/ is marginal, and often fluctuates with /s/. The liquid /r/ appears as a trill [r] when next to another consonant (either before or after), but as a flap [ɾ] elsewhere. Oguri and Erickson (1975, 40) note that some syllable-initial clusters occur with a non-syllabic transitional vocoid between the consonants, as [C^oC]. However, a few pieces of evidence point to this sound not being part of the underlying form, for a couple of reasons. First, [pr] never contains this transitional vowel (40). Second, stress typically appears in the penultimate syllable; however, this transitional vowel is never stressed, even when it would otherwise appear in the penultimate syllable of the word (40).

Vowel clusters containing [i] or [u] do occur in Isirawa, as in [sua] ‘insect type’ (Oguri and Erickson 1975, 41). However, Oguri and Erickson (1975, 41) analyze these clusters as complex syllable nuclei, not vowel glides or multiple syllables, for the following reasons. First, unambiguous vowel clusters occur in the data, as in [sɪ.sɾeəɔ. 'fani] ‘place name’ (41). Second, these vowels are part of one mora timing (41). Third, when two vowels in separate syllables are adjacent, a glide is inserted between them; this does not occur when they are part of the same syllable (41-42). Fourth, stress may fluctuate between speakers from one vowel of these clusters to another (42). Finally, the earliest place that stress may occur in a word is on the antepenultimate syllable; analyzing these as separate syllables would create a new and unnecessary generalization about stress (42).

References

Oguri, Hiroko, and Carol Erickson. 1975. “A Tentative Phonology of Isirawa.” *Irian: Bulletin of Irian Jaya Development* 5(3): 38-66. <http://www.papuaweb.org/dlib/irian/4-1.PDF>.

Exhaustive list of onset cluster types present: OG OL

Southern Zazaki

Language Overview

		Source(s)
Language Name	Southern Zazaki	Ethnologue
	Dimli	Glottolog
Code	ISO 639-3: diq	Ethnologue
	Glottocode: diml1238	Glottolog
Classification	Indo-European, Indo-Iranian, Iranian, Western, Northwestern, Zaza-Gorani	Ethnologue
	Indo-European, Classical Indo-European, Indo-Iranian, Iranian, Central Iranian PBS, Central Iranian PB, Northwestern Iranian, Adharic, Zaza	Glottolog
Number of Speakers	1,280,000	Ethnologue
Alternate Names	Dimili, Dimilî, (Southern) Zaza, Zazaca, Zazaki	Ethnologue and Glottolog
Location	Turkey (Bingöl, Diyarbakir, and Elazig provinces)	Ethnologue
	38.61 N, 39.64 E	Glottolog
Maximum Syllable Template	CCVC	Todd 2008, 25
Inventory of Consonant Phonemes (28)	(tʰ) p t k q b d g tʃ dʒ (sʰ) f s ʃ x h ħ v z ʒ ɣ ʕ m n l lʷ r r j w	Todd 2008, 2
Syllable-Initial Clusters (8)¹¹	pl kl pr tr xr pj qw zw	Todd 2008, 24-26

Examples of Onset Clusters

OG initial: [pja] ‘together’ (Todd 2008, 24)
 [zwa] ‘dry’ (Todd 2008, 25)

¹¹ Todd (2008) does not provide a canonical list of onset clusters, only noting that it must be a fricative or voiceless stop followed by either a liquid or glide (24-25). For this reason, this list may not be exhaustive, although it is still restricted to OG and OL clusters.

OL initial:	[klit] ‘key’	(Todd 2008, 24)
	[prɛn] ‘dress’	(Todd 2008, 24)
	[xrav] ‘bad’	(Todd 2008, 25)
OL medial:	[is. 'tri] ‘horn’	(Todd 2008, 26)

Other Notes

The pharyngealized stop /tˤ/ occurs “mostly, if not exclusively, in Arabic loanwords” (Todd 2008, 4). The pharyngealized fricative /sˤ/ is rare, but contrastive in words that are not traced to Arabic (7).

Although Todd (2008) does not directly address the interpretation of CG clusters, such analysis is warranted based on several pieces of evidence. First, the phonemes /j/ and /w/ occur in unambiguous consonant positions, such as [ojo] ‘he’ (9) and [wɛr. 'di] ‘small’ (10). Second, the canonical syllable structure CCVC is established through CL clusters, so CG clusters do not expand the syllable template. Conversely, an analysis as CV would unnecessarily expand the syllable template to CCVVC, adding diphthongs¹² to the possible nuclei. Third, the glides occur after multiple different obstruents; as such, analyzing them as C^j or C^w would add many more phonemes to the inventory.

However, Todd (2008, 31) notes that native speakers insisted that there is a short [i] separating initial consonant clusters, such as /bra/ [bira] ‘brother.’ Todd hypothesizes that the speakers are “reacting to the voicing of the stop which continues throughout the cluster” (31). The editor, Brigitte Werner, suggests that these CC clusters are two syllables separated by the latent vowels /i/ or /ʊ/, so that /klit/ ‘key’ would be realized as [ki. 'lit] (24, footnote a). Werner notes that writers often write these as two syllables, but shorten them to one syllable when speaking.

References

Todd, Terry Lynn. 2008. “A Grammar of Dimili: Also Known as Zaza.” PhD diss., University of Michigan. <https://forum-linguistik.de/de/zazakidimili/%E2%80%A2-grammatik>.

¹² Todd (2008, 16-17) does provide evidence for six diphthongs, but analyzes all of them as VC, as in [nɔw] ‘nine.’ This use of the term ‘diphthong’ does not imply an extra V slot in the nucleus, but rather a VC structure, which is consistent with this analysis.

Exhaustive list of onset cluster types present: OG NG OL

(Ban Lua) Bisu

Language Overview

		Source(s)
Language Name	Bisu	Ethnologue and Glottolog
Code	ISO 639-3: bzi	Ethnologue
	Glottocode: bisu1244	Glottolog
Classification	Sino-Tibetan, Tibeto-Burman, Ngwi-Burmese, Ngwi, Southern, Bisoid	Ethnologue
	Sino-Tibetan, Burmo-Qiangic, Lolo-Burmese, Loloish, Hani-Jino, Bisoid-Hanic, Bisoid, Bisu-Pyen-Laomian	Glottolog
Number of Speakers	700	Ethnologue
Alternate Names¹³	Laopin, Lawa, Lua, Mbi, Mbisu, Mibisu, Misu, Pin, Bisú	Ethnologue and Glottolog
Location	Thailand (Chiang Rai Province: Doi Chomphu and Pui Kham villages)	Ethnologue
	20.85 N, 99.99 E	Glottolog
Dialect of Focus	Ban Lua Village	Nishida 1973, 56
Maximum Syllable Template	[CCVC]	Nishida 1973, 63-65
Inventory of Consonant Phonemes (30)	p t k ? p ^h t ^h k ^h b d g ts te ts ^h te ^h (f) s ɛ h m n ɲ ŋ m̥ ŋ̊ ŋ̊ ŋ̊ l̊ j ẘ j̊	Nishida 1973, 60-62
Syllable-Initial Clusters (10)	pj kj (p ^h j) k ^h j bj m̊j pl kl p ^h l k ^h l bl	Nishida 1973, 63

¹³ According to the Ethnologue, “the terms ‘Lawa’ and ‘Lua’ generally refer to Waic language varieties in Thailand: sometimes outsiders have classified unrelated languages such as Bisu in this category.”

Examples of Onset Clusters

OG initial:	[pjà:] ‘bee’	(Nishida 1973, 63)
	[kjaù.né:] ‘to wash (cloth, hair)’	(Nishida 1973, 63)
OG medial:	[sò:.p ^h jè:] ‘tooth’	(Nishida 1973, 63)
	[ʔāŋ.bjà:] ‘many’	(Nishida 1973, 63)
NG initial:	[m̄jā.t ^h án] ‘knife’	(Nishida 1973, 63)
NG medial:	[ʔa.m̄jàŋ] ‘cow’	(Nishida 1973, 71)
OL initial:	[p ^h lā:] ‘priest’	(Nishida 1973, 63)
	[blà:] ‘arrow’	(Nishida 1973, 63)
OL medial:	[ʔāŋ.kjàm] ‘slow’	(Nishida 1973, 63)
	[mà.plàw] ‘palm’	(Nishida 1973, 66)

Other Notes

While Nishida (1973) does not address the interpretation issue of OG and NG clusters, the evidence supports such an interpretation. First, if the [j] were regarded as a vowel ([i]), it would create VV clusters that are otherwise unattested in unambiguous environments.¹⁴ Second, a C_i analysis would lead to five new phonemes in the inventory, an issue which an OG analysis avoids. It does so without increasing the maximum syllable template [CCVC], because there are unambiguous complex onsets of the type OL. Nishida (1973) lists [j] as a consonant glide and uses it consistently as such in transcriptions. Xu (2001, 18), on the other hand, lists these as contrastive palatalized phonemes, but does not give arguments for that analysis. We concur with Nishida (1973) that these are glide offset clusters, for the reasons given above.

While /l/ and /l̄/ are separate phonemes, the /l/ often becomes voiceless when in a cluster (Nishida 1973, 63). As such, it is unclear whether the [l̄] transcriptions above are part of the phoneme /l/ or /l̄/. The three tones in Bisu are High, Mid, and Low (57), transcribed here as [á], [ā], and [à], respectively. Finally, Nishida notes that there are both voiced nasals as well as nasals with voiceless on-glides (62-63). The latter are transcribed as [ᵑŋ]. We have transcribed these simply as voiceless nasals.

References

- Nishida, Tatsuo. 1973. “A Preliminary Study of the Bisu Language: A Language of Northern Thailand, Recently Discovered by Us.” In *Papers in South East Asian Linguistics No. 3*, edited by S.A. Wurm, 55-82. Canberra: Department of Linguistics, The Australian National University. <http://sealang.net/sala>.
- Xu, Shixuan. 2001. *The Bisu Language*. Translated by Cecilia Brassett. Muenchen: Lincom Europa.

¹⁴ Nishida (1973, 64-65) does address the diphthongs consisting of a non-high vowel followed by a high vocoid, as in [mà.plàw] ‘palm’ (Nishida 1973, 65), although the conclusion is unclear whether they are best analyzed as VV or VC. Even if these are analyzed as VV (i.e. [au]), interpreting the off-glide [j] as the vowel [i] would still create an nucleus of VVV (i.e. [iau]), otherwise unattested in unambiguous environments.

OG medial:	[in ³³ .k ^h wa ³¹] ‘next year’	(Fried 2000, 16)
	[ja ³¹ .pjɛ ³³] ‘duck’	(Fried 2000, 16)
NG initial:	[mjɛ ³¹] ‘ripe’	(Fried 2000, 16)
	[nja ³¹] ‘black’	(Fried 2000, 16)
NG medial:	[a ³³ .wa ³³ .ji ³¹ .mjɛ ^{ʔ13}] ‘weasel’	(Fried 2000, 80)
	[a ³¹ .nja ³¹ .k ^h u ³³] ‘next year’	(Fried 2000, 79)
LG initial:	[lja ³³ .bu ^{ʔ35}] ‘arm’	(Fried 2000, 16)
	[lja ³¹] ‘young’	(Fried 2000, 16)
LG medial:	[i ³³ .lja ³¹ .ka ³¹] ‘rice seedling’	(Fried 2000, 79)
	[su ³¹ .lje ³³] ‘forest’	(Fried 2000, 79)
OL initial:	[glu ³⁵] ‘boil water’	(Fried 2000, 16)
	[tlu ³¹] ‘white’	(Fried 2000, 16)
OL medial:	[na ³⁵ .gwi ⁵³ .k ^h la ^{ʔ35}] ‘ear’	(Fried 2000, 16)
	[a ³¹ .tlu ^{ʔ35}] ‘ginger’	(Fried 2000, 80)

Other Notes

Fried (2000) addresses the interpretation issues of both high vocoids and affricates, listing the affricates with the stops based on precedent for Loloish languages (13). However, Fried concedes that “it might be argued that the affricates act suspiciously like clusters...in that they do not occur prenasalized and they never occur with another consonant in an initial cluster” (13). Fried does not go into much more detail on the subject, and later treats them as filling only one C position (43-44). These possible interpretation issues do not pose a serious issue to the MDO, since if these were instead treated as sequences of segments, the cluster OO would be added, a language type that would still be expected based on the MDO model (as found in Parker 2017, 25).

Regarding the CG onsets, Fried (2000, 31) gives three core arguments in favor of a CGV analysis, rather than C^v/C^wV or CVV interpretations. First, there are so many types that analyzing them as palatalized or labialized would greatly expand the phoneme inventory. Second, CC onsets are already present unambiguously, so CG onsets would not expand the syllable template; conversely, analyzing these as vowels would nearly double the phoneme inventory. Finally, the maximal syllable type [CCVC] is only possible when the second onset consonant is a glide. Alternatively, if glide offsets were interpreted as part of the nucleus or as palatalization or labialization, then the maximal syllable type [CCVC] would never be fully realized as such in an actual phonetic form.

Phula has four contrastive tones: “level (33), high falling (53), low falling (31), and rising (35)” (Fried 2000, 23). Fried notes that the glottal stop is not a contrastive phoneme in Phula, and analyzes syllables ending with phonetic glottal stops as having “tense voice” (28). Finally, Fried does not give evidence for the interpretation of prenasalized stops, but clearly analyzes them as single units, and these units never cluster with another consonant in a complex onset in any lists or examples that he presents.

The segment transcribed here as [ɬ] is listed in Fried (2000, 11) as a voiceless lateral. We assume, based on the IPA symbol used, that it is a lateral fricative. When in the second position of a cluster, the lateral offset matches the voicing of the initial consonant. We analyze this segment as a liquid rather than an obstruent when it occurs in a consonant cluster, based on the

precedent of the unambiguous cluster types present. Otherwise, a sequence such as [tʃ] or [kʃ] would have to be treated as OO clusters, differently than [dl] or [gl].

Finally, Fried (2000, 11) lists all of the alveolar consonants as ‘dental,’ although it’s unclear whether they all are truly dental sounds, due to the symbols Fried uses. We have left them as alveolar in the transcriptions above.

References

- Fried, Robert Wayne. 2000. “A Preliminary Phonological Sketch of Phu Kha, a Tibeto-Burman Language Spoken in Northern Vietnam.” MA Thesis, University of Texas at Arlington.
- Parker, Steve. 2017. “Reconsidering Sonority Dispersion and Liquid vs. Glide Offsets: What Do the Typological Facts Indicate?” *Winak: Revista de Estudios Interculturales* 26: 11-42.

Exhaustive list of onset cluster types present: OG OL NL

Bumthangkha

Language Overview

		Source(s)
Language Name	Bumthangkha	Ethnologue and Glottolog
Code	ISO 639-3: kjz	Ethnologue
	Glottocode: bumt1240	Glottolog
Classification	Sino-Tibetan, Tibeto-Burman, Western Tibeto-Burman, Bodish, East Bodish, Bumthang	Ethnologue
	Sino-Tibetan, Bodic, Bodish, East Bodish, Phobjib-Chali-Bumthangic, Chali-Bumthangic, Bumthangic	Glottolog
Number of Speakers	20,000	Ethnologue
Alternate Names	Bhumtam, Bumtang, Bumtangkha, Bumtanp, Bumthang, Bumthapkha, Kebumtamp	Ethnologue and Glottolog
Location	Bhutan (Bumthang, Lhuentse, Mongar, and Trongsa districts)	Ethnologue
	27.65 N, 90.80 E	Glottolog
Dialect of Focus	Tang	van Driem 2015, 20
Maximum Syllable Template	[CCVC-s] The suffix /-s/ denoting experiential past may occur word-finally as part of a complex coda.	van Driem 2015, 21-26
Inventory of Consonant Phonemes (36)	p t̪ t̪ c k p ^h t̪ ^h t̪ ^h c ^h k ^h b d̪ d̪ ʝ g ts̪ ts̪ ^h dz̪ s f h h̃ z ʒ m ŋ ɲ ŋ l̪ l̪ r̪ r̪ ^h r̪ j w	van Driem 2015, 22
Syllable-Initial Clusters (15)	pl kl bl gl ml pr kr p ^h r k ^h r br gr mr kw k ^h w gw	van Driem 2015, 24-25

Examples of Onset Clusters

OG initial:	[kwá] ‘tooth’	(van Driem 2015, 25)
	[k ^h wí] ‘dog’	(van Driem 2015, 25)
	[gwì] ‘hip’	(van Driem 2015, 25)
OG medial:	[ḍòŋ.kwá] ‘fangs of a mammal’ ¹⁵	(van Driem 2015, 54)
OL initial:	[klát.pá] ‘brains’	(van Driem 2015, 24)
	[blè] ‘four’	(van Driem 2015, 24)
	[krón] ‘village’	(van Driem 2015, 24)
OL medial:	[mà.plá.gàè] ‘don’t make noise!’	(van Driem 2015, 24)
	[blàk.blò.gà] ‘sloppy’ ¹⁶	(van Driem 2015, 24)
	[k ^h rán.k ^h rán] ‘crane (bird)’	(van Driem 2015, 24)
NL initial:	[mlàk.sá] ‘[you] are soiling [your hand] in something sticky’	(van Driem 2015, 24)
	[mràt] ‘paddy’	(van Driem 2015, 24)
NL medial:	[mlàk.mlò.gà] ‘vicious, syrupy, oozing’	(van Driem 2015, 24)

Other Notes

Tone in Bumthangkha is somewhat predictable, with initial voiced obstruents taking low tone, and initial voiceless consonants taking high tone. Those with voiced sonorants may take a low or high tone (van Driem 2015, 23). All phonetic pitches are marked in the examples above.

Although van Driem (2015, 24-25) presents these as true consonant clusters, he does not directly address the issue of interpretation of the glide offsets. This is particularly relevant since the only OG clusters present are homorganic velar clusters. However, all relevant evidence we can glean supports an OG analysis. First, their inclusion as clusters does not expand the syllable template, due to the presence of unambiguous liquid offsets. Conversely, analyzing them as labialized consonants would add three new phonemes. Second, [w] occurs independently, as in [wì] ‘2nd singular ergative’ and [ré.wà] ‘tool for removing ears of wheat from the stalks.’ Third, van Driem (2015 25), notes that /w/ in consonant clusters is often pronounced as a labial-palatal approximant [ɥ], rather than a labial-velar approximant. Due to this detail, Cw clusters are not fully homorganic, and are less likely to be labialized consonants. In light of this evidence, there is no need to interpret [w] as a nuclear [u] (or the front glide [ɥ] as [y]), as this would create sequences of vowels (diphthongs) that do not otherwise occur (cf. van Driem 2015, 26).

A few additional clarifications are needed on the transcriptions used in van Driem (2015). First, van Driem lists a voiceless palatal-glottal fricative (represented there as [hy]) (22). We consider that this is likely intended to be a palatalized [hʲ], and is therefore represented as such in our inventory above. This segment does not appear in consonant clusters. Furthermore, even if this were treated as the cluster [hʲ], it would not change the cluster typology of this language since it would then be analyzed as OG, which is independently attested in Bumthangkha. Second, there are four trills in van Driem (2015, 22). As the IPA does not have a symbol for a trilled fricative, we have adapted the closest IPA symbol for each: voiced apical trill [r̥], voiced

¹⁵ This word is morphologically complex, as [kwá] is the word for ‘tooth.’

¹⁶ This word is fully glossed in van Driem (2015, 24) as ‘sloppy, spilt’ [said of fluid or handwriting].

apical trilled fricative [ɾ], voiceless apical trilled fricative [ɾ̥], and aspirated apical trilled fricative [ɾʰ].

References

van Driem, George. 2015. "Synoptic grammar of the Bumthang language, a language of the central Bhutan highlands." *Himalayan Linguistics Archive* 6: 1-77.
https://www.researchgate.net/publication/316655660_Synoptic_grammar_of_the_Bumthang_language_a_language_of_the_central_Bhutan_highlands.

OG medial:	[t ^h oŋ.gju] ‘will drink.1’	(Hyslop and Tshering 2008, 12)
OL initial:	[bra] ‘cliff’	(Hyslop and Tshering 2008, 9)
	[klo] ‘lose’	(Hyslop and Tshering 2008, 9)
NL initial:	[mrop] ‘scratch’	(Hyslop and Tshering 2008, 9)
	[ŋla] ‘lick’	(Hyslop and Tshering 2008, 9)

Other Notes

Hyslop and Tshering (2008) is a preliminary field report, and the authors note that it is premature to put forth a full phonemic analysis (7). More data may reveal more onset clusters (9).

While Hyslop and Tshering (2008) do not address the interpretation of glides, they do present OG examples as true consonant clusters. One apparent complication is that Dakpakha does contain four diphthongs: [au], [ui], [ei], and [ai]. In the limited data in Hyslop and Tshering (2008), there are no cases of CG clusters occurring with those four diphthongs (which supports a diphthong analysis), nor do these GV sequences occur after CC (supporting a CG analysis). However, each of the unambiguous diphthongs end with a high vocoid. Analyzing these CG clusters as diphthongs would create new (additional) diphthongs ending with a non-high vocoid. Regarding palatalization, the clusters with palatal offsets are [kj], [p^hj], and [gj], none of which are homorganic. Consequently, these are unlikely to represent true palatalization. A CG cluster analysis avoids each of these complications without expanding the syllable types or phonemic inventory.

References

Hyslop, Gwendolyn, and Karma Tshering. 2008. “Preliminary Notes on Dakpa (Tawang Monpa).” *North East Indian Linguistics* 2: 1-22.
https://www.academia.edu/5085700/Preliminary_notes_on_Dakpa_Tawang_Monpa_

Exhaustive list of onset cluster types present: OG NG OL NL

Khengkha

Language Overview

		Source(s)
Language Name	Khengkha	Ethnologue and Glottolog
Code	ISO 639-3: xkf	Ethnologue
	Glottocode: khen1241	Glottolog
Classification	Sino-Tibetan, Tibeto-Burman, Western Tibeto-Burman, Bodish, East Bodish, Bumthang	Ethnologue
	Sino-Tibetan, Bodic, Bodish, East Bodish, Phobjib-Chali-Bumthangic, Chali-Bumthangic, Bumthangic	Glottolog
Number of Speakers	65,000	Ethnologue
Alternate Names	Ken, Keng, Kenkha, Khen, Khenkha, Kyengkha	Ethnologue and Glottolog
Location	Bhutan (Mongar, Sarpang, Trongsa, and Zhemgang districts)	Ethnologue
	27.09 N, 90.86 E	Glottolog
Maximum Syllable Template	[CCVC]	Chamberlain 2004, 88
Inventory of Consonant Phonemes (27)	p t k p ^h t ^h k ^h b d g ts tʃ ts ^h tʃ ^h dz dʒ s ʃ h z ʒ m n ŋ l r j w	Chamberlain 2004, 90
Syllable-Initial Clusters (23)	pl kl p ^h l bl gl ml pr tr kr p ^h r t ^h r k ^h r br dr gr mr ŋr pj p ^h j bj mj nj kw	Chamberlain 2004, 95

Examples of Onset Clusters

OG initial:	[p ^h ja] ‘mountain side’	(Chamberlain 2004, 96)
	[kwa] ‘tooth’	(Chamberlain 2004, 97)
NG initial:	[njen] ‘marriage’	(Chamberlain 2004, 98)
	[mjə] ‘arrow’	(Chamberlain 2004, 98)

OL initial:	[pra] ‘monkey’	(Chamberlain 2004, 96)
	[glap] ‘to beat’	(Chamberlain 2004, 97)
OL medial:	[ts ^h i.trim] ‘celibate’	(Chamberlain 2004, 107)
	[toŋ.t ^h ra] ‘thousand’	(Chamberlain 2004, 132)
NL initial:	[ŋraŋ] ‘be full’	(Chamberlain 2004, 97)
	[mlak] ‘sit (tailor style)’	(Chamberlain 2004, 98)

Other Notes

Syllables in Khengkha are realized with either ‘high’ (modal) phonation, or ‘low’ (breathy) phonation (Chamberlain 2004, 88). This alternation is usually predictable based on the first consonant in the syllable; only nasal-initial syllables are not predictable (89). Chamberlain (2004) marks high register nasals with an apostrophe [’]. To represent the phonetic contrast, all ‘low’ (breathy) syllables here will be marked as [a], whether or not the breathiness is phonemic. ‘High’ (modal) syllables are unmarked here.

Chamberlain (2004) addresses three main issues of interpretation regarding ambiguous segments. First, while many Tibetan languages have retroflex /t/, /t^h/, and /d/, the corresponding Khengkha words contain the clusters /tr/, /t^hr/, and /dr/ instead of these retroflexed consonants (98). Chamberlain notes that these impressionistically sound like clusters, and gives both diachronic and distributional evidence for analyzing them as clusters. Many retroflexed consonants in Tibetan languages are written orthographically as clusters, indicating that they were once truly clusters. Chamberlain notes that Khengkha cognates of these words are pronounced as clusters (99). His distributional argument regards the unambiguous presence of Cr clusters in which the initial consonant is not coronal (98). This seamlessly allows an interpretation of coronal Cr clusters as sequences, to preserve symmetry in the phonemic inventory. These Cr clusters do not further cluster with any additional consonants, again confirming their status as clusters (99) rather than as single units.

Finally, /nj/ is often analyzed as /ɲ/ in Tibetan languages. However, Chamberlain (2004, 99) argues that if [nj] were analyzed as [ɲ] in Khengkha, it would be the only nasal that cannot appear in the coda. Otherwise, Chamberlain does not directly address the interpretation of potential glide offsets. However, we note the following facts. The language has no contrastive long vowels or diphthongs, so analyzing glide offsets as part of the nucleus would needlessly expand the maximal syllable template. Furthermore, there are unambiguous onset clusters with liquid offsets. This establishes a clear precedent for complex onsets in the language. Consequently, analyzing the glide offsets as part of a palatalized or labialized single onset segment would greatly expand the phonemic inventory, with no compensating advantage.

References

Chamberlain, Bradford Lynn. 2004. “The Khengkha Orthography: Developing a Language in the Tibetan Scriptal Environment.” MA Thesis, Graduate Institute of Applied Linguistics.

Exhaustive list of onset cluster types present: OG NG LG OL NL

Orya

Language Overview

		Source(s)
Language Name	Orya	Ethnologue and Glottolog
Code	ISO 639-3: ury	Ethnologue
	Glottocode: orya1242	Glottolog
Classification	Tor-Kwerba, Orya-Tor, Orya	Ethnologue
	Tor-Orya	Glottolog
Number of Speakers	1,600	Ethnologue
Alternate Names	Oria, Uria, Warpok, Warpu	Ethnologue and Glottolog
Location	Indonesia (Papua Province)	Ethnologue
	2.76 S, 139.89 E	Glottolog
Maximum Syllable Template	[CCVCC] All sub-combinations of this are possible (as long as the V is included), except the maximum CCVCC.	Fields 1991, 33-34
Inventory of Consonant Phonemes (16)	t k ʔ b d g d͡ʒ s h z m n ŋ l j w	Fields 1991, 36
Syllable-Initial Clusters (28)¹⁸	tl kl bl dl gl sl hl ml ŋl tj kj bj dj sj hj mj nj lj tw kw bw dw gw sw hw mw nw lw	Fields 1991, 35-36

Examples of Onset Cluster

OG initial:	[bwi] ‘cooking banana’	(Fields 1991, 35)
	[ˈtwẽn.blãn] ‘to eat (meat)’	(Fields 1991, 35)
	[ˈhwæ̃.nə] ‘but’	(Fields 1991, 36)
	[sjok] ‘fire’	(Fields 1991, 36)

¹⁸ Fields (1991, 35-36) lists out the possible CG clusters, but not the possible CL clusters. The list here is taken from the many examples he provides. As such, the CL list here may not be exhaustive. In addition, Fields (1991, 55) notes that [gj] may also exist, but no examples of that cluster have been found in his data.

	['kjaŋ.gũn] ‘to tighten’	(Fields 1991, 36)
OG medial:	['hæb.djak] ‘they cursed him’	(Fields 1991, 36)
	[lek. 'tja.kə] ‘he hit him’	(Fields 1991, 45)
	['gol.bwal] ‘I gave it to him’	(Fields 1991, 45)
NG initial:	[nwe] ‘eye’	(Fields 1991, 35)
	['mjæk.mjæk.sõn] ‘to jiggle it’	(Fields 1991, 36)
NG medial:	['mjæk.mjæk.sõn] ‘to jiggle it’	(Fields 1991, 36)
	['wẽ.njə] ‘woman.FOC’	(Fields 1991, 45) ¹⁹
LG initial:	[lwa] ‘dog’	(Fields 1991, 36)
LG medial:	['wa.ljak] ‘young’	(Fields 1991, 36) ²⁰
OL initial:	[bli] ‘frog’	(Fields 1991, 35)
	[kla. 'kat.na] ‘narrow’	(Fields 1991, 35)
	[sræ̃n] ‘different’	(Fields 1991, 35)
	[hla.l] ‘axe’	(Fields 1991, 35)
OL medial:	['bas.glām] ‘ant’	(Fields 1991, 35)
	[go.l. 'træ.blān] ‘to show him’	(Fields 1991, 35)
NL initial:	[ŋli] ‘sky’	(Fields 1991, 35)
	[mlĩn] ‘loop’	(Fields 1991, 35)

Other Notes

Stress in Orya is contrastive (Fields 1991, 32), and is marked phonetically in the examples above. Nasalization of vowels is predictable before nasal consonants (39). Finally, the phoneme /l/ has two other allophones besides itself. When a CC cluster begins with an alveolar consonant followed by /l/, the latter is realized as [r] (38). When /l/ occurs either before a [-continuant] consonant or at the end of a one-syllable word, it is realized as the lateral flap [ɺ] (37).

Fields (1991, 35) addresses the interpretation issue of CG clusters in Orya. He gives two main arguments for a CG interpretation as opposed to C^j and C^w. First, the glides appear unambiguously as phonemic consonants in words such as [trú.wẽn] ‘they sleep’ and [íje] ‘relative’ (38, 52). Second, since the syllable template allows CC onsets (with liquids as the second member), allowing glides to fill the second slot in those clusters does not expand the canonical template (35). Strengthening these two arguments, we can add two more on the basis of economy. If CG clusters were analyzed as C^j or C^w, it would expand the phonemic inventory greatly. Conversely, if they were part of the nucleus, it would require expanding the syllable template to [CCVVCC]. For these reasons, it is most simple to interpret these sequences as CG clusters.

Fields (1991, 44) describes a phonological process whereby [j] is inserted between a preceding morpheme-initial alveolar consonant and a following [a]. The triggering environment

¹⁹ This word contains the morpheme /we/ ‘woman’ and the suffix /-na/ ‘FOC.’ The [j] is inserted in certain environments; see Fields (1991, 44), as well as our discussion here, for more details.

²⁰ Fields (1991) includes this word as an example of CG and does not comment further on its syllabification. While this word could theoretically be syllabified as [wal.jak], such an analysis would be dispreferred due to the Syllable Contact Law (see Seo 2011), which otherwise is generally adhered to in Orya. Regardless of the strength of this example, LG does clearly occur in words such as [lwa] ‘dog’ (Fields 1991, 36).

includes nasals, so that /we/ ‘woman’ and /-na/ ‘FOC’ combine to form [ˈwɛ.njə] ‘woman.FOC’ (45). The productivity of the insertion of [j] after alveolar segments, including nasals, again leads to the conclusion that [nj] in [ˈwɛ.njə] ‘woman.FOC’ is a consonant cluster, not a palatalized [ɲ].

Finally, Fields (1991, 55) mentions four words that do not fit the canonical syllable patterns listed above. All of these begin with CCC, either [swɾ], [dwɾ], or [kwɾ]. One of these forms is onomatopoeic, and the rarity of the other three leads Fields to conclude that they are anomalous, possibly borrowed from another language. Consequently, these clusters are not included here in the canonical syllable types.

References

- Fields, Philip C. 1991. “A Phonology of the Orya Language.” In *Papers in Papuan Linguistics 1*, edited by Tom Dutton, 29-56. Canberra: Research School of Pacific Studies, The Australian National University.
- Seo, Misun. 2011. “Syllable Contact.” In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 1245-1262. Malden, MA: Wiley-Blackwell.

Abbreviations

1	First Person	FOC	Focus
AFF	Affirmative Verbal Prefix	G	Glide
C	Consonant	H	High Tone
Cj	Palatalized Consonant	L	Liquid (or Low Tone)
C ^w	Labialized Consonant	N	Nasal
DECL	Declarative	SG	Singular
DVNA	Deverbal noun (agentive)	V	Vowel

References

- Ahland, Michael. 2009. "Aspects of Northern Mao (Bambassi-Diddesa) Phonology." *Linguistic Discovery* 7: 1-40. <https://doi.org/10.1349/PS1.1537-0852.A.332>.
- Ahland, Michael Bryan. 2012. "A grammar of Northern Mao (Màwés Aas'è)." PhD diss., University of Oregon. <https://cla.csulb.edu/departments/linguistics/wp-content/uploads/2016/01/Ahland-2012.pdf>.
- Andvik, Erik. 2010. *A Grammar of Tshangla*. Boston: Brill.
- Chamberlain, Bradford Lynn. 2004. "The Khengkha Orthography: Developing a Language in the Tibetan Scriptal Environment." MA Thesis, Graduate Institute of Applied Linguistics.
- Eberhard, David M., Gary F. Simons, and Charles D. Fennig, eds. 2021. *Ethnologue: Languages of the World*. 24th ed. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- Fields, Philip C. 1991. "A Phonology of the Orya Language." In *Papers in Papuan Linguistics 1*, edited by Tom Dutton, 29-56. Canberra: Research School of Pacific Studies, The Australian National University.
- Fried, Robert Wayne. 2000. "A Preliminary Phonological Sketch of Phu Kha, a Tibeto-Burman Language Spoken in Northern Vietnam." MA Thesis, University of Texas at Arlington.
- Gilley, Leoma. 1992. *An Autosegmental Approach to Shilluk Phonology*. Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics 103. Dallas: Summer Institute of Linguistics.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath, and Sebastian Bank. 2021. *Glottolog* 4.4. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.4761960> (Available online at <http://glottolog.org>)
- Huisman, Ronald, Roberta Huisman, and Joy Lloyd. 1987. "Angaatihā Syllable Patterns." In *Angan Languages are Different*, edited by Phyllis M. Healey, 51-62. Huntington Beach, CA: Summer Institute of Linguistics.
- Hyslop, Gwendolyn, and Karma Tshering. 2008. "Preliminary Notes on Dakpa (Tawang Monpa)." *North East Indian Linguistics* 2: 1-22. https://www.academia.edu/5085700/Preliminary_notes_on_Dakpa_Tawang_Monpa.
- Nishida, Tatsuo. 1973. "A Preliminary Study of the Bisu Language: A Language of Northern Thailand, Recently Discovered by Us." In *Papers in South East Asian Linguistics No. 3*, edited by S.A. Wurm, 55-82. Canberra: Department of Linguistics, The Australian National University. <http://sealang.net/sala>.
- Oguri, Hiroko, and Carol Erickson. 1975. "A Tentative Phonology of Isirawa." *Irian: Bulletin of Irian Jaya Development* 5(3): 38-66. <http://www.papuaweb.org/dlib/irian/4-1.PDF>.
- Parker, Steve. 2012. "Sonority Distance vs. Sonority Dispersion: A Typological Survey." In *The Sonority Controversy*, edited by Steve Parker, 101-165. Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110261523>.

- Parker, Steve. 2017. "Reconsidering Sonority Dispersion and Liquid vs. Glide Offsets: What Do the Typological Facts Indicate?" *Winak: Revista de Estudios Interculturales* 26: 11-42.
- Remijsen, Bert, Otto G. Ayoker, and Timothy Mills. 2011. "Shilluk." *Journal of the International Phonetic Association* 41 (1): 111-125.
<https://doi.org/10.1017/S0025100310000289>.
- Seo, Misun. 2011. "Syllable Contact." In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 1245-1262. Malden, MA: Wiley-Blackwell.
- Todd, Terry Lynn. 2008. "A Grammar of Dimili: Also Known as Zaza." PhD diss., University of Michigan. <https://forum-linguistik.de/de/zazakidimili/%E2%80%A2-grammatik>.
- Troyer, Madeline. 1959. "Gaddang Phonology." *The Philippine Journal of Science* 88 (1): 95-102. https://philjournalsci.dost.gov.ph/images/pdf_upload/pjs1959/PJS_Vol_88_No1_Mar_1959.pdf.
- van Driem, George. 2015. "Synoptic grammar of the Bumthang language, a language of the central Bhutan highlands." *Himalayan Linguistics Archive* 6: 1-77.
https://www.researchgate.net/publication/316655660_Synoptic_grammar_of_the_Bumthang_language_a_language_of_the_central_Bhutan_highlands.
- Xu, Shixuan. 2001. *The Bisu Language*. Translated by Cecilia Brassett. Muenchen: Lincom Europa.