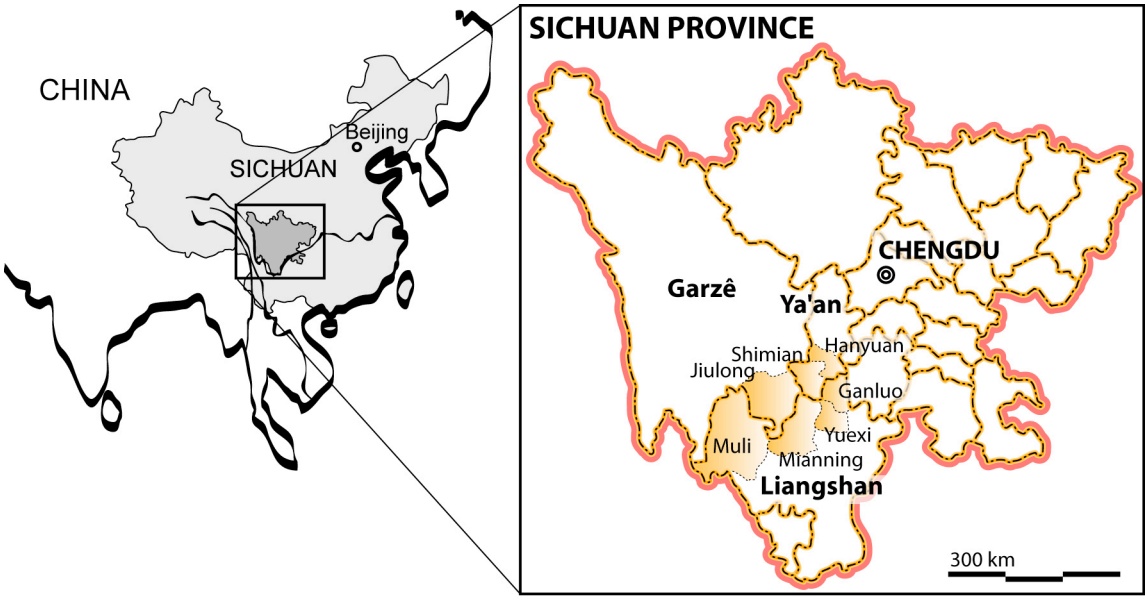
Ersu

Katia Chirkova, Dehe Wang, Yiya Chen, Angelique Amelot, Tanja Kocjančič Antolík

Ersu (Illustrations of the IPA)

Katia Chirkova, Wang Dehe, Yiya Chen, Angélique Amelot, and Tanja Kocjančič Antolík

The Ersu language (/ə̀-ʂʏ̀ xò/, 尔苏语 *ěrsūyǔ*, ISO-639 code *ers*) is spoken by approximately 16,800 people who reside in five counties in Sichuan Province (四川省) in the People's Republic of China: (i) Ganluo (甘洛县), and (ii) Yuexi (越西县) counties of Liangshan Yi Autonomous Prefecture (凉山彝族自治州), (iii) Shimian (石棉县) and (iv) Hanyuan (汉源县) counties of Ya'an Municipality (雅安市), and (iv) Jiulong (九龙县, Written Tibetan, hereafter WT *brgyad zur*) county of Ganzi (甘孜, WT *dkar mdzes*) Tibetan Autonomous Prefecture.[1] Ersu has two closely related sister languages: Lizu (/li^{55}-zu^{55}-hũ^{55}/ or /ly^{55}-zu^{55}-hũ^{55}/, 里汝语 *lǐrǔyǔ* or 栗苏语 *lìsūyǔ*) and Duoxu (/do^{55}-ʂu^{55}-na^{42}/, 多续语 *duōxùyǔ* or 多须语 *duōxūyǔ*).[2] Lizu is spoken in the neighbouring counties of (i) Muli Tibetan Autonomous County (木里藏族自治县, WT *smi li rang skyong rdzong*), (ii) Mianning (冕宁县), and (iii) Jiulong; whereas Duoxu is spoken in the county of Mianning, all in Sichuan province (see Map 1).



Map 1. Distribution of the Ersu, Lizu, and Duoxu languages

In present classifications of Tibeto-Burman languages spoken in Southwest China, Ersu, Lizu, and Duoxu are viewed as three dialects of one Ersu language. The Ersu language is, in turn, classified as a member of the Qiangic subgroup of the Tibeto-Burman language family (for more details, see Bradley 1997:36-37, Sun 2001, Chirkova 2012). In this conception, Lizu is the western dialect of the Ersu language, Duoxu is its central dialect, and Ersu proper is the eastern dialect of the Ersu language. In contrast to this received view and in accordance with the fact that differences between Lizu, Duoxu, and Ersu surpass the limit of mutual intelligibility (Sun 1982, Chirkova 2014), we consider Lizu, Duoxu, and Ersu as separate languages, and not as dialects of one Ersu language (Yu 2012: 1). The phonological, lexical, and morphosyntactic differences between Ersu, Lizu, and Duoxu are likely to be in part due to the competing influences of the languages with which they are in contact. More specifically, Ersu has been historically influenced by Pumi (普米), Nuosu (Northern Ngwi or Yi 彝), and Mandarin Chinese (the local variety of Southwest Mandarin, hereafter SW Mandarin) (Wu Da 2010:3). By contrast, Lizu has been influenced by Tibetan, Pumi, and Namuzi (纳木兹) languages (Chirkova & Chen 2013). Finally, Duoxu has been essentially influenced by SW Mandarin as well as by Nuosu (Chirkova 2014).

The main focus of this illustration is on a synchronic analysis of Ersu proper. (For more comparatively- and diachronically-oriented studies of Ersu, Lizu, and Duoxu, the interested reader is referred to Yu, 2012; Chirkova & Handel, 2013a; Chirkova, 2014.)

Ersu is relatively little researched, but the group and its language and culture have been receiving increasing attention in recent years (e.g. Wang 2010, Wu Da 2010, Schmidt 2011, Zhang 2013, 2014). Early linguistic accounts include Sun (1982, 1983) and Liu (2007 [1983]), which focus on the Ersu as spoken in Ganluo county.

Ersu is an endangered language. It is essentially used as the primary language of oral communication in family and community events. Older Ersu speakers (typically above their sixties) are mostly trilingual (Ersu, SW Mandarin, Nuosu). Over the last three decades, most Ersu speakers have been bilingual using SW Mandarin in daily life. The current trend for the school-going generation is to become practically monolingual in Mandarin. Ersu has its own pictographic writing system, known as *shaba* 沙巴 (Ersu /ʂ̩àpá/ 'ritual priest') writing, which is chiefly used by Bon priests (e.g. Sun 2009, Wang 2011).

The present illustration provides a preliminary description of Ersu on the basis of data from three speakers: two male speakers in their early sixties; and one female speaker, in her early forties; all born and raised in Ganluo county (Zela Township 则拉， Liangshan group 凉山组, Mofanggou village 磨坊沟村 /d̪z̪ɛ́ɬá lò, z̪ákáɽʐ̩́gv̩́ fv̩́, nóNd̪z̪v̩́ pá/). Given the phonetic complexity of the consonant and vowel sounds of Ersu (including a number of typologically uncommon trilled retroflex sounds and phonemic fricative vowels), further research, based on more speakers is required for a comprehensive analysis of this language. In the present illustration, the basic phonetic characteristics of Ersu are described through acoustic, palatographic, aerodynamic, electroglottographic, and video data. We have chosen to illustrate the discussion with audio files (the word list and the text provided with the paper)

as read by the second author, a male native speaker of Ersu. This is because we were in the fortunate position of recording him in a phonetics laboratory, yielding high quality audio recordings. Conversely, palatographic images in the text are from the female speaker. This is because, as the youngest speaker among our language consultants, her dental and palatal condition yielded the clearest images.

Consonants

The Ersu consonant inventory consists of 38 phonemes, listed in Table 1. There is a general three-way manner distinction in stops and affricates: voiceless unaspirated, voiceless aspirated, and voiced. Ersu has an extensive system of coronal consonant contrasts in affricates at the dental, alveolar, alveolopalatal, and retroflex places of articulation. Alveolopalatals are marginal. They mainly occur in (recent) loanwords from Mandarin Chinese, such as /kʰà-tɕà-tɕá/ 'to pick up with chopsticks' (Chinese *jiā* 夹, SW Mandarin /tɕia^{44}/), /kòtɕó/ 'legging, puttee' (Chinese *guǒjiǎo* 裹脚, SW Mandarin /ko^{53}tɕio^{21}/), /ɕá/ 'incense' (Chinese *xiāng* 香, SW Mandarin /ɕiaŋ^{44}/).[3] In the native vocabulary, alveolopalatals have a restricted distribution, co-occurring only with the vowels /i, o, a/. Examples include /tɕí-tɕí/ 'to squeeze, to pick up with chopsticks; to cut with scissors' (possibly, a loanword from Chinese *jiǎn* 剪 'to cut with scissors', SW Mandarin /tɕian^{53}/), /tɕó/ 'to wrap, to bind', /ódʑá/ 'pear' (see below on vowels).

Table 1: Ersu consonant phonemes

	Bilabial	**Labio-**	**Dental**	**Alveolar**	**Alveolo-**	**Retroflex**	**Velar**

		dental			palatal		
Plosive	p p^h b		t t^h d				k k^h g
Affricate			t̪s̪ t̪s̪^{h} d̪z̪	ʈʂ ʈʂ^h ɖʐ	tɕ tɕ^h dʑ	ʈɕ ʈɕ^h ɖʑ	
Nasal	m		n				ŋ
Trill						ɾ	
Fricative		f v	s̪ z̪	ʂ ʐ	ɕ ʑ		x
Approximant	w				j		
Lateral				l			
Lateral fricative				ɬ			

p	pɛ́	‘to give as a present’	ʂ	ʂɛ́	‘louse’
p^h	p^{h}ɛ́	‘male’	ʐ	ʐɛ́	‘to flow’
b	bɛ́	‘insect’	l	lá	‘fertilizer, manure’
m	mɛ́	‘ink’	ɬ	ɬá	‘month’
w	wá	‘Chinese yam’	tɕ	tɕó	‘to wrap, to bind’
f	fí	‘mildew’	tɕ^h	tɕ^{h}ó	‘pepper’
v	ví	‘tinder’	dʑ	dʑó	‘to push’
t	tɛ́	‘one’	ɕ	ɕó	‘to sweep’
t^h	t^{h}ɛ́	‘he, she, it’	ʑ	ʑó	‘to soften (skin)’

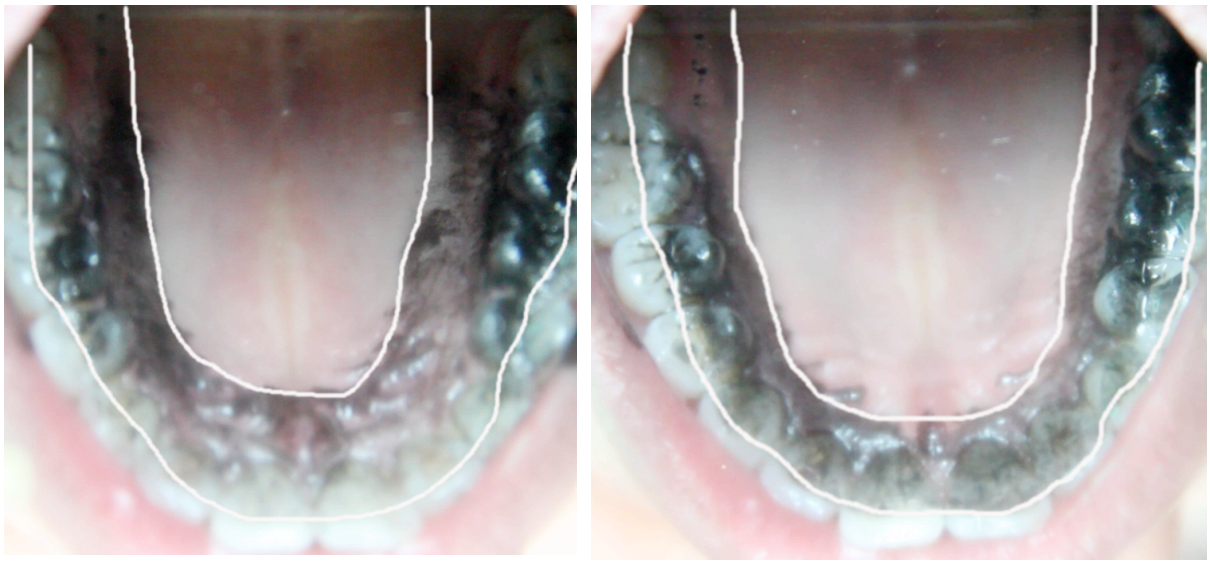
d	dɛ́	‘to weave’	j	jó	‘sheep’
t̪s̪	t̪s̪ɛ́	‘cloud’	ʈɾ	ʈɾɛ́	‘steelyard’
t̪s̪^{h}	t̪s̪^{h}ɛ́	‘to wash’	ʈɾ^{h}	ʈɾ^{h}ɛ́	‘fence’
d̪z̪	d̪z̪ɛ́	‘rooster (zodiac sign)’	ɖɾ	ɖɾɛ́	‘pair’
n	nɛ́	‘you, thou’	ɽ	ɽɛ́	‘earth, soil’
s̪	s̪ɛ́	‘air’	k	ká	‘to be stupid’
z̪	z̪ɛ̀	‘wife’	k^{h}	k^{h}á	‘barley’
ṭs̪	ṭs̪ó	‘oats’	g	gá	‘to sing’
ṭs̪^{h}	ṭs̪^{h}ó	‘to rot (of wood)’	ŋ	ŋá	‘to be hungry’
ḍz̪	ḍz̪ó	‘water’	x	xá	‘to have’

We studied the four-way contrast in coronals using palatographic analysis techniques (as described in Marchal 1988, Ladefoged 2003: 36-42, and Anderson 2008). The list of words used in the palatographic study consisted of monosyllabic words in common use in the language, each of which included only one coronal consonant. The palatography procedures were to paint the tip, blade, and front of the tongue with a solution of one part olive oil and one part of finely ground activated charcoal. After each word was pronounced, a mirror was placed in the speaker’s mouth, resting against the lower teeth and the reflection of the upper palate and teeth was recorded using a video camera.

Stops and affricates

Ersu dental stops and affricates are both produced in the dental region. Dental stops involve contact on both the teeth and most of the alveolar ridge, making them (laminal) denti-alveolar. By contrast, dental affricates involve a smaller contact area, which includes the upper front teeth and the front part of the alveolar ridge. This is illustrated in Figure 1 with the words /dɛ́/ ‘to weave’ and /d̪z̪ɛ́/ ‘rooster (zodiac sign)’.

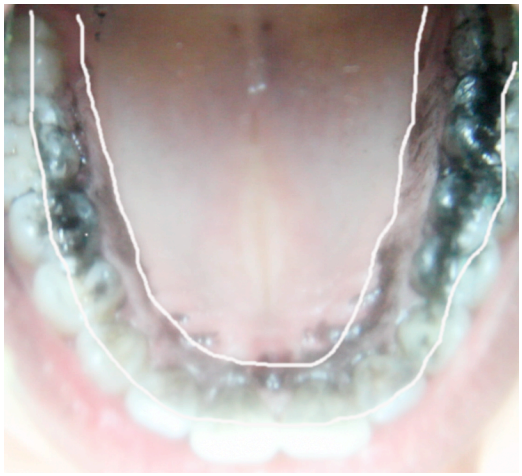
Figure 1: Palatograms of the words /dɛ́/ ‘to weave’ and /d̪z̪ɛ́/ ‘rooster (zodiac sign)’



/dɛ́/ ‘to weave’ /d̪z̪ɛ́/ ‘rooster (zodiac sign)’

Alveolar affricates are produced with the tongue touching the middle of the alveolar ridge (laminal flat alveolar). The contrast between dental and alveolar affricates is illustrated in Figure 2 with the minimal pair /d̪z̪ó/ ‘to lay bricks’ and /d̪ẓó/ ‘water’.

Figure 2: Palatograms of the words /d̪z̪ó/ ‘to lay bricks’ and /d̪ẓó/ ‘water’



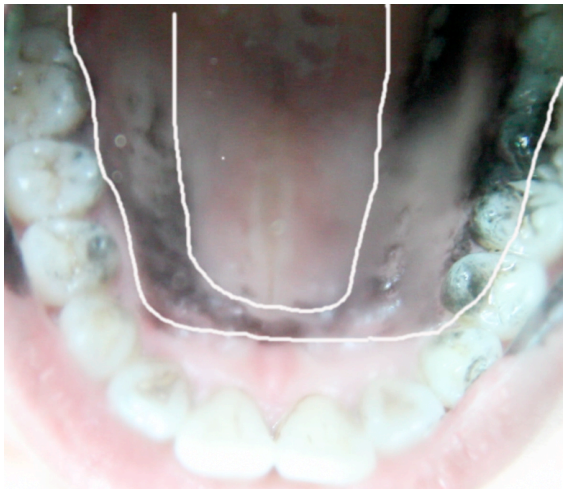
/d̪z̪ó/ ‘to lay bricks’

/ḍẓó/ ‘water’

Alveolopalatal affricates are produced with the blade of the tongue behind the alveolar ridge and the body of the tongue raised toward the palate, thus involving simultaneous alveolar and palatal articulation (cf. “laminal palatalized post-alveolar”, as described in Ladefoged & Maddieson 1996: 153-154). The contrast between alveolar and alveolopalatal affricates is illustrated in Figure 3 with the words /ḍẓó/ ‘water’ and /dʑó/ ‘to push’.

Figure 3: Palatograms of the alveolar and alveolopalatal affricates in the words /ḍẓó/ ‘water’ and /dʑó/ ‘to push’

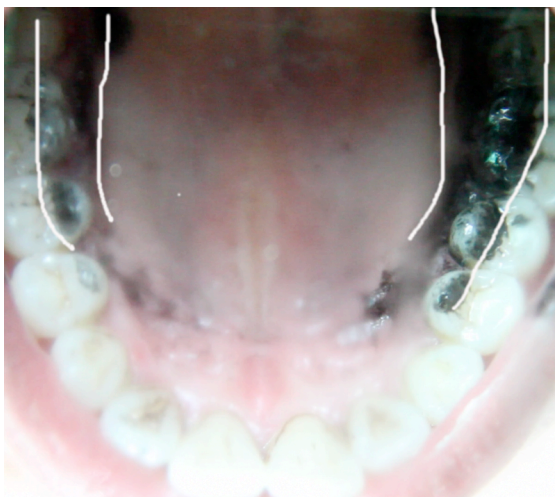
/ḍẓó/ ‘water’

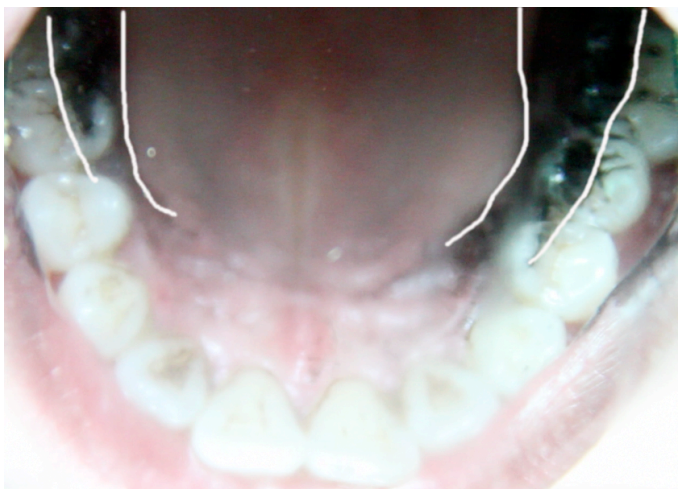


/dʑó/ ‘to push’

Ersu retroflex affricates are produced with the point of contact on the roof of the mouth, that is, in the hard palate. The contact is made with the underside of the tongue (subapical). The articulation of retroflex affricates involves lateral bracing of the tongue against the teeth, so that the tongue tip is free to move to and from the hard palate. Figure 4 contains palatograms of the words /ʈɾò/ ‘gallbladder’ and /ɖɾò/ ‘pot, pan’. As we did not paint the underside of the tongue with a mixture of olive oil and activated charcoal powder, the area of contact is not visible on the image.

Figure 4: Palatograms of the words /ʈɾò/ ‘gallbladder’ and /ɖɾò/ ‘pot, pan’

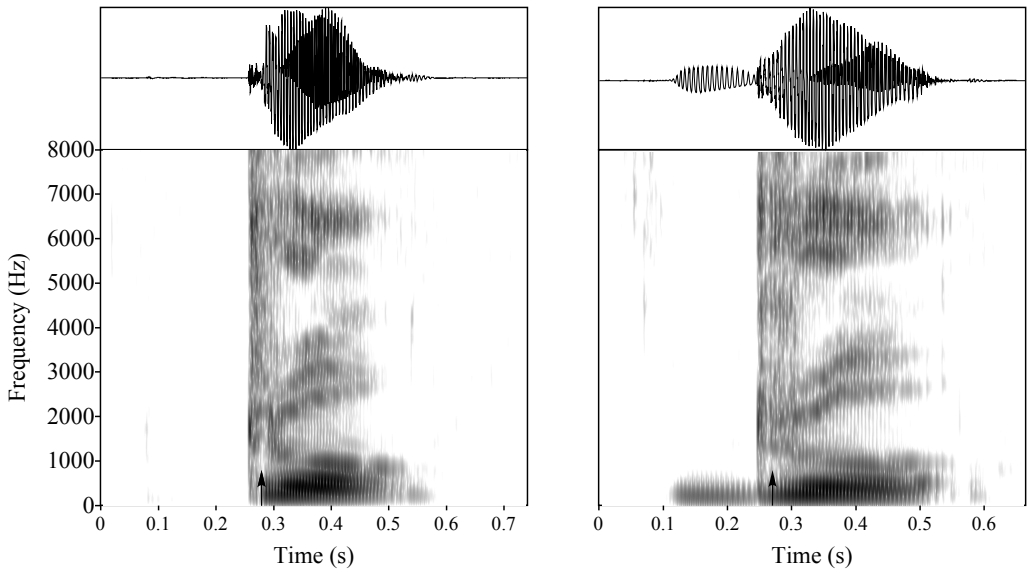




/ʈɽò/ ‘gallbladder’ /ɖɽò/ ‘pot, pan’

Ersu retroflex affricates have a trill release. They involve an aerodynamically induced movement of the tip of the tongue, causing intermittent contact between the tip of the tongue and the roof of the mouth. Ersu retroflex affricates are typically single-contact trills. This is illustrated in Figure 5 with the words /ʈɽò/ ‘gallbladder’ and /ɖɽò/ ‘pot, pan’. A contact, that is, a moment of closure of the oral cavity, is reflected on the spectrograms by a period of white space.

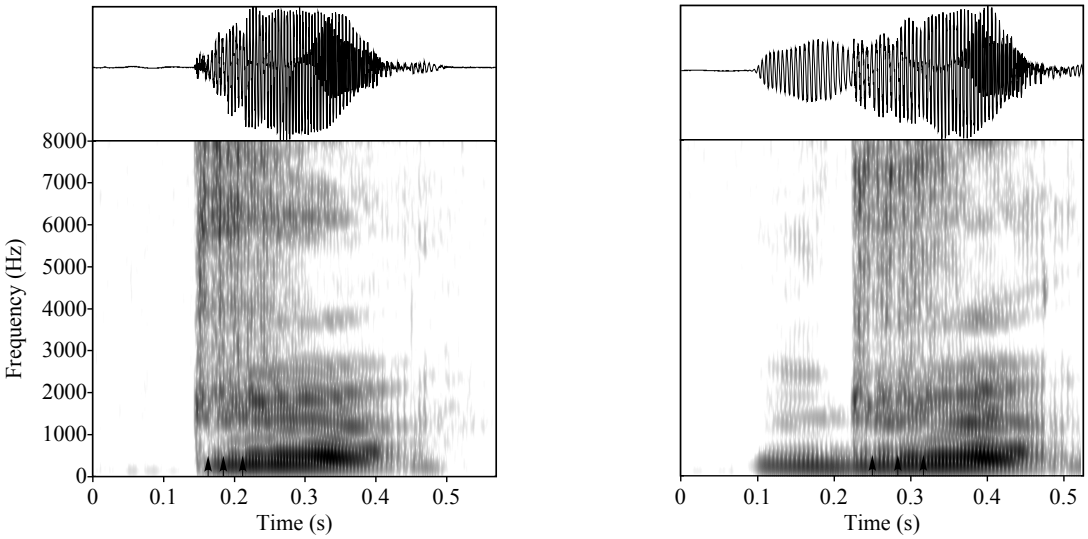
Figure 5: Waveforms and spectrograms of /ʈɽò/ ‘gallbladder’ and /ɖɽò/ ‘pot, pan’ (arrows indicate contacts between the tip of the tongue and the roof of the mouth)



/ʈɽò/ ‘gallbladder’ /ɖɽò/ ‘pot, pan’

When retroflex affricates are followed by fricative vowels, the number of contacts between the tip of the tongue and the roof of the mouth may be increased to three. This is illustrated in Figure 6 with the words /ʈɽ̩́/ ‘sweat’ and /Nɖɽ̩́/ ‘tile’.

Figure 6: Waveforms and spectrograms of /ʈɽ̩́/ ‘sweat’ and /Nɖɽ̩́/ ‘tile’ (arrows indicate contacts between the tip of the tongue and the roof of the mouth)



/ʈɽ̩́/ ‘sweat’ /Nɖɽ̩́/ ‘tile’

Fricatives

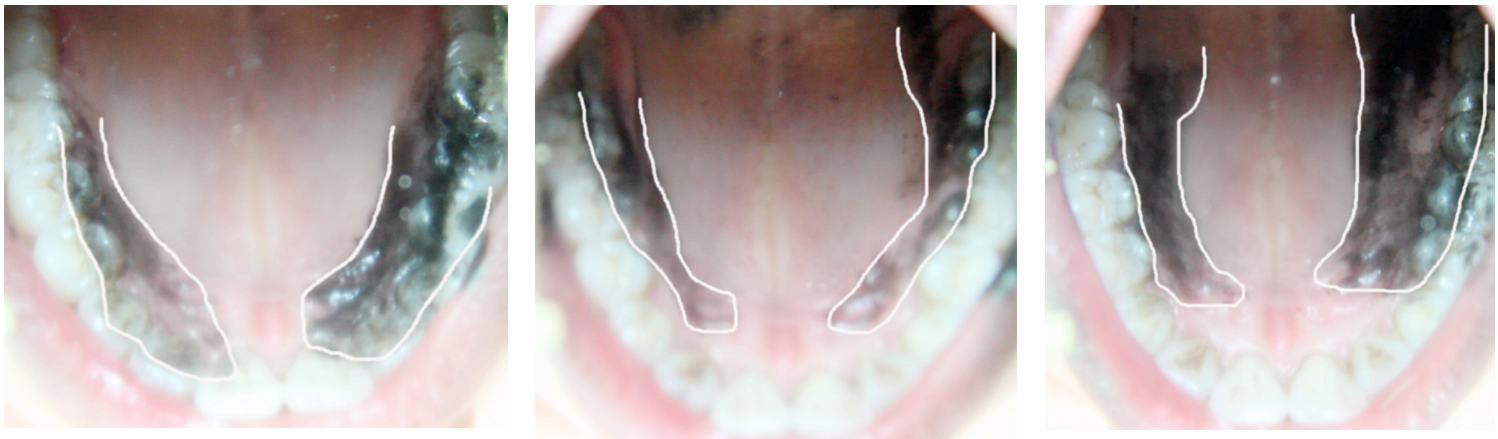
Ersu contrasts fricatives at five places of articulation: (i) labiodental, (ii) dental, (iii) alveolar, (iv) alveolopalatal (palatalized laminal post-alveolar), and (v) velar. All but velar show a two-way contrast between voiceless and voiced: /f, v/ (e.g. /fí/ ‘mildew’, /ví/ ‘tinder’), /s̪, z̪/ (e.g. /s̪ɛ́/ ‘air’, /z̪ɛ̀/ ‘wife’), /s̩, z̩/ (e.g. /s̩ò/ ‘blood’, /z̩ó/ ‘to scold; to be in debt’), /ɕ, ʑ/ (e.g. /ɕó/ ‘to sweep’, /ʑó/ ‘to soften (skin)’), /x/ (e.g. /xì/ ‘bamboo’).

The voiced labiodental fricative /v/ is contrastive with the voiced labial-velar approximant /w/ before /a/ and /o/, as in /vá/ ‘net’, /wá/ ‘to be full, satisfied’. /f/ has a restricted distribution, occurring only before /i/ (as above) and /ɽ̩/ (as in /fɽ̩̀/ ‘garlic’). In the

latter environment, /f/ can be alternatively regarded as the allophone of /x/, as is also the case in SW Mandarin, with which Ersu is in close contact. Consider, for example, the Ersu word for 'kettle': /fỳ-fý/, which is a loanword from SW Mandarin /fu^{21}-fu^{44}/, corresponding to Standard Mandarin *hú* 壺 [xu^{35}].

Ersu exploits contrasts between grooved fricatives at the dental place of articulation and flat fricatives at the alveolar place of articulation. Figure 7 illustrates the three-way contrast between the dental, alveolar, and alveolopalatal places of articulation in fricatives with the words /ʐ̪è/ 'wife', /ʐó/ 'to scold; to be in debt', and /ʑó/ 'to soften (skin)'.

Figure 7: Palatograms of the words /ʐ̪è/ 'wife', /ʐó/ 'to scold; to be in debt', and /ʑó/ 'to soften (skin)'



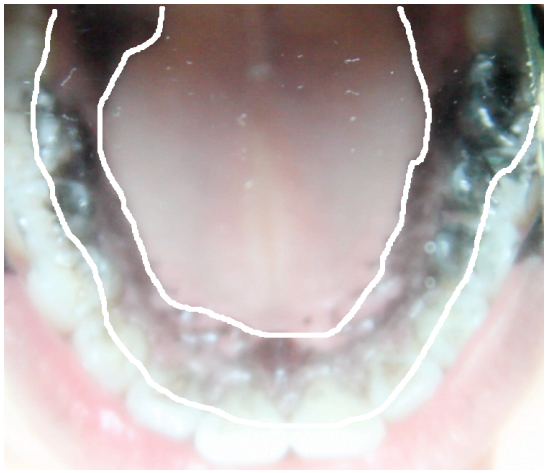
/ʐ̪è/ 'wife' /ʐó/ 'to scold; to be in debt' /ʑó/ 'to soften (skin)'

Before /o/, /x/ has an allophone [ç], which we analyse phonemically as a sequence of /x/ and the palatal approximant /j/, e.g. /xjó/ [çó] 'to cry out' (see below on clusters with approximants).

Nasals

Ersu has nasals at three places of articulation: (i) bilabial (/m/), (ii) dental (denti-alveolar) (/n/) (see Figure 8), and (iii) velar (/ŋ/).

Figure 8: Palatogram of the word /nɛ́/ ‘you, thou’

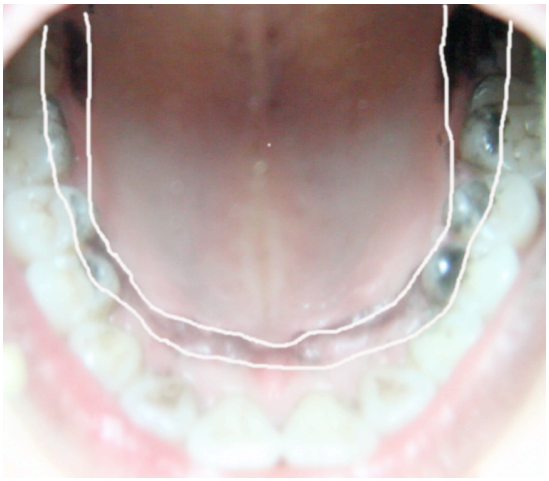


Ersu has one syllabic nasal, /ŋ̩/, as in /ʂʐ̩-ŋ̩/ ‘seven’, /lwáŋ̩ká/ ‘ridge of a building’, /ə˧ŋ̩ʤʒɛ́/ ‘satin, silk’ (possibly from Chinese *língzi* 绫子 ‘damask silk’, SW Mandarin /nin^{53}tsɿ^{53}/).

Liquids

Ersu laterals are made with an occlusion in the alveolar region extending back to the back molars. This is illustrated in Figure 9 with the minimal pair /lá/ ‘fertilizer, manure’ and /ɬá/ ‘month’.

Figure 9: Palatograms of the words /lá/ ‘fertilizer, manure’ and /ɬá/ ‘month’



/lá/ ‘fertilizer, manure’

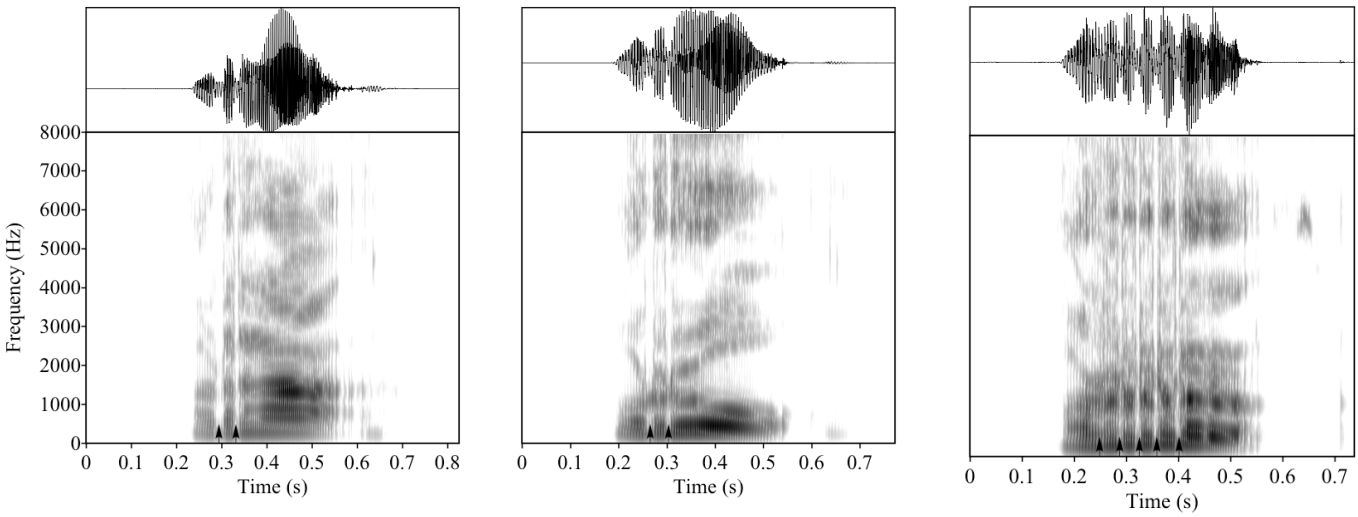
/ɬá/ ‘month’

Similarly to retroflex affricates, Ersu retroflex trill (/ʈ/) is produced with the point of contact on the roof of the mouth in the hard palate and with the tongue body braced against the sides of the teeth to allow for an aerodynamically induced movement of the tongue tip (cf. McGowan 1992:2903, Spajić et al. 1996:3; see Figure 10). Similarly to the retroflex trill in Toda (Spajić et al. 1996: 13), the first contact of the tongue tip is made back of the alveolar ridge, whereas subsequent contacts are made slightly further forward near or at the alveolar ridge, so that the trill is realized as [ʈr].

Figure 10: Palatogram of the word /ʈɛ́/ ‘earth, soil’

Ersu trill typically has two contacts. However, similarly to trilled retroflex affricates, a trill followed by a fricative vowel may have a larger number of contacts (four to five). This is illustrated in Figure 11.

Figure 11: Waveforms and spectrograms of the words /ʈá/ ‘chicken’, /ʈó/ ‘bone; horn’, and /ʈʐ́/ ‘to laugh’ (arrows indicate contacts between the tip of the tongue and the roof of the mouth)



/ʈá/ ‘chicken’ /ʈó/ ‘bone; horn’ /ʈʐ́/ ‘to laugh’

Clusters

Ersu has a rich inventory of clusters, including (i) clusters with approximants, (ii) prenasalized clusters, and (iii) clusters with a schwa-like segment.

Clusters with approximants

The approximants /w/ and /j/ may occur after a broad range of initials and they may be realized as secondary labialization or palatalization of the first position consonant. Of the two approximants, /w/ has the broadest distribution, occurring after bilabial, dental, and velar stops, dental affricates, and laterals. However, most of these clusters with /w/ can only be followed by /a/ (as in /kwá/ ‘to take off’, compare to /ká/ ‘to be stupid’). In addition, clusters with velar stops and /w/ can also be followed by /aʷ/ (as in /nà-kwáʷ/ ‘to put inside’, compare to /d̪è-z̪ɻ́-kàʷ/ ‘to bear a grudge’).

/j/ has a more restricted distribution, occurring only after (bilabial and dental) nasals and laterals, and it can only be followed by the vowels /o/ and /a/. (/j/ may also occur after /x/, which combination can only be followed by /o/. The cluster /xjo/ is realized as [ç], as in /xjó/ [çó] ‘to cry out’, see above). Consider the following (near) minimal pairs: /má/ ‘mother, female’ vs. /mjábó/ ‘tear’; /nà-ná/ ‘to occupy’ vs. /njá/ ‘child’; /lá/ ‘fertilizer, manure’ vs. /ljá/ ‘to paint’; and /Nt̪s̪ʰòɬó/ ‘flea’ vs. /mɛ́-ɬjó/ ‘lightning’.

It is important to note that the realization of /w/ and /j/ in clusters ranges between a separate segment with a clear segmental boundary (mostly before /a/) and a segment with a

secondary articulation, that is, a segment with a lesser degree of stricture that accompanies a primary articulation of a higher degree (mostly before /o/) (cf. Ladefoged & Maddieson 1996: 354). Examples include /njá/ [njá] ‘child’, /ljá/ [ljá] ‘to paint’, /njó/ or /nʲó/ [ɲó] ‘day’, /ljò/ or /lʲò/ [ʎò] ‘arm spread (measure of length)’. We note that the addition of the lip rounding gesture (in the case of /w/) and raising of the body of the tongue (in the case of /j/) have a strong acoustic effect on both the preceding consonant and the following vowel. Overall, compared to clusters with /w/ and /j/ in the closely related Lizu and Duoxu languages, Ersu medials /w/ and /j/ exhibit stronger assimilatory influence on neighboring segments (see Chirkova & Handel 2013b for a detailed discussion).

In light of the above, Ersu alveolopalatal affricates and fricatives, which only occur in native vocabulary before the vowels /i, o, a/, can be alternatively analysed as palatalized allophones of dentals, alveolars and/or velars, followed by the high front vowel or the palatal approximant /j/.

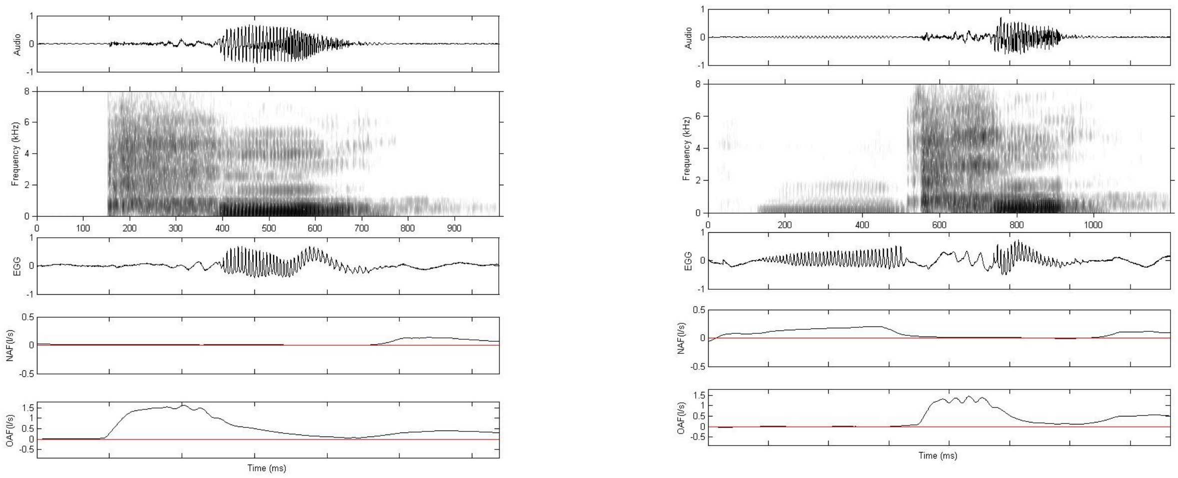
Prenasalized clusters

Prenasalization in Ersu is contrastive and occurs not only before voiced stops and affricates, but also before voiceless aspirated ones. Prenasalized stops and affricates are found in all places of articulation. The place of articulation is always homorganic with that of the obstruent in the cluster. So, we use the archiphoneme “N” to refer to the homorganic nasal in prenasalized clusters. Compare the contrast between plain onsets in the following minimal pairs: /bò/ ‘string’ vs. /Nbò/ ‘horse’, /pʰó/ ‘to escape’ vs. /Npʰó/ ‘to steal’, /d̪z̪à/ ‘fodder’ vs.

/Nd̪z̪à/ ‘Chinese people’, /t̪s̪^{h}á/ ‘hot’ vs. /Nt̪s̪^{h}á/ ‘mark, sign’, /ɖɽò/ ‘pot, pan’ vs. /Nɖɽò/ ‘dirt, filth’, /(mɛ́lí) tɛ́ ʈɽ^{h}ó/ ‘one piece (of land)’ vs. /tɛ́ Nʈɽ^{h}ó/ ‘one handful (e.g. of rice)’.

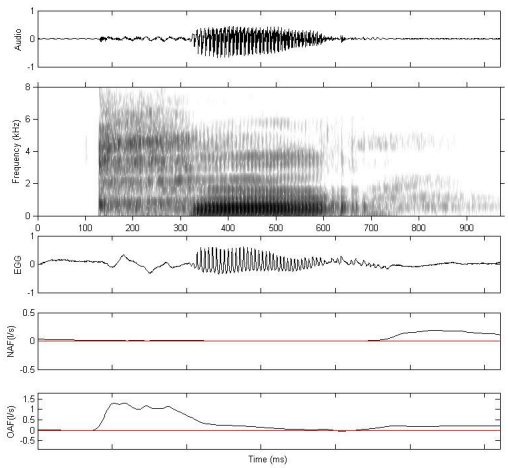
In prenasalized voiceless aspirated stops and affricates, we observe regular glottal pulsing during the nasal section but not during the voiceless stop, as detailed in Figure 12. This figure shows that the nasal section is produced with a complete closure within the oral tract (no oral airflow) and with air moving through the nasal cavity (nasal airflow for the entire duration of the segment). It is therefore a characteristic nasal.

Figure 12: Audio waveforms, spectrograms, electroglottograph waveforms (EGG), nasal airflows (NAF), and oral airflows (OAF) for the pairs /p^{h}ó/ ‘to escape’ vs. /Np^{h}ó/ ‘to steal’ and /t^{h}wá/ ‘mule’ vs. /Nt^{h}wá/ ‘to be sharp’

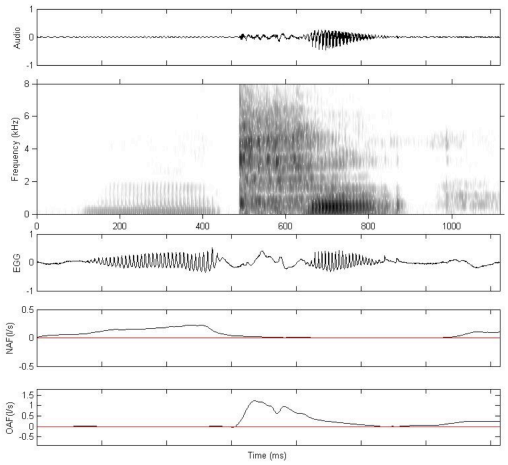


/p^{h}ó/ ‘to escape’

/Np^{h}ó/ ‘to steal’



/t^{h}wá/ ‘mule’

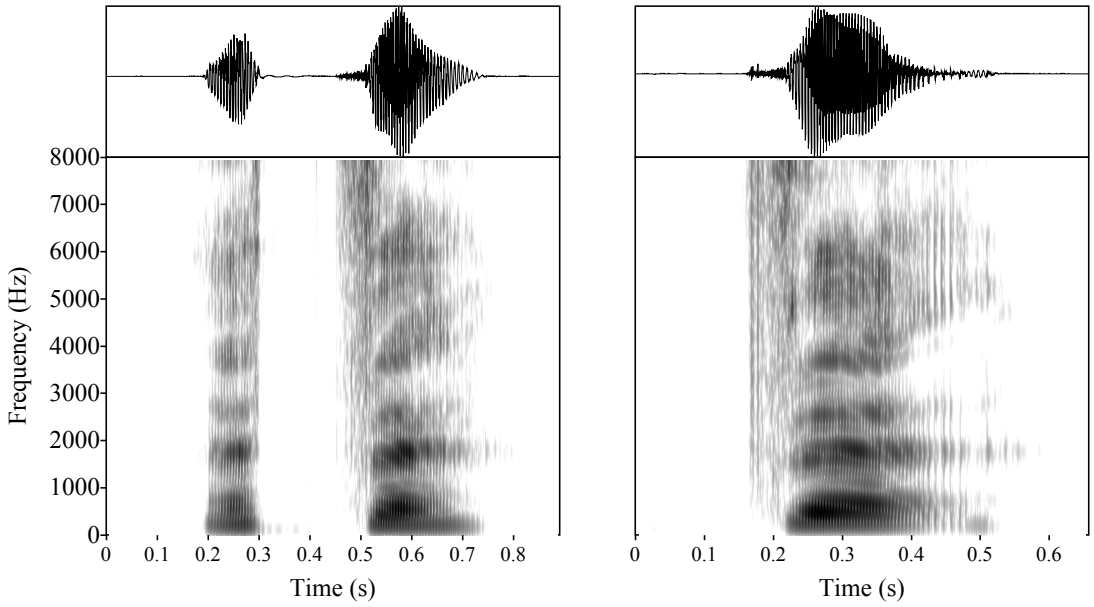


/Nt^{h}wá/ ‘to be sharp’

Clusters with a schwa-like segment

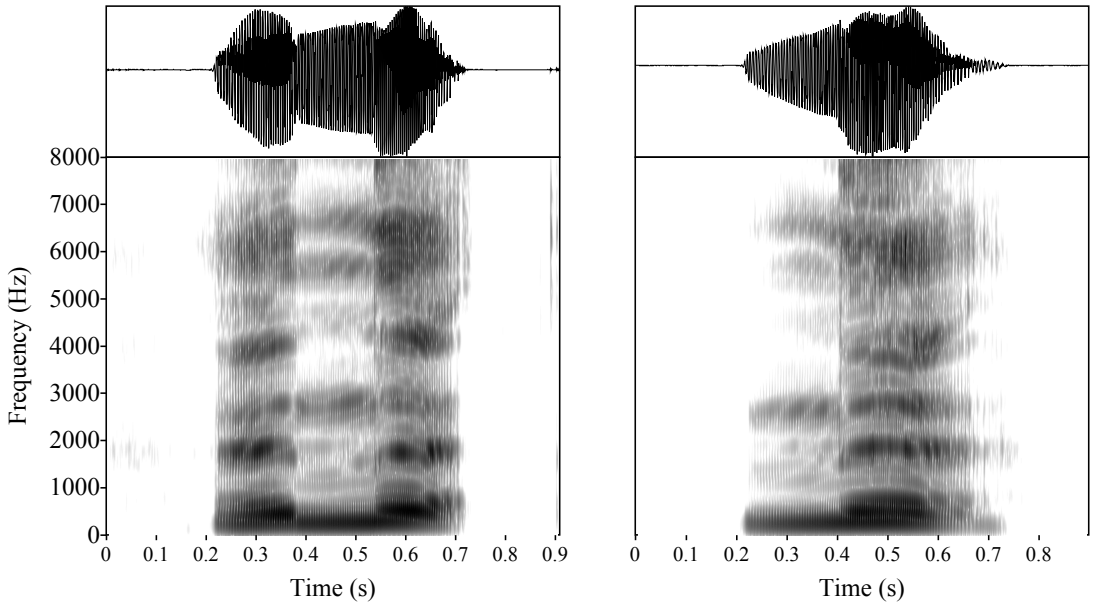
Ersu has seven voiceless (unaspirated) stops, affricates, and fricatives (/əp, ət, ək, ə̪ʦ̪, əʈʂ, ətɕ, əx/) and seven voiced stops, fricatives, and nasals (/əb, əd, əg, əʐ̪, əm, ən, əŋ/) that can be preceded by a segment with a schwa-like formant structure. Compare the minimal pairs /ə̪ʦ̪é/ ‘interest’ vs. /ʦ̪é/ ‘cloud’, and /ənɛ́/ ‘to be heavy; to be deep’ vs. /nɛ́/ ‘two’ in Figures 13 and 14.

Figure 13: Waveforms and spectrograms for the minimal pair /ə̪ʦ̪é/ ‘interest’ vs. /ʦ̪é/ ‘cloud’



/əʈʂɛ́/ ‘interest’ /ʈʂɛ́/ ‘cloud’

Figure 14: Waveforms and spectrograms for the minimal pair /ənɛ́/ ‘to be heavy; to be deep’ vs. /nɛ́/ ‘two’

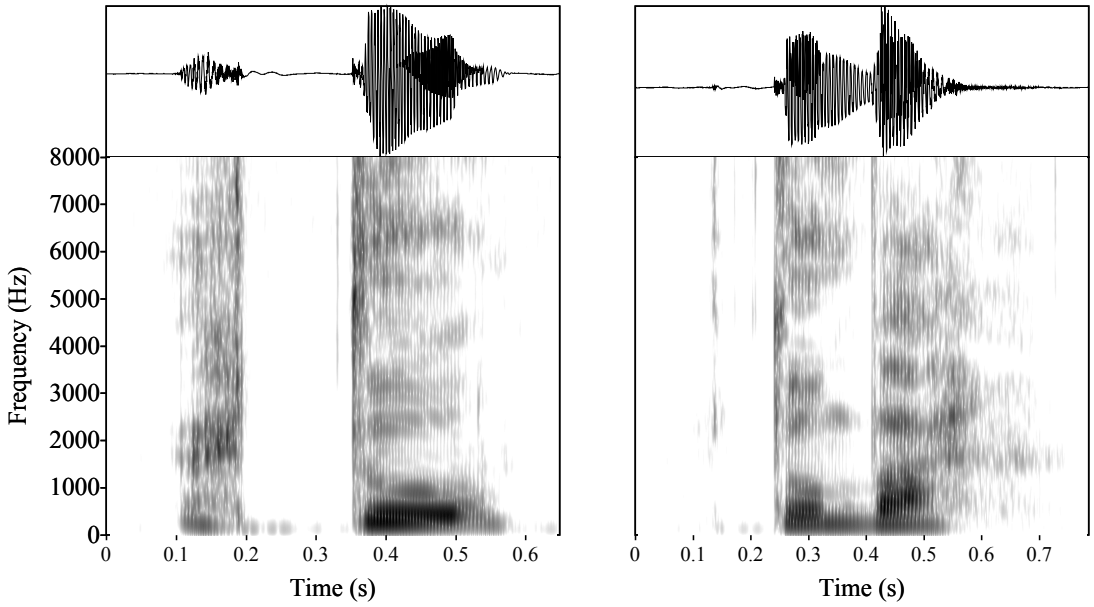


/ənɛ́/ ‘to be heavy; to be deep’ /nɛ́/ ‘two’

Clusters with a schwa-like segment can be tentatively identified with earlier consonant clusters. Independent evidence for such clusters can be found in Ersu words that etymologically correspond to Proto-Lolo-Burmese and Proto-Tibeto-Burman forms with initial clusters (cf. Bradley 1979:144, 1985:242, Matisoff 2003:37). For example, Ersu /ənɛ́/ ‘to be deep’ corresponds to Proto-Loloish *ʔ-nak^{L} ‘deep’ (Matisoff 2003:37) (see Chirkova & Handel 2013a for a detailed discussion). Clusters with a schwa-like segment are in the process of disappearing from this language, merging with corresponding simple onsets.[4]

In addition to clusters with a schwa-like segment, Ersu also marginally has preaspirated clusters, which are restricted to loanwords from Tibetan. Examples include /htóNbá/ ‘to be empty’ (WT *stong pa*), /hkwàʈá/ ‘to turn, to circle’ (WT *skor ba*). In these words, preaspiration diachronically derives from stop clusters with the preradical *s-* in Old Tibetan. The acoustic quality of Ersu preaspirated clusters is different from voiceless segments preceded by a schwa-like segment. The main differences include (i) the longer duration and (ii) the clear formant structure of the schwa-like segment, as compared to preaspiration. These differences are illustrated in Figure 15 with the pair /ətó/ ‘to jump’ vs. /htóNbá/ ‘to be empty’.

Figure 15: Waveforms and spectrograms for the pair /ətó/ ‘to jump’ vs. /htóNbá/ ‘to be empty’



/ətó/ ‘to jump’ /htóNbá/ ‘to be empty’

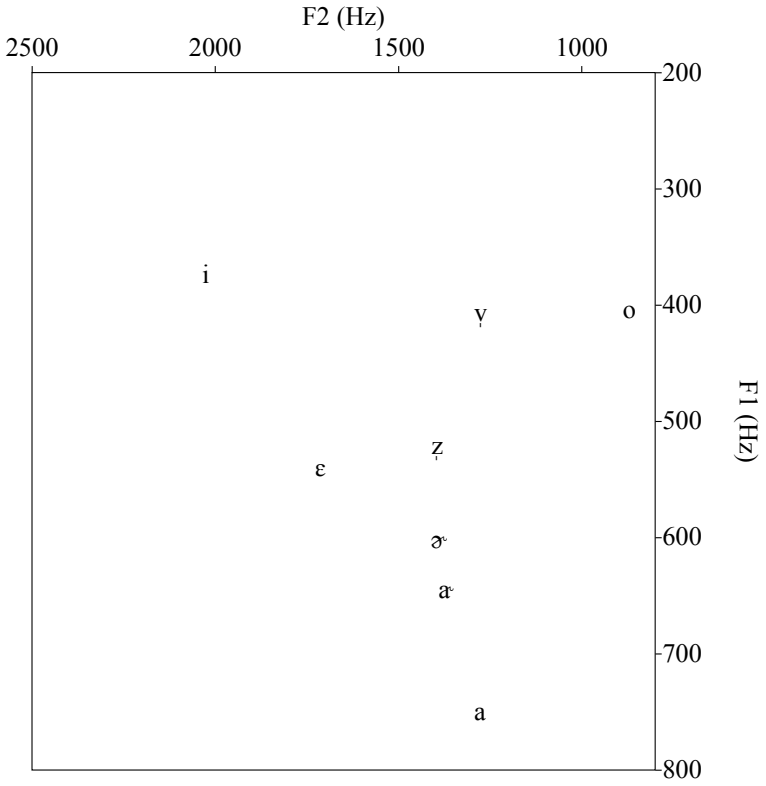
Vowels

Ersu has eight vowel phonemes, of which four are plain (/i, ɛ, a, o/), two are fricative (/z̩, ʐ̩/), and two are rhotacized (/ə˞, a˞/). See the vowel chart plotted on the relative F1/F2 formant values.

i	íz̪á	‘son’	pí	‘dregs’	t̪s̪í	‘hair’
ɛ	ɛ́lá	‘ram’	pɛ́	‘to give as a present’	t̪s̪ɛ́	‘cloud’
a	á	‘I’	pá	‘place of provenance’	t̪s̪á	‘to tie’
o	ódʑ̪á	‘pear’	pó	‘wrapping’	t̪s̪ó	‘grain cabinet’
z̩	ʐ̩̀	‘snow’	pʑ̩́-pź̩	‘to be flat’	t̪s̪ʑ̩́	‘to feed’

ɣ̩	ý̩	‘wine’	pý̩	‘potato’		ʈʂý̩	‘hearth’
aˠ			pàˠ	‘grain’			
əˠ	ə́ˠ	‘year’	pəˀ [pz̩̀]	‘thin rope’			

Table 2: Ersu vowel phonemes



Plain vowels

/i/ has a fairly broad range of realization. It ranges from [i] (essentially in loan vocabulary from SW Mandarin and Nuosu) to [jẹ] (mostly after bilabial initials) and [e] (mostly after dental, alveolar, and alveolopalatal initials). The value range for F1 is between 264 and 347 Hz, and for F2 between 1920 and 2274 Hz. Examples include: /ɕíkwá/ [ɕíkwá] ‘watermelon’ (Chinese *xīguā* 西瓜, SW Mandarin /ɕi^{44}kua^{44}/), /íʈʂʰý̩/ [éʈʂʰý̩] ‘ladle, a long handled spoon,

generally made of wood' (Nuosu *it chyp* [i^{55}ʈʂ^{h}ɻ^{21}]);[5] /pí/ [pjé̱] 'dregs', /mí/ [mjé̱] 'monkey', and /s̱í/ [s̱é] 'wood' (compare to /s̱ɛ́/ 'air').

After alveolar affricates, dental and alveolar fricatives, and velar stops, /ɛ/ may be realized as retroflex (characterized by lowered F3 values). For example, compare the realization of the word /z̩ɛ́/ 'to crawl, to climb' in isolation and the three repetitions of that word in the compound /Nbí z̩ɛ́/ 'to climb mountains'.

Fricative vowels

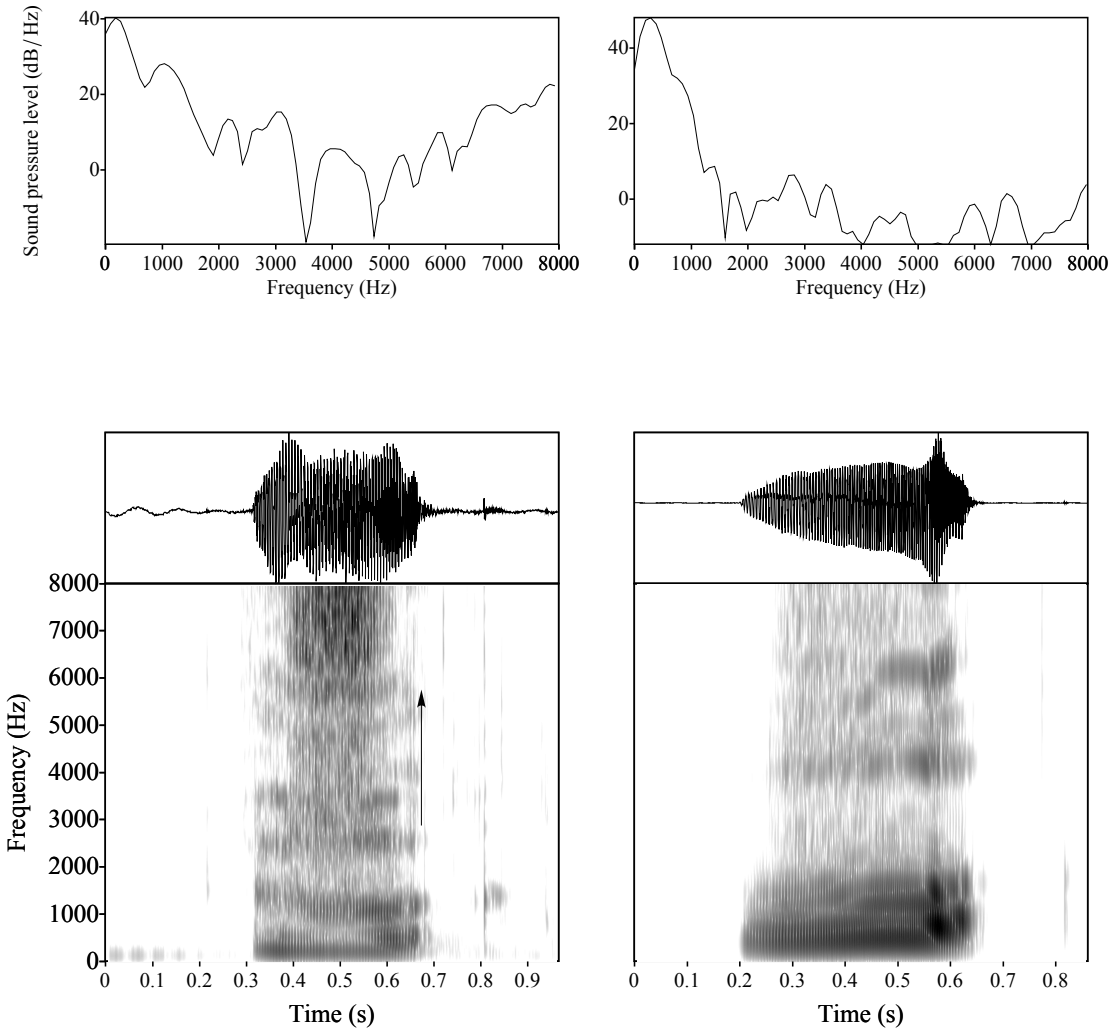
Fricative vowels (Ersu /z̩/ and /v̩/) are defined as vowels that are produced with the tongue in essentially the same position as in the corresponding fricatives (Ladefoged & Maddieson 1990:117, 1996:314). The constriction of the tongue tip or lips produces alveolar and labiodental frication, respectively.

The two fricative vowels in Ersu (/z̩, v̩/) are independent phonemes that co-occur with a broad range of initials. Therefore, they are distinct from the known cases of fricative vowels in Mandarin or Nuosu (e.g. Chao 1972 [1948], Li & Ma 1983:36, Ladefoged & Maddieson 1996:314), where syllabic fricatives can be viewed as conditioned variants of other (high) vowels.

Ersu fricative vowels display periodic vocal fold vibration and clear formant structure, as is typical of vowels. The two are differentiated by the configuration of the lips: spread for /z̩/ and rounded for /v̩/. /v̩/ is produced with a pronounced lip compression, whereby the lower lip is raised, while the upper lip remains in a static position (as characteristic for the articulation of labiodental fricatives, cf. Laver 1994: 250). Examples include /z̩̀/ 'snow', /v̩́/

‘wine’, /ʐ̪ʑ̪/ ‘shoe’ and /ʐ̪ʏ̪/ ‘oil’ (see also video clips ‘shoe’ and ‘oil’). Ersu fricative vowels are accompanied by fricative noise: the high frequency energy noise in the 3000Hz-6000Hz region for /ʐ̪/, and a relatively flat spectrum for /ʏ̪/ (as typical of labiodental fricatives) (cf. Ladefoged & Maddieson 1996:173-176). The fricative noise is more diffuse and weaker in intensity than that found for other fricatives (see Figure 16).

Figure 16: Acoustic spectra and waveforms and spectrograms of the words /ʐ̪̀/ ‘snow’, /ʏ̪́/ ‘wine’ (an arrow indicates high frequency noise in the 3000-6000 Hz region)



/z̀/ ‘snow’ /v́/ ‘wine’

The vowel /z̩/ may occur after bilabial stops, dental and alveolar affricates and fricatives, retroflex affricates, /ɽ/, and, in a very few cases, also after the voiced velar initials /g, Ng, ŋ/. It has a broad range of realization:

(1) After bilabial stop initials, /z̩/ is realized as [zz̩] after non-aspirated initials and as devoiced, [s̩], after aspirated stop initials. Examples include /bz̀/ [bzz̀] ‘bee’, /pź-pź/ [pzź-pzź] ‘to be flat’, /pʰz̀pó/ [ps̀pó] ‘wood shavings’.

(2) After dental and alveolar affricates and fricatives, retroflex affricates and /ɽ/, /z̩/ is realized as homorganic to the preceding consonant onset. Examples include /s̪z̀/ ‘matter, affair’ (Chinese *shì* 事, SW Mandarin /sɿ^{213}/), /z̪z̀/ ‘shoe’, /t̪s̪ź/ ‘to feed’, /t̪s̪ʰz̀/ ‘salt’, /s̪z̀/ [s̪s̀] ‘meat’, /z̪ź/ [z̪z̩] ‘to wear, to put on’, /ʈʈź/ [ʈʈ́] ‘star’, /ʈʈʰź/ [ʈʈʰ́] ‘to cut with a sickle’, /ɖʈź/ [ɖʈ́] ‘glutinous rice’, /ɽź/ [ɽ́] ‘to laugh’.

(3) After velar initials, /z̩/ is realized close to [ɤ], as in /gźɽóət̪s̪ź/ [gɤ́ɽóət̪s̪ź] ‘spine, backbone’ (compare to /gv́/ ‘boat’, /gó-gó/ ‘light (adj.)’).

/v̩/ has a broader distribution and may occur after bilabial and velar stops; dental, alveolar and retroflex affricates; nasals; and dental and alveolar fricatives. It may be realized as [v̩] or [ʋ̩]

in free variation (compare the two realizations of the word /gṷ/ ‘boat’). Similarly to /ẓ/, /ṿ/ has a fairly broad range of realization:

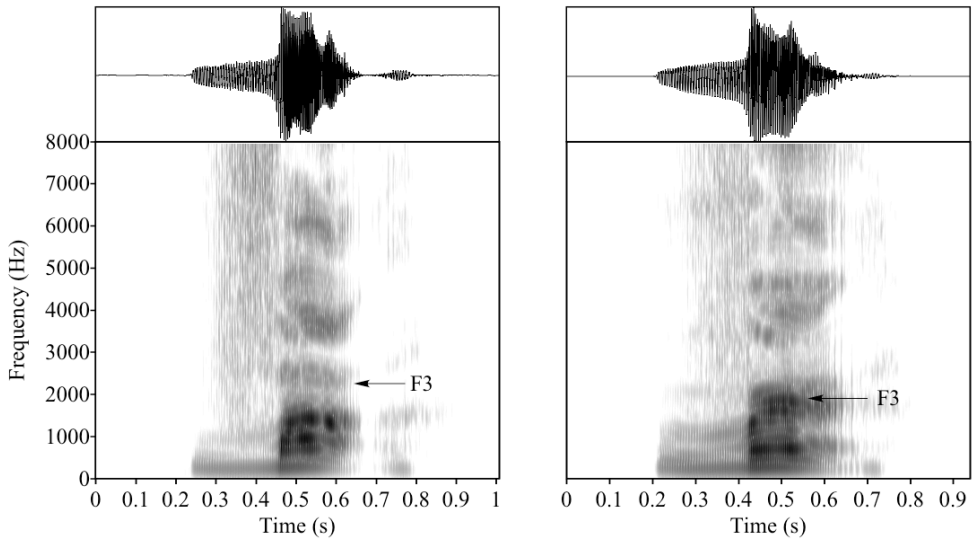
(1) After bilabial stop initials and retroflex affricates, /ṿ/ is realized close to [ʙ̥]. This is similarly to its realization in the closely related Lizu language. Examples include /pṷ/ [pʙ̥] ‘potato’, /tʰè-pʰṷ/ [tʰè-pʰʙ̥] ‘to change, to turn into, to transform’, /bṷ/ [ʙ̥] ‘wild cat’, /ʈʈṷ/ ‘sweat’, /ʈʈʰṷ/ ‘six’, /Nɖʈṷ/ ‘tile’ (see also Figure 6 and video clips ‘potato’, ‘to change, to turn into, to transform’, ‘wild cat’). The bilabial trill is particularly evident in the minimal contrastive pair /bṷ/ [ʙ̥] ‘wild cat’ vs. /bó/ [bó] ‘to have, to possess’.

(2) After /m/, /ṿ/ is realized as a voiced syllabic bilabial nasal ([m̩]). For example: /mṿ̀ʈʂɛ́/ [m̩̀ʈʂɛ́] ‘carpenter’, /mṿ́ʈʂɛ́/ [ḿ̩ʈʂɛ́] ‘cat’.[6]

Rhotacized vowels

Ersu has two rhotacized vowels: /ɚ̀/ (as in /ɚ̀/ ‘ashes; year; to be white; to bark’, /xɚ̀mìNtɕʰí/ ‘south’, reportedly from Nuosu (*yyx*) *hmy* [(ʑɻ^{34})m̥ɳ^{33}] ‘south’) and /a˞/ (as in /vá˞/ ‘slave’, /xá˞/ ‘bear; needle’). Compared to their oral counterparts (/ẓ/ and /a/, respectively), /ɚ̀/ and /a˞/ have a lowered frequency of the third formant. This is illustrated in Figure 17 with the minimal pair /vá/ ‘net’ vs. /vá˞/ ‘slave’.

Figure 17: Spectrograms of /vá/ ‘net’ vs. /vá˞/ ‘slave’ (arrows indicate F3)



/vá/ ‘net’ /və́/ ‘slave’

/ə̀/ has a broad distribution and co-occurs with bilabial and velar stops, /m/, /ŋ/, /v/, and /x/. (After bilabial initials, /ə̀/ may be realized as the sequence [ɹa], e.g. /bə́/ [bɹá] ‘to be full’, /mə́/ [mɹá] ‘to sleep’.) Conversely, /ə̀/ mostly occurs in isolation (as in /ə́/ ‘ashes; year; to be white; to bark’). In addition, /ə̀/ also occurs after bilabial stops and /x/ (as in /xə̀mìNtɕʰí/ ‘south’). After bilabial stops /ə̀/ is realized with frication. This is similar to the realization of /ʐ̩/ after bilabial stops. However, in contrast to /ʐ̩/, /ə̀/ is pronounced with the tongue curved in a convex shape. Examples include /bə́/ [bʑ̩́] ‘crown of a head’, /pə̀/ [pʑ̩̀] ‘thin rope’, /pʰə́/ [pʂ̩́] ‘Tibetan’.

Overall, Ersu can be said to have an unbalanced vowel system in that it has a high front vowel /i/ without a corresponding high back vowel /u/; a back close-mid vowel /o/ without a

corresponding front close-mid vowel /e/; and an open-mid low vowel /ɛ/ without a corresponding back open-mid vowel /ɔ/. In our analysis, this may be due to an ongoing realignment of the vowel system of Ersu, following the development of the fricative vowels /ʐ̩/ and /ʋ̩/ from the high vowels /i/ and /u/ in this language. The development of the fricative vowels from /i/ and /u/ is suggested on the one hand, by the synchronic distributional evidence, considered in the light of aerodynamic constraints outlined in Ohala (1983), and on the other hand, by comparative evidence from the closely related Lizu and Duoxu languages. We note that Ersu /ʐ̩/ does not co-occur with (dental and velar) stop initials, but it co-occurs with affricate initials instead. The vowel /ʋ̩/ does not co-occur with dental stop initials, but it co-occurs with alveolar affricate initials. This complementary distribution can be explained as an outcome of sound change whereby dental and velar stops developed an affricated release when followed by high vowels. This is due to the fact that the high velocity of the airflow created upon release of a stop lasts longer when the stop precedes a close vowel as opposed to an open vowel (Ohala 1983:204-205). From a comparative perspective, Ersu /ʐ̩/ has multiple correspondences in Lizu and Duoxu, including /i/ (as in 'bee': Ersu /bʑ̩́/, Lizu /^{R}bi/, Duoxu /bi^{31}/; 'shoe': Ersu /ʐ̈ʑ̩́/, Duoxu /ʑ̩i^{33}/), /e/ (as in 'hair': Ersu /t̪s̪í/, Lizu /^{F}tɕe/) as well as non-high vowels preceded by the palatal approximant /j/ (as in 'mountain': Ersu /Nbí/ [Nbjɛ́], Lizu /^{R}Nbje/) (see Chirkova & Handel 2013b). These correspondence patterns reveal complex developments, which contribute to a realignment of the vowel system of Ersu, whereby the earlier phoneme /e/ is moving into the vacated /i/ space.

Nasalized vowels

Ersu marginally has a set of nasalized vowels. Nasalized vowels are generally restricted to recent loanwords from Mandarin Chinese, where the donor language has the nasal codas /n/ or /ŋ/ (as in /kǎ/ ‘steel’, Chinese *gāng* 钢, SW Mandarin /kaŋ^{44}/). (In addition, nasalized vowels are attested in two native Ersu words in our corpus, /ɔ̀ɔ́/ ‘goose’ and /ɛ̀ɛ́/ ‘duck’). For that reason, vowel nasalization in Ersu must be regarded as subphonemic, and only needs to be marked in those cases where it is unpredictable (i.e. in recent loanwords). It is interesting to note that in older loanwords, where the original nasal coda is followed by a syllable that begins with a vowel or a nasal or when it is word final, the original nasal element is in most cases lost without compensation, as in /ʈʂwá/ ‘brick’ from *zhuān* (SW Mandarin 砖 /ʈʂuan^{44}/), /pɛ̀tì/ ‘silver’ from *báidìng* 白锭 ‘white ingot’ (SW Mandarin /pei^{21}tin^{213}/). This is similar to the situation in the closely related Lizu and Duoxu languages (Chirkova & Chen 2013, Chirkova & Handel 2013b).

Vowel harmony

In disyllabic domains, we observe regressive vowel assimilation. The vowel qualities can be divided into two sets: (1) the low vowels /a, aʷ/, and (2) the remaining, non-low vowels, that is, /i, ɛ, o, ʐ̩, ɻ̩/. Vowel harmony appears to only apply to directional prefixes and the number ‘one’ (that is, it is restricted to high frequency morphemes). Consider expressions consisting of the numeral ‘one’ (/tɛ́/ in isolation) followed by various nouns: /tá ʐ̩á/ ‘one hundred’, /tá ká/ ‘one strip’, /Nʈʂʰè tà páʷ/ ‘one grain of rice’, /tɛ́ pʰó/ ‘one set (of clothing)’, /tɛ́ pɻ̩́/ ‘one

tree; ten cents', /tɛ́ ḍẓ́ḍ́/ 'one sentence'. Like many languages that display vowel harmony in polysyllabic lexical items Ersu has two forms for affixes, such as verbal directional affixes, e.g. /kʰá-lá/ 'to come in (in the direction to the speaker)' vs. /kʰɛ́-jí/ 'to enter (in the direction away from the speaker)'.

Syllable Structure

The canonical Ersu syllable minimally consists of an obligatory nucleus and a tone. It may also contain up to three optional elements in the following linear structure: (C1)(C2)(C3)V, where C1 can be nasal (/N/) or a schwa-like segment (/ə/); C2 can be any consonant; C3 can be either /w/ or /j/; and V stands for vowel, and parentheses indicate optional constituents. Zero-initial words can be preceded by a non-phonemic glottal stop (e.g. /ɔ̀ɑ́/ [ʔɔ̀ɑ́] 'goose', /ódẓá/ [ʔódẓá] 'pear').

(1)	V	/á/ 'I, first person singular pronoun'		
(2)	CV	/tá/ 'flag' (WT *dar*)	/tʰá/ 'millstone'	/bò/ 'string'
(3)	CCV	/ətá/ 'wether, castrated ram'	/tʰwá/ 'mule'	/Nbò/ 'horse'
(4)	CCCV	/Ntʰwá/ 'to be sharp'		

In addition, (recent) loanwords may have the following structures:

(5) CVŋ /pǐtʰàŋ/ ‘crystal sugar’ (from Chinese *bīngtáng* 冰糖, SW Mandarin /pin^{44}tʰaŋ^{21}/)

(6) CVV(V) /fɛ̌tʰiào/ ‘rice noodle’ (from Chinese *fěntiáo* 粉条, SW Mandarin /fen^{53}tʰiao^{21}/)

Loanwords from Tibetan may also have /h/ in the C1 slot, as in /hkwàɽá/ ‘to turn, to circle’.

Similarly to its linguistic neighbours, Ersu is phonologically monosyllabic with a strong tendency towards disyllabicity in its lexicon. Trisyllabic and quadrisyllabic words are mostly composite, e.g. /lɛ́mákâ/ ‘thumb’ (< /lɛ́/ ‘hand’); /ʂ̩ʏ́Nbʏ́ əpɛ́əkʏ́/ ‘nostril’ (< /ʂ̩ʏ́Nbʏ́/ ‘nose’, /əpɛ́əkʏ́/ ‘hole, cavity’), although a handful of trisyllabic monomorphemic words (both native and loanwords) do exist (e.g. /xə̀mìNtɕʰí/ ‘south’).

In disyllabic composite forms, where the second syllable has zero initial, the two adjacent vowels merge into one vowel or a diphthong, a process that typically results in a long vowel or diphthong, as well as a tone change. This change characteristically occurs when the perfective marker /á/ (which has an etymological high-register tone) is added to a verb stem. For example, compare the realization of the verb /dɛ̀-pʰwá/ [dà-pʰʷá] ‘to smash’ in isolation and when followed by the perfective marker /á/, i.e. /dɛ̀-pʰwà-á/ ‘have smashed’.

The diminutive morpheme /ji/ often fuses with the preceding vowel resulting in a diphthong that combines the original vowel with the offglide [j]. Examples include: ‘armpit’, which is /jíbá-jì/ in careful pronunciation and /jíbáj/ in a more rapid speech tempo (note a

clear falling tone contour on the second syllable of the fused form, resulting from the low tone of /jì/ being conjoined with the high tone of /bá/); 'lamb', which is /lábɻ́-jì/ in careful pronunciation and /lábɻ́j/ in a more rapid speech tempo.

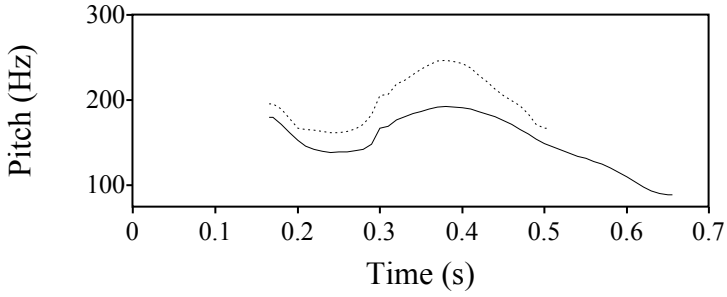
Prosodic organization

Ersu is a register tone language with two registers: Low and High (hereafter L and H), thus bearing resemblance to the tonal system of Tibetan (e.g. J. Sun 1997). In polysyllabic domains, there is tone reduction in non-initial syllables, resulting in highly restricted tone patterns in polysyllabic words.

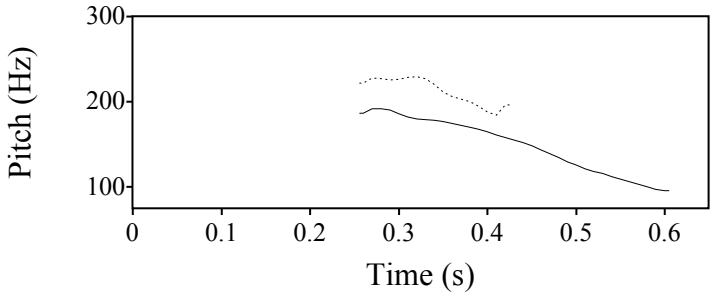
Monosyllabic words and compounds

The register contrast in Ersu can be exemplified with the following minimal pairs: /bɻ̀/ 'to plough' vs. /bɻ́/ 'wild cat', /là/ 'musk deer' vs. /lá/ 'fertilizer, manure', /Nɖʈò/ 'dirt, filth' vs. /Nɖʈó/ 'leather, skin', /ʈʈʰò/ 'dog' vs. /ʈʈʰó/ 'sound, melody'. While the register contrast is fundamental in this language, surface pitch contours are subject to variation in both registers (as illustrated in Figure 18).

Figure 18: Pitch contours of the low and high register tones by a male speaker illustrated with the minimal pairs /bɻ̀/ 'to plough' vs. /bɻ́/ 'wild cat', and /ʈʈʰò/ 'dog' vs. /ʈʈʰó/ 'sound, melody'. The low tone contour is represented by a solid line, the high tone contour is represented by a dotted line.



/bə̀/ ‘to plough’ and /bə́/ ‘wild cat’



/ʈʈʰò/ ‘dog’ and /ʈʈʰó/ ‘sound, melody’

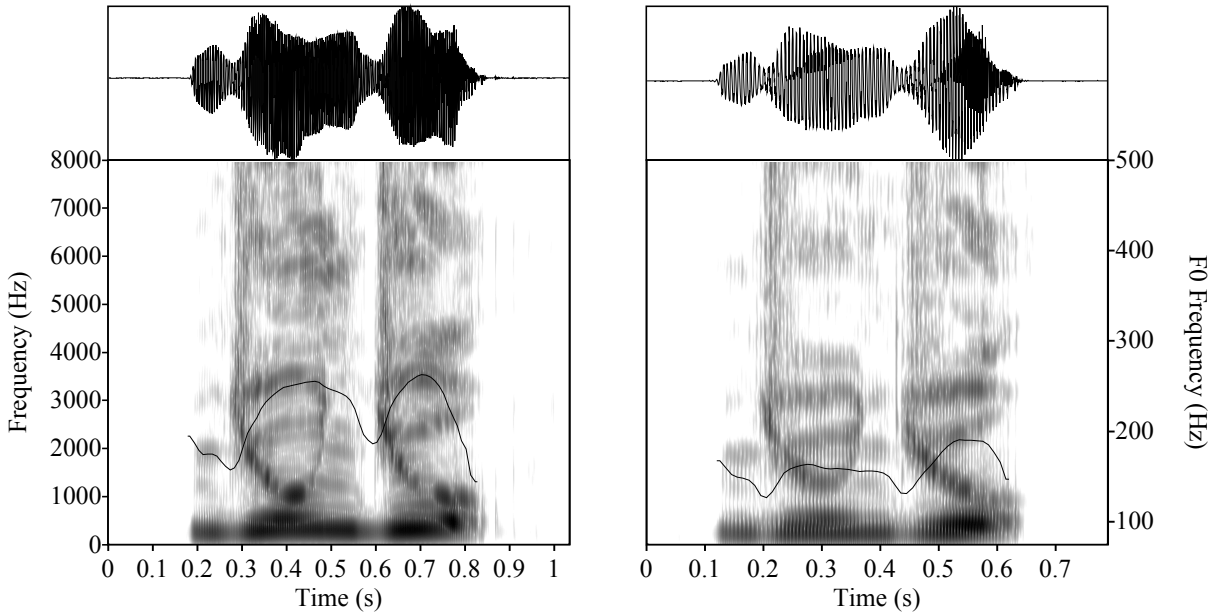
We note that monosyllabic words that are rarely used in isolation (such as verbs and measure words) and loanwords are often realized in the context of elicitation in the high register, which we, for that reason, regard as the unmarked register in this language.

Polysyllabic words and compounds

In disyllabic domains (both compounds and composite lexical words), the domain-initial syllable retains its tone; whereas the non-initial syllable does not. Disyllabic words and compounds that begin with a morpheme or a word in the high register are invariably realized with H tone on both syllables. Conversely, in the case of disyllabic words and compounds that begin with a morpheme or a word in the low register, two tonal patterns are possible: (i) L tone on both syllables, and (ii) L tone on the domain-initial syllable and H on the second syllable. The two patterns are in free variation, as is the case in the word /ỳɽà/ 'cloth': [vụ^{33}ɽra^{44}] or [vụ^{33}ɽra^{31}]. Examples of the tonal patterns on disyllabic domains include: (i) the compounds /ŋwà-dʑí/ 'cow shed' and /Nbò-dʑí/ 'horse pen', which begin each with a L register word (/ŋwà/ 'cow' and /Nbò/ 'horse', respectively; the word 'shed, pen' has an etymological low-register tone, i.e. /dʑì/); and (ii) the compounds /vé-dʑí/ 'pig shed' and /jó-dʑí/ 'sheep pen', which begin each with a H register word (/vé/ 'pig' and /jó/ 'sheep', respectively).

The same two patterns are attested in composite lexical words. Examples of minimal pairs for the two tonal patterns include: /Ndʑó-Ndʑó/ 'to make friends', /Ndʑò-Ndʑó/ 'ear of millet' (see Figure 19); /njó-njó/ 'milk; breast', /njò-njó/ 'to be soft'; /tɕó-tɕó/ 'to wrap, to bind', /tɕò-tɕó/ 'maternal uncle' (Chinese *jiùjiu* 舅舅, SW Mandarin /tɕiəu^{213}tɕiəu^{44}/); /k^{h}é-jí/ 'to enter', /k^{h}è-jí/ 'to go live (inside)'.

Figure 19: Illustration of the H vs. L contrast in disyllabic items /Ndʐó-Ndʐó/ ‘to make friends’ and /Ndʐò-Ndʐó/ ‘ear of millet’ by a male speaker



/Ndʐó-Ndʐó/ ‘to make friends’ /Ndʐò-Ndʐó/ ‘ear of millet’

In addition, Ersu disyllabic words marginally have a third tonal pattern, in which the high F0 peak is realized before the end of the first syllable, where the pitch starts to fall already and continues to fall in the second syllable. Perceptually, the second syllable of these words sounds much less prominent than the first syllable, giving rise to the impression that it is unstressed. This tonal pattern is essentially observed in Chinese loanwords, in fusions as well as in a handful of native Ersu words. Examples include /fɛ̋tʰiào/ ‘rice noodle’, /pǐtʰaŋ̀/ ‘crystal sugar’, /ʈʂʰálà/ ‘tomb’, /ʂɛ́tɛ̀/ ‘who’.

The same three tonal patterns are observed in trisyllabic compounds. These include:

(i) H tone on all syllables within the domain, for those domains that begin with an H tone word (HHH), as in:

/z̠v̩́/ ‘fish’ + /mέNṭṣʰέ/ ‘tail’ > /zv̩́ mέNṭṣʰέ/ ‘fish tail’

/jó/ ‘sheep’ + /Nɖɽv̩̀pí/ ‘skin, leather’ > /jó Nɖɽv̩́pí/ ‘sheep skin’

(ii) L tone on all syllables of the domain, for those compounds that begin with an L tone word (optionally, with the H tone on the last syllable of the domain) (LLL/H), as in:

/Nbò/ ‘horse’ + /mέNṭṣʰέ/ ‘tail’ > /Nbò mὲNṭṣʰέ/ ‘horse tail’

/ŋwɚ̀/ ‘cow’ + /Nɖɽv̩̀pí/ ‘skin, leather’ > /ŋwɚ̀ Nɖɽv̩̀pí/ ‘cowhide’

(iii) The third, less frequent, tonal pattern is that in which the domain of tone change appears to be restricted to the first two syllables, whereas the remaining syllable(s) is realized with a low tone (HHL), as in:

/ṣέ/ ‘iron’ + /Nd̠z̠v̩́-Nd̠z̠v̩́/ ‘button’ > /ṣέ Nd̠z̠v̩́-Nd̠z̠v̩̀/ ‘iron button’

/əmέ/ ‘soldier’ + /Nd̠z̠òmó/ ‘official’ > /əmέ Nd̠z̠ómò/ ‘general’

/ɚ́-s̠v̩́/ ‘the Ersu people’ + /xó/ ‘language’ > /ɚ́-s̠v̩́ xò/ ‘the Ersu language’

Ersu function words and discourse particles (e.g. the genitive particle /ji/, the focus particle /n̠ε/ in the recorded text) are never pronounced in isolation. Their surface tone realization depends on the tone of the preceding (host) lexical word (as in tonal contours in compounds).

Transcription of the recorded text 'The North Wind and the Sun'

The original audio and video recordings (made with a Digidesign 003 Rack soundcard, Pro Tools LE software for iMac, an AKG C520L headset microphone, and video cameras Sony HDR-XR 520E and Sony HDR-PJ650) have been made available to the JIPA along with this analysis.

Semi-narrow phonetic transcription

tɛ́ njó tɛ́ kɛ́ nɛ́ | mɛ́ə̀ lá máʈʂʰá = ɖʐ́ ɖɛ̀-ɖʐ̀-ɖʐ̀-á dʑ̀ì || nɛ̀ nɛ́ = wó tʰɛ́ tá-wá xíbá ká | xíbá ká xá nɛ́ nɛ́ = wó dɛ̀-Ndʑ̀ì-Ndʑ́í | ʂɛ́tɛ́ = wó xá nɛ́ ʂɛ́ʈɛ́ = wó já-ŋá dʑ̀ì-á | ʂɛ́tɛ́ = wó xá nɛ́ ʂɛ́tɛ́ = wó ʂòmó jà-ɖʐ́ó dʑ̀ì-á | Ndʑ́í-Ndʑ́í || nɛ́-Ndʑ́í-Ndʑ́í nɛ́-Ndʑ́í-Ndʑ́í nɛ́ tʰɛ̀-á tɛ́ = wó tɛ́-á tʰà-kʰá mà = pʰá | tʰɛ́ = kɛ́ nɛ́ Nɖʐ́á ʈʐ́-kʰwá Nɖʈʐ̀ʂ̀-ʂʐ́ tɛ́ ŋà-lá dʑ̀ì-gɛ̀-á | Nɖʈʐ̀ʂ̀-ʂʐ́ lɛ́ ətɕí Ngámɛ́ já-ʈʂʰá jà-kʰwá tá-ʈʂʰá dɛ̀-ʐʐ́ ʈʂá dʑ̀ì-gɛ́ || nɛ̀ nɛ́ = wó tʰɛ́ ná-tɕá-jí | jó = ɖʐ́ ʂɛ́tɛ́ = wó já-ŋá nɛ́ tʰɛ́ Nɖʈʐ̀ʂ̀-ʂʐ́ Ngámɛ́ jà-kʰwà = ʈʂʰà | tʰà-kwá ná-lá ʂʐ́ tʰɛ́-já dʑ̀ì-á || nɛ́ mɛ́ə̀ = wó | jò ʂó tɕʰó-gɛ́ tʰɛ́-já dʑ̀ì-gɛ̀-á || mɛ́ə̀ = wó dɛ̀-dɛ̀-mà nɛ́ | átʰɛ́ Nɖʈʐ̀ʂ̀-ʂʐ́ Ngámɛ́ = ʈʂʰá tʰɛ́-tʰɛ́-pʐ́-pʐ́ | mɛ́ə̀ = wó dɛ̀-dɛ̀-ə̀ nɛ́ | Ngámɛ́ = ʈʂʰá tʰɛ́-tʰɛ́-pʐ́-pʐ́ | áNdʑ́í tɕʰò-á xá lá Ngámɛ́ = ʈʂʰá tʰà-kwà-jí ná-lá mà = pʰá || nɛ́ tʰɛ́ = kɛ́ nɛ́ máʈʂʰá = wó tɕʰó = gɛ́ tʰɛ́-já dʑ̀ì-á || máʈʂʰá = wó tɛ́ kʰó ŋà-sá xá nɛ́ | átʰɛ́ ʂʐ́ dà-ʈʂʰà-jí dà-là nɛ́ | Ngámɛ́ = ʈʂʰá ɖʐ́ógwá tʰà-kwà-jí ná-lá dʑ̀ì-gɛ́ || nɛ́ tʰɛ́ = kɛ́ nɛ́ mɛ́ə̀ = wó ɛ̀j | á ná kʰɛ̀-ʂʐ̀ má-tó-á tʰɛ́-já dʑ̀ì-á ||

Interlinear morphemic glossing

Abbreviations used in the gloss below follow the Leipzig Glossing Rules (LGR, http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Non-standard abbreviations (those not included in the LGR) are: AGT = agentive, ITSF = intensifying.

tɛ́ njó tɛ́ kɛ́ nɛ́ | mɛ́ɚ́ lá mát̺s̺ʰá = d̺z̺í

one day one LOC TOP wind and sunshine=DU

dɛ̀-d̺z̺v̩̀-d̺z̺v̩̀-á dʑì || nɛ̀ nɛ́ = wó tʰɛ́

upward-meet-meet-PFV say TOP two=item that

tá-wá xíbá ká | xíbá ká xá nɛ́ nɛ́ = wó

one-together conversation hit conversation hit time TOP two=item

dɛ̀-Ndʑì-Ndʑí | s̺ɛ́tɛ́ = wó xá nɛ́ s̺ɛ́tɛ́ = wó

upward-dispute-dispute who=item time TOP who=item

já-ŋá dʑì-á | s̺ɛ́tɛ́ = wó xá nɛ́ s̺ɛ́tɛ́ = wó

ITSF-ruthless say-PFV who=item time TOP who=item

ṣòmó jà-ḍẓó dʑì-á | Ndʑí-Ndʑí || nɛ́-Ndʑí-Ndʑí

strength ITSF-exist say-PFV dispute-dispute downward-dispute-dispute

nɛ́-Ndʑí-Ndʑí nɛ́ tʰɛ́-á tɛ́ = wó tɛ́-á

downward-dispute-dispute TOP that-N-AGT one=item one-N-AGT

tʰà-kʰá mà = pʰá | tʰɛ́ = kɛ́ nɛ́ Nḍẓá ɽ̩́-kʰwá

outward-win NEG=be.able that=LOC TOP Chinese road-big

Nɖɽv̩̀ṣẓ̩̀-s̺v̩́ tɛ́ ŋà-lá dʑì-gɛ̀-á | Nɖɽv̩̀ṣẓ̩̀-s̺v̩́ lɛ́

walk-NMLZ.AGT one outward-come say-PROG-PFV walk-NMLZ.AGT FOC

ətɕí Ngámɛ́ já-ʈʂʰá jà-kʰwá tá-ʈʂʰá dè-ʐ́ʑ́

really clothes ITSF-set ITSF-big one-set upward-wear

ʈʂá dʑì-gɛ́ || nè nɛ́=wó tʰɛ́ ná-tɕá-jí | jó=ɖʐí

DUR say-PROG you you=item that downward-agree-RES self=DU

ʂɛ́tɛ́=wó já-ŋá nɛ́ tʰɛ́ Ndʅ̀ʂ̀-ʂ́ Ngámɛ́

who=item ITSF-ruthless TOP that walk-NMLZ.AGT clothes

jà-kʰwà=ʈʂʰà | tʰà-kwá ná-lá ʂ́ tʰɛ́-já

ITSF-big=set here-take.off downward-come CAUS that-kind

dʑì-á || nɛ́ mɛ́ə́=wó | jò ʂó tɕʰó-gɛ́ tʰɛ́-já

say-PFV TOP wind=item self first perform-PROG that-kind

dʑì-gè-á || mɛ́ə́=wó dè-dè-má nɛ́ | átʰɛ́

say-PROG-PFV wind=item upward-upward-blow TOP that

Ndʅ̀ʂ̀-ʂ́ Ngámɛ́=ʈʂʰá tʰɛ́-tʰɛ́-pʼ́-pʼ́ | mɛ́ə́=wó

walk-NMLZ.AGT clothes=set here-here-wrap-wrap wind=item

dè-dè-ə́ nɛ́ | Ngámɛ́=ʈʂʰá tʰɛ́-tʰɛ́-pʼ́-pʼ́ |

upward-upward-blow TOP clothes=set here-here-wrap-wrap

áNdʑí tɕʰó-á xá lá Ngámɛ́=ʈʂʰá tʰà-kwà-jí

how perform-PFV time and clothes=set here-take.off-RES

ná-lá mà=pʰá || nè tʰɛ́=kɛ́ nɛ́

downward-come NEG=be.able TOP that=LOC TOP

máʈʂʰá = wó	tɕʰó-gɛ́	tʰɛ́-já	dʑì-á ||	máʈʂʰá = wó			
sunshine=item	perform-PROG	that-kind	say-PFV	sunshine=item			
tɛ́	kʰó	ŋà-ʂá	xá	nɛ́|	átʰɛ́	ʂʏ́	dà-ʈʂʰà-
one	shine	outward-put	time	TOP	that	person	upward-hot-
jí	dà-là	nɛ́ |	Ngámɛ́ = ʈʂʰá	ɖʐógwá	tʰà-kwà-jí		
RES	upward-come	TOP	clothes=set	immediately	here-take.off-RES		
ná-lá	dʑì-gɛ́ ||	nɛ́	tʰɛ́ = kɛ	nɛ	mɛ́ə́ = wó	ɛ̀j |	
downward-come	say-PROG	TOP	that=LOC	TOP	wind=item	ITRJ	
á	ná	kʰɛ̀-ʂʐ́	má = tó-á	tʰɛ́-já	dʑì-á ||		
I	you.N-AGT	inward-defeat	NEG=can-PFV	that-kind	say-PFV		

Acknowledgements

We would like to thank our Ersu language consultants, Mrs. Wang Ying 王英 (Jiman Keli 吉满柯莉) and Mr. Chen Guofu 陈国富 for their patience and assistance. We are grateful to the Laboratoire de Phonétique et Phonologie (LPP) of the Centre National de la Recherche Scientifique (CNRS) for making their phonetic laboratory available to us through the operation “Phonetic and Phonological Complexity” of the project LabEx-EFL (Laboratoire d’excellence “Fondements Empiriques de la Linguistique—Empirical Foundations of Linguistics”). We are also grateful to Jacqueline Vaissière, Zev Handel, and the two anonymous reviewers of this illustration for helpful comments and suggestions, and to Matthew Faytak for a discussion on fricative vowels. Thanks are also due to Jos Pacilly (Leiden University) for helping us with PRAAT scripts, to Franz Huber (ETH Zürich) for creating the map in this paper, and to Wang Ke 王轲 (Jiman Shalima 吉满莎丽玛) and Gu

Tao 古涛 for help with the recording of palatograms. We gratefully acknowledge the financial support of the Agence Nationale de la Recherche of France (grant number ANR-07-JCJC-0063) and LabEx-EFL (axe PPC2) to Katia Chirkova, Wang Dehe, and Tanja Kocjančič Antolík; of the Endangered Languages Documentation Programme (ELDP, grant number MPD0257) to Katia Chirkova, Wang Dehe, and Yiya Chen; and of the European Research Council (ERC-Starting Independent Researcher Grant, 206198) to Yiya Chen.

References

Anderson, Victoria B. 2008. Static palatography for language fieldwork. *Language Documentation and Conservation* 2(1), 1–27.

Bradley, David. 1979. *Proto-Loloish*. (Scandinavian Institute of Asian Studies, monograph series, 39). London & Malmo: Curzon Press.

Bradley, David. 1985. Nasality in Bisu and Bisoid. In: Ratanakul, Suriya, David Thomas, and Suwilai Premsrirat (eds.), *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*, 234-263. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.

Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), *Tibeto-Burman languages of the Himalayas (Papers in Southeast Asian Linguistics no.14)*, 1–71. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Bradley, David. 2008. Labiovelars in Adur. Presented at the Workshop on Tibeto-Burman Languages of Sichuan, Academia Sinica, Taipei, 21-22 November 2008.

Chao, Yuen Ren. 1972. *Mandarin primer: An intensive course in spoken Chinese*. Cambridge: Harvard University Press; London: Oxford University Press. First published in 1948 by Harvard University Press.

Chirkova, Katia. 2012. The Qiangic subgroup from an areal perspective: A case study of languages of Muli. *Language and Linguistics* 13(1), 133–170.

Chirkova, Katia. 2014. The Duoxu language and the Ersu-Lizu-Duoxu relationship. *Linguistics of the Tibeto-Burman Area* 37(1), 104–146.

Chirkova, Katia & Yiya Chen. 2013. Lizu (Illustrations of the IPA). *Journal of the International Phonetic Association* 43(1), 75–86.

Chirkova, Katia & Zev Handel. 2013a. Diachronic developments of voiceless nasals in Ersu, Lizu, and related languages. Presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, New Hampshire, August 7-10. [Written version of the presentation is available at https://www.academia.edu/7637522/, accessed July 2014.]

Chirkova, Katia & Zev Handel. 2013b. Fricativization of high vowels and glides in Ersu, Lizu, and related languages. Presented at the 3rd Workshop on Sino-Tibetan Languages of Sichuan, Paris, September 2-4.

Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques.* Malden: Blackwell Publishing.

Ladefoged, Peter & Ian Maddieson. 1990. Vowels of the world's languages. *Journal of Phonetics* 18:93-122.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Malden, MA: Blackwell Publishers.

Lama Ziwo. 1994. Namuyiyu zhishu yanjiu [On the genetic affiliation of the Namuyi language], *Minzu Yuwen* [Minority languages] 1, 50–60.

Laver, John. 1994. *Principles of phonetics.* Cambridge: Cambridge University Press.

Li, Lan. 2010. Sichuan Muli xian fangyan jilue [The Chinese dialect of Muli, Sichuan: A survey]. *Fangyan* [Dialects] 2, 114–133.

Li, Min & Ma Ming. 1983. *Liangshan yiyu yuyin gailun* [Introduction to the phonetics of the Yi language of Liangshan Prefecture]. Chengdu: Sichuan Nationalities Press.

Liu, Huiqiang. 2007. Ersuyu gaikuang [An outline of the Ersu language]. In Li Shaoming & Liu Junbo (eds.), *Ersu Zangzu yanjiu* [Studies on the Ersu Tibetans], 462–500. Beijing: Minzu Chubanshe. First published in 1983 in Sichuan sheng Minzu Yanjiusuo (eds.), *Minzu yanjiu lunwenji* [Collected papers on minority languages], vol. 1.

Marchal, Alain. 1988. *La palatographie*. Paris: Editions du Centre National de la Recherche Scientifique.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction.* Berkeley, CA: University of California Press.

Matisoff, James A. 2006. Much Adu about something: extrusional labiovelars in a Northern Yi patois. *Linguistics of the Tibeto-Burman Area*, 29(1), 95–106.

McGowan, R.S. 1992. Tongue-tip trills and vocal-tract wall compliance. *Journal of the Acoustical Society of America*, 91, 2903-2910.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Pan, Zhengyun. 2001. Yiyu Aduhua chunruan'e fufuyin shengmu bijiao yanjiu [A comparative study of labiovelar cluster initials in the Adu patois of the Yi language]. *Minzu Yuwen* [Minority languages] 2, 17–22.

Schmidt, Edwin. 2011. Commodification in an Ersu Tibetan Village of Sichuan, China. MA Dissertation: Oregon State University.

Spajić, Siniša, Peter Ladefoged, and P. Bhaskararao. 1996. The trills of Toda. *Journal of the International Phonetic Association* 26(1), 1–21.

Sun, Hongkai. 1982. Ersu (Duoxu) hua jianjie [An outline of Ersu (Duoxu)]. *Yuyan Yanjiu* [Linguistic Study] 2, 241–264.

Sun, Hongkai. 1983. Liujiang liuyu de minzu yuyan ji qi xishu fenlei [Minority languages of the Six River Valley and their genetic classification]. *Minzu Xuebao* [Scholarly Journal of Nationalities] 3, 99–274.

Sun, Hongkai. 2001. Lun Zang-Mian yuzu zhong de Qiang yuzhi yuyan [On the Qiangic branch of the Tibeto-Burman language family]. *Language and Linguistics* 2(1), 157–181.

Sun, Hongkai. 2009. The Ersu Shaba Pictographic Writing System. *Journal Asian Highlands Perspectives* 1, 159–186.

Sun, Jackson T.-S. 1997. The typology of tone in Tibetan. *Proceedings of the Symposium "Chinese Languages and Linguistics IV: Typological studies of languages in China" (Symposium Series of the Institute of History and Philology-Academia Sinica)*, 485–521. Taipei: Taiwan.

Wang, Dehe. 2010. *Ersu zangzu wenhua yanjiu* [Study of Ersu Tibetan culture]. Chengdu: Sichuan University Press.

Wang, Dehe. 2011. Ersu shaba xiangxingwen lishu yanjiu [Divination books and pictographic script of the Ersu Shaba priests]. *Cahiers de linguistique – Asie orientale* 40(2), 225–248.

Wu Da. 2010. Zuqunxing yu zuqun rentong jiangou—Sichuan Ersuren de minzuzhi yanjiu [The construction of ethnic identities among the Ersu of Sichuan]. Beijing: Minzu Chubanshe.

Yu, Dominic. 2012. *Proto-Ersuic*. Ph.D. Dissertation: University of California, Berkeley.

Zhang, Sihong. 2013. *A reference grammar of Ersu: A Tibeto-Burman language of China*. PhD Dissertation: James Cook University.

Zhang, Sihong. 2014. The expression of knowledge in Ersu. In Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds.). *The grammar of knowledge: A cross-linguistic typology*, 132-

147. Oxford: Oxford University Press.

---

[1] In transcriptions of Ersu words "-" stands for morpheme boundary and "=" stands for clitic boundary. In transcriptions, the Low tone is marked as "à", and the High tone is marked as "á" (for more detail, see section "Prosodic organisation").

[2] Tone notation in Lizu is provided in superscript letters to the left of the lexical word. The following abbreviations are used: "F" stands for falling tone, "R" for stands for rising tone, "EP" stands for "equally-prominent pattern", "LP" stands for "left-prominent pattern", and "RP" stands for "right-prominent pattern" (see Chirkova & Chen 2013 for a detailed discussion).

[3] SW Mandarin readings are based on Li (2010) and provided in the original transcription.

[4] We note that Ersu nasals preceded by a schwa-like segment correspond to what we impressionistically call geminated nasals in the closely related Lizu language. Our principal Lizu language consultant marginally distinguishes between, according to his own analysis, a relatively shorter nasal initial and a relatively longer ("geminated") nasal initial in a few minimal pairs, such as /^{F}ne/ 'you, thou' vs. /^{F}ne/ [^{F}nne] 'deep'. A more systematic comparison between Ersu words with this type of clusters and their cognates in Lizu will be undertaken in forthcoming fieldwork.

[5] Nuosu examples are from the online *Nuosu Yi-Chinese-English Glossary* (http://www.yihanyingcihui.net/?page_id=6065&lang=en), accessed January 2013. In the transcription system of Nuosu, final consonants (e.g. *t* and *p* in the word for 'ladle') are used to mark tone (respectively, the high level, 55, and the low falling tone, 21).

[6] We note that this distribution has a close parallel in most Nuosu varieties, where the voiced syllabic bilabial nasal [m̩] is the usual allophone of a syllable with a voiced or voiceless bilabial nasal initial followed by /u/ (Bradley 2008). Furthermore, as in Ersu, syllables with

the high vowels /i/ and /u/ are an environment conducive to the development of syllables with unusual phonetic forms in many Ngwi languages (such as Adur, various Nuosu or Liangshan Yi varieties, and Sanie, Pan 2001, Matisoff 2006, Bradley 2008) as well as some Qiangic languages, such as Namuzi (Lama Ziwo 1994). These unusual forms include syllabic fricatives or trills after initial bilabial and velar stops, and labial-velar nasals after initial bilabial and velar nasals.