संकटग्रस्त भाषाओं का प्रलेखन एवं क्षेत्रकार्य

डॉ. परमान सिंह
सीनियर रिसोर्स पर्सन (एस.पी.पी.एल)
भारतीय भाषा संस्थान मैसूर

भूमिका

भाषाओं के अध्ययन, संरक्षण एवं प्रलेखन का प्रथम सोपान भाषा की रिकार्डिंग एवं उसका लिप्यंकन है। किसी भाषा का चाहे उच्चारण के स्तर पर अध्ययन करना हो या उसका व्याकरण बनाना हो, महानगरों में हो रहे भाषा प्रयोग को जानना हो या शहरी सभ्यता से बहुत दूर पहाड़ों, जंगलों, रेगिस्तानों, टापुओं, द्वीपों इत्यादि में रचे-बसे लोगों की भाषा का अध्ययन करना हो अथवा समुदाय में सामाजिक संरचना के आधार पर भाषाई विविधता का निरुपण करना हो, तो उसका पहला सोपान डेटा संग्रहण होता है। डेटा संग्रहण के लिए क्षेत्रकार्य (Fieldwork) अनिवार्य है। हालाँकि विभिन्न परिस्थितियों में क्षेत्रकार्य की रणनीति विभिन्न होती है, यहाँ हमारा केंद्र बिंदु संकटग्रस्त भाषाओं के प्रलेखन के लिए किए जाने वाले क्षेत्रकार्य होगा। इस लेख में क्षेत्रकार्य के लिए आवश्यक साधनों, उपकरणों, सूचनाप्रदाता (Informant) के चुनाव के साथ साथ शोधकर्ता से अपेक्षित व्यवहार की भी विवेचना की जाएगी।

क्षेत्रकार्य

जिस भाषा को प्रलेखित करने का आप ने निश्चय किया है उस भाषा समुदाय में जाना, वहाँ से उस भाषा के बोलने वाले सक्षम व्यक्ति/व्यक्तियों से उस भाषा में डेटा रिकार्ड करना क्षेत्रकार्य है। वास्तव में मूल रूप में यही क्षेत्रकार्य है लेकिन अच्छे भाषा वैज्ञानिक क्षेत्रकार्य को थोड़ा लंबा रखते हैं और उस समुदाय में रहकर ही इकट्ठा किए गए आँकड़ों का डेटा तैयार करते हैं, आँकड़ों का विवेचन करते हैं, विवेचन के आधार पर उनकी व्याख्या करते हैं, उस भाषा का व्याकरण तैयार करते हैं, उस भाषा में संभवतया कुछ सृजन करते हैं एवं उसका कोश तैयार करते हैं। उनके लिए यह पूरी प्रक्रिया ही क्षेत्रकार्य है।

पीटर लैडफोगेड (2003:1) ने क्षेत्रकार्य की तुलना हृदय की सर्जरी से की है। लैडफोगेड के अनुसार एक डॉक्टर हृदय की सर्जरी को दूसरों के हृदय की सर्जरी कर के सीखता है ठीक उसी प्रकार दूसरों की भाषा के ऊपर प्रयोग करके ही आप क्षेत्रकार्य में दक्षता हासिल कर सकते हैं। ठीक इसी से मिलती-जुलती कहावत है कि “कटे काहू का, सीखे नाउ का” तात्पर्य यह है कि नाई दूसरों के बाल काट कर ही बाल काटने में दक्ष होता है।

भाषाओं के संदर्भ में क्षेत्रकार्य को करने के कई उद्देश्य हो सकते हैं लेकिन यहाँ सारा संदर्भ संकटग्रस्त भाषाओं का चल रहा है, तो हम क्षेत्रकार्य को करने के उद्देश्यों की चर्चा भी इसी दृष्टिकोण से करेंगे। भाषावैज्ञानिकों के द्वारा किए जाने वाले क्षेत्रकार्य के निम्न उद्देश्य हो सकते हैं-

1. **भाषा का प्रलेखन-** वर्तमान दौर अधिक से अधिक संकटग्रस्त भाषाओं के प्रलेखित कर लेने का दौर है। इसलिए ज्यादातर भाषावैज्ञानिकों के क्षेत्रकार्य का उद्देश्य संकटग्रस्त भाषाओं को प्रलेखित कर उनका, व्याकरण, उनका कोश इत्यादि बनाने से है।

2. **भाषा के द्वारा सांस्कृतिक प्रलेखन–** संकटग्रस्त भाषा जिस समुदाय में बोली जा रही है उसकी पारिवारिक, सामाजिक, सांस्कृतिक एवं धार्मिक संरचना का भाषा के माध्यम से अध्ययन करने एवं उनके देशी वैज्ञानिक ज्ञान को प्रलेखित करने के लिए।
3. **अनुवाद-** कुछ मिशनरियों ने बाइबिल का अनुवाद देशज भाषाओं में करने के क्रम में क्षेत्रकार्य किया और उन भाषाओं का अपनी सुलभता के लिए व्याकरण एवं कोश तैयार किया।
4. **भाषावैज्ञानिक परिकल्पना को प्रमाणित करने के लिए-** कुछ भाषाविद् थोड़े से आँकड़ों के आधार पर अपनी परिकल्पना कर लेते और फिर उनको प्रमाणित करने हेतु अधिक स्पष्ट और विश्वसनीय आँकड़ों के लिए क्षेत्रकार्य करते हैं। सैमरिन ने सही ही कहा है कि क्षेत्र (Field) भाषावैज्ञानिकों की प्रयोगशाला है (Samarin 1967: 4)।

डिक्सन (Dixon 2010:309-310) ने भी क्षेत्रकार्य के उद्देश्यों के संदर्भ में अपने विचार प्रकट किए हैं। इन्होंने क्षेत्रकार्य करने के उद्देश्यों को दो भागों में बाँटा है- मुख्य और गौण उद्देश्य। ऊपर जो उद्देश्य बताए गए हैं वे सारे डिक्सन साहब के वर्गीकरण में गौण उद्देश्यों में आते हैं। डिक्सन साहब के अनुसार क्षेत्रकार्य करने के जो दो मुख्य उद्देश्य होने चाहिए वे निम्न हैं-

1. **भाषाविज्ञान सीखने के लिए-** डिक्सन के अनुसार क्षेत्रकार्य ही एक ऐसा रास्ता है जिससे हम भाषाविज्ञान को वास्तव में सीखते हैं। पहले भाषाविज्ञान के सिद्धांतों को पढ़ा जाता है लेकिन उनका प्रयोग क्षेत्रकार्य में करके पूर्णतः समझा जाता है । औषधि विज्ञान का छात्र शरीर रचना एव शल्य क्रिया के बारे में पढ़ता है लेकिन सर्जन बनने से पूर्व वह वास्तविक शल्य को करके सीखता है और धीर धीरे वह सीख कर इतना परिपक्व हो जाता है कि भविष्य में शल्य क्रिया के सिद्दांतों पर किताब भी लिख लेता है। ठीक इसी प्रकार भाषावैज्ञानिक भी क्षेत्रकार्य के दौरान भाषाओं पर (जिन पर किए गए शोधों की संख्या नगण्य हो) वास्तविक रूप में भाषावैज्ञानिक सिद्धांतों का प्रयोग कर भाषाविज्ञान को सीखता है।
2. **नई भाषाओं पर कार्य करने का बौद्धिक उत्साह-** भाषावैज्ञानिक नई (अप्रलेखित) भाषाओं पर कार्य करके कुछ विशिष्ट निष्कर्षों के प्रति आशान्वित रहते हैं। ऐसे भाषावैज्ञानिक बौद्धिक उत्साह की संतुष्टि के लिए भाषावैज्ञानिक क्षेत्रकार्य में आनंद लेते है और उस भाषा का व्याकरण, कोश इत्यादि लिखकर पढ़ने वालों को भी आनंदित करते हैं। डिक्सन महोदय सामाजिक जिम्मेदारी समझकर संकटग्रस्त भाषाओं के प्रलेखन के लिए किए जाने वाले क्षेत्रकार्य को अच्छा उद्देश्य नहीं मानते। डिक्सन (Dixon 2010: 310)

माननीय डिक्सन जी द्वारा सूचित दोनों मुख्य उद्देश्य आदर्शवादी ज्यादा हैं व्यवहारिक कम। संकटग्रस्त भाषाओं के प्रलेखन के लिए किया गया क्षेत्रकार्य भाषा को विवेचित कर सीखने के उद्देश्य से कम संकटग्रस्त भाषाओं के अनुरक्षण एवं प्रलेखन के दबाव में ज्यादा किया जाता है। संकटग्रस्त भाषाओं का अनुरक्षण एवं प्रलेखन सिर्फ भाषावैज्ञानिकों का बौद्धिक उत्साह का परिणाम नहीं अपितु यह पूरा कार्य भाषावैज्ञानिकों, नृजातिवैज्ञानिकों, फोकलोरिस्टों, कंप्यूटर विशेषज्ञों एवं सबसे ज्यादा संकटग्रस्त भाषा बोलने वाले समुदाय के सहयोग से संपादित होता है। यदि व्यक्ति विशेष का भाषा विज्ञान सीखने में कोई बाधा (आर्थिक, पारिवारिक, सामाजिक, मानसिक कोई भी बाधा) आ जाती है या किसी कारणवश उसके बौद्धिक उत्साह में कम आनंद आने लगता है उस परिस्थिति में व्यक्ति शायद संकटग्रस्त भाषाओं के अनुरक्षण एवं प्रलेखन के निमित्त किए जाने वाले क्षेत्रकार्य को छोड़ सकता है परिणामस्वरूप यह कार्य अधर में लटक सकता है। इस प्रकार बहुतेरी संकटग्रस्त भाषाएँ विश्व की नज़र में आए बिना ही समाप्त हो सकती हैं। दूसरे शब्दों में व्यक्ति विशेष के उमंग में कमी संकटग्रस्त भाषाओं का भविष्य डुबो सकता है। मेंरी निगाह में सिर्फ रोजगार के लिए या आर्थिक फायदे को ध्यान में रखकर किया जाने वाला क्षेत्रकार्य प्रशंसनीय नहीं है।

क्षेत्रकार्य प्रारम्भ करने से पूर्व हमें यह सुनिश्चित कर लेना चाहिए कि हमें क्या रिकार्ड करना है। हालाँकि यह भी किसी भाषाविद् का यादृच्छिक निर्णय नहीं होगा अपितु रिकार्ड करने वाली सामाग्री का निर्धारण परियोजना की प्रकृति एवं परियोजना की समयावधि से अभिप्रेरित होता है। इसलिए हमें सर्वप्रथम व्याकरण बनाना है तो आपको उस भाषा में वाक्यों की रिकार्डिंग पर केन्द्रित होना होगा। कोश बनाना है तो सर्वप्रथम मूलभूत शब्दावली को केन्द्र में लाना होगा और स्वनिमिक तालिका (phonemic inventory) बनाना हो तो शब्द सूची (word list) की रिकार्डिंग करनी चाहिए जिससे उस भाषा में उपलब्ध सभी व्यंजन एवं स्वर चाहे वे स्वनिमिक हो या सह स्वनिमिक हो ये रिकार्डिंग हो जाये।

संकटग्रस्त भाषाओं के संरक्षण एवं प्रलेखन की भारतीय भाषा संस्थान की परियोजना के मुख्य तीन उद्देश्य हैं-

I. व्याकरण बनाना (स्वनिमिक विश्लेषण भी शामिल है)
II. त्रिभाषिक कोश बनाना
III. सामाजिक एवं सांस्कृतिक जानकारी संग्रहण एवं विश्लेषण

अच्छे से तैयार की गई शब्द सूची एवं वाक्य सूची के अलावा किसी भी क्षेत्रकार्य के प्रमुख सोपान निम्न हैं-

1. उपकरण एवं तकनीकी का चुनाव
2. सूचना प्रदाता का चुनाव
3. सूचना का प्रलेखन

इन सभी बिन्दुओं पर क्रमवार जानकारी नीचे दी जाएगी-

उपकरण एवं तकनीकी का चुनाव

भाषाविज्ञान के प्रारम्भिक दौर में जब तकनीकी इतनी विकसित नहीं थी तो भाषा के लिए डेटा संग्रहण पेन से पेपर पर होता था लेकिन जब टेप रिकार्डर का आविष्कार हुआ तो भाषाविदों नें उसका काफी उपयोग किया। 20 वीं शताब्दी के अन्तिम दशक तक भारत में डेटा रिकार्डिंग के लिए टेप रिकार्डिर का प्रयोग होता रहा है। कुछ समय पश्चात भाषा विश्लेषण में कंप्यूटरों का प्रयोग होने से रिकार्डिंग की तकनीक में भी परिवर्तन आया। ऐसे रिकार्डरों का आविष्कार हुआ जिनमें रिकार्डिंग के DAT (Digital Audio Tape) तकनीकी का प्रयोग होता या लेकिन वर्तमान में बहुत छोटे और प्रचालन में सहज रिकार्डर भी मौजूद हैं जिनमें एक छोटे चिप में कई घंटों का डेटा रिकार्ड करके सीधा कम्प्यूटर में स्थानांतरित किया जा सकता है। ये रिकार्डर ध्वनियों को वेब के रूप (Wave form) में रिकार्ड करते हैं जिससे ध्वनियों के सूक्ष्मतम स्तर तक जा करके भी उनका विश्लेषण किया जा सकता है। वर्तमान में कुछ मोबाइलों में भी ध्वनि प्रलेखन की सुविधा है लेकिन उनका प्रयोग ध्वनि विज्ञान में शोध के लिए बहुत उपयोगी साबित नहीं हुआ है। वर्तमान तकनीकी युग में किसी भी भाषा प्रलेखन के क्षेत्रकार्य में निम्न उपकरणों का होना अति आवश्यक है-

क. डिजीटल डेटा रिकार्डर
ख. लैपटॉप
ग. सॉफ्टवेयर
घ. कैमरा
ङ. पेपर, पेन, पेंसिल, रबर इत्यादि

टेप रिकार्डर के दिन लद गए क्योंकि उन्हें कंप्यूटर में प्रयोग करने के लिए डिजिटाइट्ज करना पड़ता है । इसमें डेटा क्षय होता परिणामस्वरूप डेटा की गुणवत्ता में गिरावट आती है। वर्तमान युग डिजिटल युग है और इसके अनेकानेक लाभ हैं इसलिए समय के साथ कदम मिलाते हुए डिजिटल डेटा रिकार्डर का प्रयोग किया जाना चाहिए।

वैसे तो बाजार में अनेकों प्रकार के डिजिटल डेटा रिकार्डर विभिन्न कीमतों में उपलब्ध हैं लेकिन उनको खरीदने से पहले उनके बारे में जानकारी किसी अनुभवी व्यक्ति या इण्टरनेट पर उपलब्ध सूचना के प्राप्त कर लेनी चाहिए। कुछ निम्न बातों पर भी गौर कर लेना चाहिए-

1. रिकार्डर छोटे एवं मजबूत होने चाहिए जिससे उनको क्षेत्रकार्य के दौरान सुरक्षित रखने एवं ढोने में असुविधा न हो।
2. कम से कम 4 GB तक का मेंमोरी कार्ड उसमें लगा हो।
3. रिकार्डर प्रयोक्ता अनुकूल (यूजर फ्रेंडली) होना चाहिए । प्रयोक्ता अनुकूल होने से तात्पर्य ऐच्छिक ध्वनि को बार-बार सुन सकने तथा ध्वनि के छोटे से छोटे खण्ड को जितनी बार चाहे उतनी बार आसानी से सुनने की सुविधा एवं रिकार्डर आसानी से पूरी दक्षता के साथ प्रचालन कर सके।
4. फ्रीक्वेन्सी के प्रति उसका अच्छा रिस्पान्स होना तात्पर्य यह है कि वह कितनी क्षमता के पिच (pitch) की ध्वनि को अच्छे ढंग से प्रलेखित कर सकता है।
5. बिजली जाने पर रिकॉर्डर बैटरी पर स्वतः ही चला जाय अन्यथा बिना रक्षित किया हुआ डेटा क्षय हो सकता है।

मैंने खासी, भीली, गहरी, स्पीती, हक्कीपिक्की इत्यादि भाषाओं में अपने क्षेत्रकार्य दौरान जूम एच 4 हैण्डी रिकार्डर (ZOOM H4 Handy Recorder) (चीन में बना हुआ) एवं इडिरॉल (Edirol R-O9 24 bit WAVE/MP 3 Linear PCM by Roland) रिकार्डर का प्रयोग किया और इनकी गुणवत्ता, प्रचालन में सहजता से काफी प्रभावित रहा। मैने नियंत्रित वातावरण (controlled environment) में डेटा रिकार्डिंग के लिए टसकैम एच डी-पी 2 पोर्टेबल स्टीरियो ऑडियो रिकार्डर (Tascam HD-P2 Portable Stereo Audio Recorder) जो कि जापान का बना हुआ है का उपयोग किया और इसकी गुणवत्ता ने मुझे काफी प्रभावित किया। इसकी गुणवत्ता के बावज़ूद क्षेत्रकार्य में दो कारणों से इसकी अनुशंसा मैं नहीं करूँगा- पहला, इसका आकार बड़ा होने से इसको क्षेत्रकार्य में लाने ले जाने में असुविधा होती है और दूसरी, इसकी सेटिंग्स एवं प्रचालन को सीखने में थोड़े समय की आवश्यकता होती है क्योंकि यह थोड़ी क्लिष्ट है। वैसे वर्तमान में सोनी एवं पैनासोनिक के रिकार्डरों की भी काफी प्रशंसा की जा रही है।

अच्छे रिकार्डर के साथ अच्छी बैटरियों का भी प्रयोग करना ना भूलें। अगर आप रिचार्जेबल बैटरियों का प्रयोग करना चाहते हैं तो अलकालाइन या लीथियम बैटरियों का प्रयोग करें। अलकालाइन बैटरियाँ अच्छे ब्राँड की अच्छी बैटरी खरीदें। सस्ती बैटरियों के प्रयोग से बचें क्योंकि एक तो ये कम चलती हैं और दूसरे इनके लीक होने के संयोग अधिक होते हैं जिससे आप का महँगा रिकार्डर खराब हो सकता है। लीथियम बैटरियाँ महँगी तो होती हैं लेकिन ये अधिक समय तक चलती हैं और अलकालाइन से हल्की होती हैं। आप रिचार्जेबल बैटरियों का भी प्रयोग कर सकते हैं जिनपर समग्र खर्च कम आता है लेकिन चलती भी कम ही हैं। वैसे रिचार्जेबल बैटरियों का प्रयोग करना हो तो नी-काड (Nickel-Cadmium) या नी-एम एच (Nickel-Metal Hydride) बैटरियों का प्रयोग करें।

रिकार्डर में प्रयोग के लिए मेंमोरी कार्ड की क्षमता के बारे में निर्णय लेने से पहले आप यह देख लें कि आपको कितने घंटे की और किस गुणवत्ता की रिकार्डिंग करनी है। सामान्यतया 16 बिट डेप्थ, 44.1 kHz सैंपलिंग रेट पर मोनो चैनल में रिकार्डिंग की सेटिंग पर आप 1 GB की मेंमोरी कार्ड में 3 घंटे की रिकार्डिंग कर

सकते हैं। लेकिन जब आपको 34 बिट डेप्थ, 44.1 kHz सैंपलिंग रेट पर स्टीरियो चैनल में रिकार्डिंग करना हो तो आप 1 GB की मेमोरी कार्ड में मात्र 45 मिनट की ही रिकार्डिंग कर सकते हैं।

कंप्यूटरों के प्रयोग से ज्ञान के सभी क्षेत्रों में क्रांतिकारी परिवर्तन आया। क्षेत्रकार्य में लैपटॉप साथ होने से रिकार्ड किए गए डेटा को कंप्यूटर में भेज कर सहेज लेते हैं जिससे रिकार्डर खराब होने या खो जाने पर भी पूर्व में रिकार्ड किया हुआ डेटा कंप्यूटर में सुरक्षित रहता है। क्षेत्रकार्य के दौरान खाली समय में हम कंप्यूटर के सहयोग से रिकार्ड किए गए डेटा का मेटाडेटा बना सकते हैं जिससे भविष्य इच्छित डेटा को ढूँढने में परेशानी नहीं होगी। लंबे समय तक घर से बाहर रहते हुए क्षेत्रकार्य के दौरान हमें घर की याद सताने लगती है या उबन महसूस होने लगती है लेकिन इस वातावरण में कभी-कभी हम कंप्यूटर के सहयोग से गाना सुनकर या सिनेमा देखकर हम अपनी उबन को मिटा सकते हैं या कम कर सकते हैं। हालाँकि यह कार्य कंप्यूटर का गौण कार्य है परंतु कभी-कभी परिस्थिति विशेष में गौण कार्य भी प्राथमिक लगने लगता है।

ऐसे सॉफ्टवेयरों के बिना जिनका भाषावैज्ञानिक अध्ययनों में उपयोग होता है, कंप्यूटर क्षेत्रकार्य के लिए अनुपयोगी होगा। रिकार्ड किए हुए डेटा के संपादन के लिए कूल इडिट प्रो (Cool Edit Pro) एक आवश्यक सॉफ्टवेयर जो सिर्फ विण्डोज वाले कंप्यूटरों पर कार्य करता है। वाक् ध्वनियों के विश्लेषण के लिए प्रात (PRAAT) सॉफ्टवेयर एक बेहतर सॉफ्टवेयर है। इसमें रिकार्डेड डेटा का संपादन कर सकते हैं, उसका लिप्याँकन (Transcription) कर सकते है, स्पेक्टोग्राफ देख सकते हैं, टोन एवं पिच का भी विश्लेषण इस सॉफ्टवेयर की मदद से कर सकते हैं। इसका नवीनतम संस्करण 1.4.8 नवंबर 2014 में आया है। भाषा की फोनोलॉजी बनाने में मदद के लिए फ़ोनोलॉजी असिस्टैंट (Phonology Assistant{ XE "Phonology Assistant" }) सॉफ्टवेयर काफी कारगर है। यह एक बार इनपुट दे देने पर ध्वनियों का वितरण, व्यंजन गुच्छ, फोनोटैक्टिक नियम, सिलैवल इत्यादि का विश्लेषण स्वतः ही कर देता है लेकिन उसके बाद भी हमें एक बार उसके देख लेना चाहिए। संकटग्रस्त भाषाओं के प्रलेखन में सिर्फ रिकार्डेड ऑडियो डेटा का ही महत्व नहीं होता अपितु हम उनके समुदाय की सामाजिक सांस्कृतिक संरचना, उनका सदियों से उनकी स्मृति में रचा-बसा ज्ञान, उनके धार्मिक क्रिया-कलाप, उनके सामाजिक धार्मिक अनुष्ठान, उनका जीवन-दर्शन इत्यादि का प्रलेखन फोटो एवं विडियो रिकार्डिंग के द्वारा भी करते हैं। भाषा वैज्ञानिकों के लिए फोटो एवं विडियो रिकार्डिंग का संपादन, विश्लेषण एवं संरक्षण के लिए उपयुक्त सॉफ्टवेयर इलैन (ELAN Linguistic Annotator) है। इसको नीदरलैण्ड के मैक्स प्लैंक फ़ॉर साइकोलिंगविस्टिक्सन ने मूक-बधिरों की भाषा (Sign languages) के विश्लेषण एवं व्याख्या को ध्यान में रखकर निर्माण किया है लेकिन इसका प्रयोग सामान्य भाषाओं के लिए प्रलेखित विडियो डेटा के संपादन एवं प्रलेखन के लिए भी भाषावैज्ञानिकों द्वारा खूब प्रयोग किया जा रहा है। अंत में हम कोश बनाने वाले सॉफ्टवेयर लेक्सिक प्रो (Lexique Pro) की चर्चा करेंगे। संकटग्रस्त भाषाओं के प्रलेखन पर कार्य करने वाले भाषा वैज्ञानिक कोश बनाने के लिए लेक्सिक प्रो (Lexique Pro) सॉफ्टवेयर का प्रयोग कर रहे हैं। इस सॉफ्टवेयर को सिल (SIL- Summer Institute of Linguistics) की वेवसाइट से मुफ्त में डाउनलोड किया जा सकता है। अभी हाल में ही प्रो. अन्विता अब्बी की संकटग्रस्त अंदमानी भाषा का बहुचर्चित कोश भी इसी सॉफ्टवेयर में बना है। भारतीय भाषा संस्थान में भी एस.पी.पी.एल. प्रोजेक्ट में कोश बनाने के लिए इसी सॉफ्टवेयर का प्रयोग किया जा रहा है। इसकी सबसे बड़ी विशेषता है कि यह सॉफ्टवेयर डेटाबेस आधारित है। यह टूलबाक्स एवं शूबाक्स (Toolbox and Shoebox) द्वारा .txt फाइल में दिए गए उत्पाद को आसानी से अपना डेटाबेस बना सकता है। सबसे अन्त में दो बातों का ध्यान अवश्य रखें। पहला; सभी सॉफ्टवेयरों में हो रहे सुधार एवं

उनके नए संस्करणों पर नज़र अवश्य रखें, एवं दूसरा किसी भी सॉफ्टवेयर में कार्य शुरू करने से पहले कम से कम दो बार उसके प्रयोक्ता नियमावली (User's Manual) को ज़रूर पढ़ें।

अन्त में हमें इतने उत्कृष्ट उपकरणों को साथ रखने के बावज़ूद पेपर, पेन, पेंसिल, रबर इत्यादि को साथ रखना नहीं भूलना चाहिए। कभ-कभी उपकरण छोटे मोटे कारणों से कार्य करना बंद कर देते है ऐसे समय में उन पर अपना समय बर्बाद करने से बेहतर होगा कि आप अपने विश्वसनीय साथी अपने हाथ का उपयोग कर पेपर पेंसिल के साथ प्रलेखन का कार्य प्रारंभ कर दें। वैसे भी आप को क्षेत्रकार्य के दौरान डेटा रिकार्डिंग से बचे समय का उपयोग मेटा डेटा बनाने एवं लिप्यांकन (transcription) के लिए करना चाहिए और लिप्याँकन करने के लिए सबसे उपयोगी उपकरण पेपर और पेंसिल ही हैं।

सूचना प्रदाता का चुनाव

चुनाव शब्द से यहाँ तात्पर्य अच्छे सूचना प्रदाता से डेटा रिकार्ड करने से है। यह 'अच्छा' अपने आप में कई गुण समेटे हुए है जिसकी चर्चा नीचे की जाएगी। सर्वप्रथम हम यहाँ यह कहना चाह रहे हैं कि संकटग्रस्त भाषाओं के सक्षम वक्ता क्या हमें इतनी मात्रा मिलेंगे की हमें चुनने की सहूलियत रहे या बिना चुने हुए सूचना प्रदाता से डेटा रिकार्ड करने से प्रलेखन की गुणवत्ता पर असर पड़ेगा? कितने भाषा वैज्ञानिकों ने सिर्फ एक या दो बचे हुए वक्ताओं को ही (भले ही मज़बूरी में) सूचना प्रदाता के रूप में लेकर उत्तम गुणवत्ता का प्रलेखन कार्य किया है। इसलिए भाषा की संकटग्रस्तता, भाषा के बोलने वालों की उपलब्धि, भाषा बोलने में सक्षमता एवं वक्ता पर प्रभावशाली भाषा का प्रभाव इत्यादि को ध्यान में रखकर ही सूचना प्रदाता के चुनाव के प्रश्न पर विचार करना चाहिए। यह कोई बहुत आसान कार्य नहीं है। इसमें बड़ी ही सावधानी की आवश्यकता पड़ती है अन्यथा ध्वनि की गुणवत्ता से समझौता करना पड़ेगा या कभी-कभी मशीन (कंप्यूटर) में प्रयोग करने के उपयुक्त डेटा संग्रहण नहीं हो पाता और अंततः सूचना प्रत्यानयन (information retrieval) में परेशानियाँ आती हैं। अतः क्षेत्रकार्य शुरु करने से पूर्व ही सूचना प्रदाता के चुनाव पर विचार कर लेना समझदारी का काम होगा।

जब आप संकटग्रस्त समुदाय, भाषा या अनजानजनजाति क्षेत्र या सरकार द्वारा रक्षित क्षेत्र में क्षेत्रकार्य के लिए जा रहे हैं तो सर्वप्रथम आपको वहाँ सक्षम क्षेत्रीय प्रशासनिक अधिकारी से अनुमति ले लेनी चाहिए। नहीं तो कम से कम उस क्षेत्र के सक्षम प्रशासनिक अधिकारी को सूचना प्रेषित कर देनी चाहिए। दूसरे, आपको वहाँ एक क्षेत्रीय सहयोगी की भी आवश्यकता होगी जो वहाँ की जलवायु, परिवेश, सामाजिक ढाँचे एवं उस जनजाति की सामान्य आदतों से भली-भाँति अवगत हो। आपका सहयोगी कम से कम दो भाषाएँ जानता हो, एक वह भाषा जो संकटग्रस्त भाषा समुदाय के लोग समझते हों और दूसरी वह भाषा जिसे आप समझते हों। यही आप और उस समुदाय के बीच संप्रेषण की कड़ी होगा जिसका आप प्रलेखन करने जा रहें। अगर कोई शिक्षक मिल जाय तो बहुत अच्छी बात है नहीं तो स्थानीय डाकिया भी आपकी मदद कर सकता है। पुजारी और पादरी भी उस कार्य के लिए उपयुक्त हो सकते हैं बशर्ते उन लोगों की आपके कार्य में रुचि हो और वे इसके लिए समय निकाल पाएँ।

संकटग्रस्त जनजातीय भाषाओं में कार्य करना थोड़ा मुश्किल होता है क्योंकि आज भी कुछ जनजातियाँ बाहरी लोगों को पसन्द नहीं करती। इसके कई कारण हो सकते हैं जैसे उनका मानना कि बाहर के लोगों से संपर्क होने से प्राकृतिक आपदाएँ एवं बीमारियाँ आती हैं। हाँ, लेकिन प्रयास करके वहाँ के ग्राम प्रमुख या जनजातीय समिति के प्रमुख से आग्रह करके प्रलेखन का कार्य किया जा सकता है। कभी-कभी सूचना प्रदाता के मन में यह धारणा घर कर जाती है कि अन्वेषणकर्त्ता उनका शोषण कर रहे हैं और अपने इस कार्य से पैसा बनाएँगे। ऐसी स्थिति का भी सामना करने के लिए अन्वेषक को तैयार होना चाहिए। इन परिस्थियों से निबटने के लिए किसी सिद्धांत की आवश्यकता नहीं होती वरन अन्वेषक को खुद अपनी सूझ-बूझ एवं व्यवहार कुशलता से इनका सामना करके प्रलेखन करना होता है।

सूचनाप्रदाता के चुनाव से पूर्व संकटग्रस्त भाषाओं में वक्ताओं के वर्गीकरण पर विचार करना समीचीन होगा। इस विषय पर क्षेत्रकार्य संबंधित साहित्यों में मुख्यतः सैमरिन (Samarin 1967), न्यूमैन एवं रैटलिफ (Newman and Ratliff 2001), क्रॉउले (Crowley 2007), बॉउअर्न (Bowern 2008) इत्यादि ने काम किया है। ग्रीनेवाल्ड एवं बर्ट (Grinevald and Bert 2011) ने अपने लेख वक्ता एवं समुदाय (Speakers and Communities) में संकटग्रस्त भाषाओं एवं इनके प्रलेखन को ध्यान में रखते हुए इनके वक्ताओं के वर्गीकरण पर वृहद् विवेचन प्रस्तुत किया है। ग्रीनेवाल्ड एवं बर्ट ने अपने वर्गीकरण में कुल सात प्रकार के वक्ताओं का वर्णन किया है जो निम्न प्रकार हैं-

1) **कुशल वक्ता** (Fluent Speakers): ऐसे वक्ताओं की भाषाविदों में भाषा प्रलेखन के लिए काफी माँग है। ऐसे वक्ताओं को रूढ़ वक्ता (Traditional Speaker) भी कहा जाता है क्योंकि इनकी भाषा बहुत ही परंपरागत होती है। भाषा की शुद्धता और परंपरागत भाषा के संरक्षण को लेकर दूसरे वक्ताओं की अपेक्षा काफी सतर्क रहते हैं।
2) **अर्धकुशल वक्ता** (Semi Speakers): संकटग्रस्त भाषा समुदाय में ऐसे इस वर्ग के वक्ताओं की भरमार होती है। ऐसे वक्ताओं में ग्रहण कौशल (receptive skills) तो पर्याप्त होता है लेकिन विभिन्न वक्ताओं के उत्पादन कौशल (productive skills) में भिन्नता होती है। इनकी भाषा में थोड़ा परिवर्तन रहता है जिसको समुदाय का कुशल वक्ता अर्धकुशल वक्ता की गलती समझता है। अर्धकुशल वक्ता अपने अधिकांश सामाजिक जीवन का निर्वाह संकटग्रस्त भाषा की अपेक्षा प्रभावशाली भाषाओं में करता है।
3) **अंशतः कुशल / टर्मिनल वक्ता** (Terminal Speakers): अंशतः कुशल या टर्मिनल वक्ता वे हैं जिनका ग्रहण कौशल अर्धकुशल वक्ता से काफी कम होता है और उत्पादन कौशल तो बहुत ही कम होता है। कभी-कभी इनका संकटग्रस्त भाषा ज्ञान बहुत प्रचलित या रूढ़ वाक्यों या शब्दों के प्रयोग तक ही सीमित होता है।
4) **स्मरणशील वक्ता** (Rememberes): कभी-कभी किसी भाषा समुदाय के लोगों के साथ ऐसी घटनाएँ घट जाती हैं कि उनके दबाव के कारण उन्हें अपनी भाषा या अपनी पहचान छुपानी पड़ती है (जैसे नरसंहार के पश्चात पलायन)। क्रमशः वे समय बीतने के क्रम में अपनी भाषा को भूल जाते हैं और सिर्फ प्रभावशाली भाषा का ही व्यवहार करते हैं। ऐसे वक्ताओं को विश्वास में लेकर उनकी भाषा को याद करवाने का प्रयत्न किया जाय तो वे उस भाषा को याद कर के उसके प्रयोग की सीमित दक्षता हासिल कर सकते हैं। ऐसे ही वक्ताओं को स्मरणशील वक्ता कहा जाता है।
5) **आभासी वक्ता** (Ghost Speakers): आभासी वक्ता संकटग्रस्त भाषाओं के ऐसे वक्ता हैं जो कुछ हद तक तो भाषा जानते है लेकिन किसी के सामने वे स्वीकार नहीं करते कि वे उस भाषा को थोड़ा भी जानते हैं। अपनी भाषा की जानकारी को नकारने का मुख्य कारण उस भाषा के प्रति सशक्त नकारात्मक दृष्टिकोण या दूसरों की निगाह में अपनी पहचान न ज़ाहिर करना होता है। ऐसा उन स्थितियों में भी होता है जब लोग प्रभावशाली भाषाओं की तुलना में क्षेत्रीय बोलियों को हेय दृष्टि से देखते हैं।
6) **नव वक्ता** (Neo Speakers): वास्तव में भाषा पुनरुत्थान (language revitalization) का मुख्य उद्देश्य होता है इस प्रकार के वक्ताओं को तैयार करना जो उस भाषा को पहले से नहीं जानते अपितु सीखकर उस भाषा को कुछ सीमा तक प्रयोग करना सीख लेते हैं। इनकी भाषा दक्षता भाषा पुनरुत्थान में पढ़ाई की गुणवत्ता एवं समय, व्यक्ति की भाषा को सीखने की लगन इत्यादि पर निर्भर करती है। इनके द्वारा हासिल भाषा दक्षता के आधार पर इनको पुनः अर्धकुशल वक्ता की श्रेणी में वर्गीकृत किया जा सकता है।

7) **अंतिम वक्ता** (Last Speakers): हालाँकि इस प्रकार की श्रेणी का अस्तित्व संकटग्रस्त भाषाओं के वक्ताओं के वर्गीकरण में नहीं है लेकिन इन वक्ताओं की काफी चर्चा जनसंचार के माध्यमों एवं संकटग्रस्त भाषाओं के प्रलेखन के संदर्भ में होती है। ऐसा प्रतीत होता है कि यह श्रेणी व्यक्ति विशेष को समुदाय के लोगों द्वारा प्रदान की जाती है या कभी-कभी वक्ता अपने को अंतिम वक्ता घोषित कर देता है। इनका क़द समुदाय में काफी ऊँचा होता है और इस प्रकार ये समुदाय के किसी अन्य सदस्य द्वारा बोली गई भाषा या भाषावैज्ञानिकों द्वारा पहचाने गए भाषा के किसी उच्चारण या वाक्य को आसानी से ग़लत घोषित कर देते है। यह आवश्यक नहीं कि अंतिम वक्ता उसा भाषा का कुशल या अर्धकुशल वक्ता ही हो। कभी-कभी अंतिम वक्ता भी उस भाषा का टर्मिनल वक्ता ही होता है।

सूचना प्रदाता के चुनाव में हमें निम्न बातों का ध्यान रखना चाहिए –

1. पहले यह सुनिश्चित कर लेना चाहिए कि जिस संकटग्रस्त भाषा का प्रलेखन हम कर रहे हैं वह भाषा सूचना प्रदाता की मातृभाषा हो न की द्वितीय भाषा।
2. मातृभाषा के साथ-साथ सूचना प्रदाता इस भाषा का प्रयोग अपनी रोजमर्रा की जिन्दगी में करता हो।
3. सर्वप्रथम हमें संकटग्रस्त भाषा के प्रलेखन के लिए उस भाषा के कुशल, अर्धकुशल वक्ता को सूचनाप्रदाता के रूप में ढूँढने का प्रयत्न करना चाहिए। लेकिन आपको ऐसा वक्ता न मिले तो आपको कम से कम टर्मिनल वक्ता तो चाहिए ही।
4. भाषा की ध्वनि व्यवस्था (फ़ोनोलॉज़ी) का प्रलेखन करने के लिए ऐसे सूचना प्रदाता की आवश्यकता होती है जिनके मुँह में पूरे दाँत हो, दाँत अखंड हों, उनके बीच रिक्त जगह न हो । इसके अलावा उसके होंठ कटे न हों, नाक से न बोलता हो, उसे सुनने में कोई समस्या न हो अन्यथा वह हमेशा उच्च आवाज (High pitch) का प्रयोग कर सकता है जिससे रिकार्डेड डेटा प्राकृतिक न होकर कृत्रिम लगने लगेगा। हमें यह भी देख लेना चाहिए की सूचना प्रदाता हकलाता, तुतलाता न हो।
5. डेटा प्रलेखन के लिए प्रयास करके ऐसे वक्ता का चुनाव करना चाहिए जिसके बारे में उस भाषा के बोलने वाले लोग भी उसे एक सक्षम और अच्छा वक्ता मानते हो। खासकर संकटग्रस्त भाषाओं के संदर्भ में इसका ध्यान रखना आवश्यक हो जाता है क्योंकि उस भाषा की मातृभाषा तो उस भाषा समुदाय के सभी लोग बोल लेते हैं लेकिन वास्तव में उस भाषा को सक्षम रूप से एवं धारा प्रवाह बोलने वालों की संख्या कम ही होती है। यह भी देखने में आया है कि समुदाय के सबसे बुजुर्ग व्यक्ति को समुदाय के लोग उस भाषा का अच्छा जानकार या वक्ता बताते हैं लेकिन ध्वनि के स्तर पर डेटा के लिए बहुत बुजुर्ग आदमी उपयुक्त नहीं होगा क्योंकि हो सकता है कि उसके आगे के कुछ दाँत टूटे हों या उसकी जबान लड़खड़ाती हो या कभी-कभी एक निश्चित उम्र के बाद लोगों की आवाज में एक कम्पन्न सी होने लगती है जो स्वरों एवं व्यंजनों की गुणवत्ता पर असर डाल सकता है। लेकिन जब आप संकटग्रस्त भाषाओं के सामाजिक, सांस्कृतिक एव धार्मिक विचारों का प्रलेखन कर रहे हों, उसके ग्राम्य साहित्य का अध्ययन कर रहे हों या संकटग्रस्त भाषा का कोश बना रहे हों तो बेशक एक वृद्ध व्यक्ति एक अच्छा सूचना प्रदाता बन सकता है और आपको अच्छी जानकारी उपलब्ध करा सकता है।
6. कुछ वक्ताओं की आदत होती है प्रत्येक वाक्य के अंत में कुछ विशिष्ट शब्द बोलने की, जिसको आम बोलचाल की भाषा में ठेका या तक़ियाकलाम कहा जाता है। ऐसे सूचना प्रदाताओं से भी बचना चाहिए क्योंकि सेगमेंटेशन, एनोटेशन एवं विश्लेषण के दौरान ऐसे विशिष्ट शब्द काफी समस्याएँ खड़ी करते हैं। इनको संपादित कर हटाने पर एक तो नैतिक मुद्दे खड़े होते हैं और दूसरे संपादित होने पर डेटा पर ही प्रश्न चिह्न खड़ा कर देते हैं। एस.पी.पी.एल. परियोजना के अंतर्गत प्रलेखित संकटग्रस्त भाषा गोजापूरी (ग्राम- खालू, जिला- डोडा, जम्मू एवं कश्मीर) के दिसम्बर 2014 में भेजी गई रिपोर्ट के मूल्याँकन करने के क्रम मुझे एक ऐसा उदाहरण मिला जिसमें सूचनाप्रदाता हर

शब्द या वाक्य के अंत में 'है जी' शब्द का प्रयोग करता था। जैसे रेत के लिए 'रेत है जी रेत', ज़मीन के लिए 'ज़मीन है जी ज़मीन' जिसको सेगमेंट करना बहुत ही कठिन कार्य होगा।

7. हमें ऐसे लोगों को भी सूचना प्रदाता बनाने से बचना चाहिए जो अपनी भाषा को विशुद्ध रूप देने के क्रम में आवश्यकता से अधिक शुद्ध उच्चारण करके अतिशोधन (hyper correction) करने लगते हैं। कुछ भोजपुरी बोलने वालों को मैने भारतीय भाषासंस्थान में ही जो अपने को भोजपुरी भाषी मानने में झिझक महसूस करते हैं वे CIIL (सी.आई.आई.एल.) को अतिशोधन से शी.आई.आई.एल. का भी प्रयोग कर देते हैं। हमें इस प्रकार के सूचना प्रदाता से भी बचना चाहिए।
8. कभी-कभी उस संकटग्रस्त भाषा समुदाय के प्रमुख या सूचना प्रदाता के परिवार वालों को छोटा-मोटा उपहार भी देना चाहिए जिससे संबंधों में नवीनता बनी रहेगी और समुदाय आपको अपना सदस्य समझने लगे।
9. एक सामान्य प्रश्न पर और विचार कर लेना चाहिए कि किसी भाषा के प्रलेखन के लिए एक भाषा से कम से कम कितने सूचना प्रदाताओं से डेटा लेना चाहिए? वास्तव में इस प्रश्न का उत्तर आपकी परियोजना के उद्देश्य, प्रकृति एवं डेटा के उपयोग पर निर्भर करता है। वैसे सामान्यतः एक संकटग्रस्त भाष के प्रलेखन के लिए एक स्थान से कम से कम 6 सूचना प्रदाता से डेटा लेना आदर्श माना जाता है। इससे अधिक हों तो अच्छा ही है लेकिन कभी-कभी 6 सक्षम एवं अच्छा सूचना प्रदाता संकटग्रस्त भाषाओं में मिलना मुश्किल हो जाता है। कितने भाषाविदों ने सिर्फ एक सूचना प्रदाता के डेटा से भाषा-विशेष की ध्वनि व्यवस्था का अच्छा निरूपण किया है। सूचना प्रदाता का विभाजन इस प्रकार होना चाहिए–

उम्र	पुरुष	महिला
16-20	1	1
21-50	1	1
50- से ऊपर	1	1

अगर संभव हो तो इस प्रकार के कुछ और मापदण्डों में से भी इस प्रकार सूचना प्रदाता चुनना चाहिए जैसे- शिक्षित-अशिक्षित, शहरी-ग्रामीण, गरीब-अमीर इत्यादि।

10. महिला एवं पुरुष दोनों समूहों से डेटा इसलिए लेना चाहिए कि महिला एवं पुरुषों की भाषा में शब्दों के स्तर पर उच्चारण के स्तर पर एवं स्वनों के स्तर पर भी कुछ भाषाओं में भिन्नता मिलने की संभावना रहती है। पीटर लैडफोगेड (2003:14) ने अमेजन के वर्षा वनों में बोली जाने वाली भाषा पिराहाँ (Pirahã) का उदाहरण देते हुए बताया है कि जिस परिस्थिति में महिलाएँ /ह्/ ध्वनि का प्रयोग करती हैं उसी परिस्थिति में पुरूष /स्/ ध्वनि का प्रयोग करते हैं। अतः महिला एवं पुरूष दोनों से डेटा लेना अनिवार्य होना चाहिए।

निष्कर्षतः सूचनाप्रदाता के चुनाव में हमें काफी सतर्कता बरतनी चाहिए। इसके लिए अपनी सोच के साथ-साथ अपने सहयोगियों (अगर कोई साथ में हो तो) और भाषा समुदाय के सदस्यों की भी राय लेनी चाहिए। सूचना प्रदाता एक कुशल या अर्धकुशल वक्ता से साथ-साथ स्वभाव में सहज और प्रलेखन के लिए स्वेच्छा से सुलभ हो।

सूचना का प्रलेखन

सूचना प्रदाता का चुनाव कर लेने के पश्चात क्षेत्रकार्य का जो दूसरा प्रमुख कार्य होता है वह है भाषा के प्रलेखन (Recording) का । भाषा प्रलेखन के लिए बहुत तरह के उपकरण प्रचलन में हैं जिनकी चर्चा

ऊपर उपकरण एवं तकनीकी का चुनाव शीर्षक के अंतर्गत की जा चुकी है। भाषा की ध्वनि रिकार्डिंग के समय निम्न बातों का ख्याल रखें-

1. रिकार्डर की सेटिंग्स ठीक रखें। अच्छी रिकार्डिंग के लिए सेटिंग्स में बिट डेप्थ 34 तथा सैंपलिंग रेट 44.1 kHz रखें। मोनो या स्टीरियो (mono or stereo) चैनल अपनी आवश्यकता के अनुसार सेट करें। अगर आपका सूचना प्रदाता सिर्फ एक व्यक्ति है तो मोनो चैनल में रिकार्डिंग करना अच्छा होगा लेकिन अगर सूचना प्रदाता एक से अधिक हैं तो स्टीरियो चैनल में रिकार्डिंग करना ही अच्छा होगा। अगर आप संकटग्रस्त भाषा का ग्राम्यसंगीत का प्रलेखन कर रहें हैं तो स्टीरियो चैनल में रिकार्डिंग करना ही अच्छा होगा।
2. रिकार्डर शुरू करने से पुर्व रिकार्डर की सेटिंग्स, बैटरी की स्थिति, एवं वहाँ का वातावरण को खूब बारीकी से देख लें, अच्छी रिकार्डिंग के लिए खुले स्थान की अपेक्षा बंद कमरा ज्यादा उपयुक्त होगा।
3. शोधकर्ता को वास्तविक रिकार्ड करने वाली ध्वनि और शोर (background noise) के बीच का अन्तर पहचानना होगा। ध्वनि और शोर के बीच के अन्तर के अनुपात को डी. बी (dB) कहा जाता है। सामान्यतया 1 डी. बी का अन्तर ध्वनि और शोर में बहुत ज्यादा फर्क नहीं डालता लेकिन इससे ऊपर का अन्तर समस्या खड़ी कर सकता है (Ladefoged 2003:18–20) । इसलिए वास्तविक रिकार्डिंग प्रारम्भ करने से पहले थोड़ी रिकार्डिंग करके उसके ध्वनि और शोर के बीच का अनुपात देख लेना चाहिए और उसके अनुसार अपने रिकार्डर की सेटिंग ठीक करके ही वास्तविक रिकार्डिंग प्रारम्भ करनी चाहिए। हालाँकि आजकल कुछ सॉफ्टवेयर ऐसे हैं जिनके उपयोग से शोर को कम या बिल्कुल खत्म किया जा सकता है (noise reduction) लेकिन उससे आपके जरूरत की ध्वनियों की भी प्रबलता (loudness) या गुणवत्ता कम हो जाएगी । इसलिए हमें अधिकतम सम्भव प्रबलता पर ही ध्वनियों को रिकार्ड करना चाहिए लेकिन अगर आप रिकार्डिंग नियंत्रित वातावरण (controlled environment) में कर रहे हैं तो रिकार्डर की सेटिंग्स में प्रबलता को ज्यादा बढ़ाना नहीं चाहिए। तकनीकी रूप से एक अच्छी रिकार्डिंग का मतलब होता है शोर रहित ध्वनि रिकार्डिंग।
4. रिकार्डिंग करते समय माइक्रोफोन को हमेंशा मुँह के किनारे पर तकरीबन 3 सेंटीमीटर की दूरी पर रखें। किनारे पर रखने के पीछे का तर्क यह है कि अगर कभी अचानक सूचनाप्रदाता सामान्य से अधिक उँची आवाज में बात करे तो यह अवस्था मुहँ से अचानक तेज निकलने वाले हवा के झोंके के प्रभाव को कम कर रिकार्डिंग की गुणवत्ता को खराब नहीं होने देगा।

लेखक हिमाचल के केलाँग में सूचना प्रदाता श्री दोर्जे मने पा से संकटग्रस्त भाषा गहरी की रिकार्डिंग करता हुआ

5. आपको को यह भी प्रयास करना चाहिए कि अनावश्यक ध्वनियों (background noise) जैसे- हवा की तेज आवाज, मोटर-गाड़ियों की आवाजें, पशु-पक्षियों की आवाजें, खिड़कियों-दरवाजों के खुलने-बंद होने की आवाजें इत्यादि न हो अन्यथा आपकी रिकार्डिंग की गुणवत्ता पर असर पड़ सकता है।
6. वर्तमान समय में विज्ञान हर घर तक किसी न किसी रूप में पहुँच चुका है। ऐसे दौर में विद्युत उपकरणों का हर घर में बोलबाला है। लेकिन ये उपकरण जैसे- टी.वी, पंखे, कूलर की आवाज एवं ट्यूबलाइट की झनझनाहट की आवाज आपकी रिकार्डिंग की गुणवत्ता पर असर डाल सकते हैं।
7. मोबाइल आज सबकी जरूरत है, इसके बिना जिंदगी अधूरी लगती है लेकिन रिकार्डिंग करते समय मोबाइल की घंटी रिकार्डिंग को उतने समय तक के लिए खराब कर सकती है। कुछ शोधकर्ता मोबाइल को वाइब्रेशन मोड में रख कर यह सोच लेते हैं कि अब यह रिकार्डिंग को प्रभावित नहीं करेगा लेकिन यह सच नहीं है। वाइब्रेशन में भी एक ध्वनि निकलती है जो रिकार्डिंग को प्रभावित करती है। अगर आप हेड फोन लगाकर रिकार्डिंग कर रहों तो आप घंटी बजने से पूर्व ही महसूस कर सकते हैं कि किसी के मोबाइल पर कॉल आ रही है।

संदर्भ ग्रंथ-

Bowern, Claire (2008). *Linguistic Fieldwork: A Practical Guide*. New York: Palgrave Macmillan.

Cameron, D. et al. (1992). *Researching language: issues of power and method*. New York: Routledge.

Crowley, Terry (2007). *Field Linguistics: A Beginner's Guide*. Oxford University Press.

Daes, Erica-Irene (1993). *Discrimination against Indigenous peoples: Study on the protection of the cultural and intellectual property of indigenous peoples*. Paper presented to the 45th session of the Commission on Human Rights, Economic and Social Council, United Nations, New York.

Dixon, R.M.W. (2010). *Basic Linguistic Theory: Volume 1-Methodology*. Oxford: Oxford University Press.

Dwyer, Arienne M. (2006). *Ethics and practicalities of cooperative fieldwork and analysis*. In Gippert, Jost, Mosel, Ulrike and Nicolaus Himmelmann, (eds). Fundamentals of Language Documentation: A Handbook (pp. 31-66). Berlin: Mouton de Gruyter.

Grinevald, Colette and Michel Bert (2011). *Speakers and communities*. In The Cambridge handbook of Endangered Languages (45-65). Eds. By Austin, Peter K and Jullia Sallabank. Cambridge: Cambridge University Press.

Ladefoged, Peter (2003). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.

Newman, Paul and Martha Ratliff (eds.) (2001). *Linguistic Fieldwork*. Cambridge University Press.

Samarin, William J. (1967). *Field Linguistics: A Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston.

शोध दृष्टि

क्रियोल भाषा या संस्कृति (मॉरिशियन परिप्रेक्ष्य में)

सविता तिवारी

कोई भी भाषा किसी अन्य भाषा से मिलती जुलती हो सकती है, पर उसे उस भाषा का बिगड़ा स्वरूप कहना उचित नहीं है। जिस प्रकार भोजपुरी, राजस्थानी, गुजराती, पंजाबी, अवधी, आदि भाषाएँ हिंदी के करीब होते हुए भी अपने आप में पूर्ण रूप से स्वतंत्र भाषा एवं संस्कृति को दर्शाती है, उसी प्रकार मॉरिशस की क्रियोल भाषा फ्रेंच के करीब होते हुए भी उससे पूरी तरह भिन्न है। यह भाषा गुलामी और बंधुआ मजदूरों के इतिहास की द्योतक है। इस भाषा में छिपे दर्द का अंदाजा इसी बात से लगाया जा सकता है कि घर को क्रियोली में "काज" कहा जाता है काज शब्द इंग्लिश के केज शब्द से बना है, जिसका अर्थ है- 'पिंजरा' गुलामों को पिंजरों में जानवरों की तरह बंद रखा जाता था और वही उनका घर बना। उन्हीं पिंजरों में किन परिस्थितियों में क्रियोली भाषा का जन्म एवं विकास हुआ यह अपने आप में इतिहास का विषय है। दोस्तों को आज भी "कमाराद" की संज्ञा प्राप्त है, जिसका अर्थ है कोठरी का साथी। गुलामी की कोठरियों में बच्चों और रिश्तों का जन्म होता था। किसी बात को पूछने के लिए आज भी "कोमो" कहते हैं, जो अंग्रेजी के शब्द कमांड से बना है। गुलामों को केवल आदेश लेने के लिए जबान खोलने की इजाजत थी।

मॉरिशस में 90 प्रतिशत लोगों की बोलचाल की भाषा क्रियोली है। ऐसी विशेषताओं के बावजूद इस भाषा का भविष्य संकट में है। इन्फीरियोरिटी कॉम्प्लेक्स ने इस भाषा को जकड़ लिया है। जो लोग आपस में आराम से क्रियोली में बात कर सकतें है, उन्हें टूटी-फूटी फ्रेंच से जूझते देखना यहां एक आम नजारा होता जा रहा है, खास कर युवाओं में यह प्रवृत्ति अधिक देखी जाती है। 12 लाख की आबादी वाले इस छोटे से टापू में किसी भी चीज को प्रचलित होने या लुप्त होने में समय नही लगता। इस वर्ष फ्रांसीसियों के आगमन की 300वीं वर्षगांठ मना रहा, यह टापू पूरी तरह फ्रेंच भाषा के रंग में रंगा नजर आ रहा है।

क्रियोल क्या है -

दो या अधिक संस्कृति या रीतियों के मेल को मानवविज्ञान में क्रियोलीकरण कहा जाता है। वैश्वीकरण के दौर में भाषाओं का क्रियोलीकरण पूरे विश्व में देखा जा सकता है। मुख्य तौर पर अंग्रेजों की प्लांटेशन कॉलोनियों में इस परंपरा एवं भाषा का जन्म हुआ। लुसियाना में क्रियोल शब्द का इस्तेमाल एक विशेष प्रकार के आर्कीटेक्ट के लिए किया जाता है। रीयूनियन आईलैंड में लोकल व्यक्तियों के लिए इस शब्द का इस्तेमाल होता है,मॉरिशस की बात करें तो यहां क्रियोल भाषा है और अफ्रीकी मूल के व्यक्ति के लिए क्रियोल शब्द का उपयोग होता है। भाषा के रूप में क्रियोल मॉरिशस के अलावा जमाइका, त्रिनिडाड, रियूनियन, लुसियाना के अलावा अन्य कई अफ्रीकी देशों में प्रयोग में लायी जाती है।

क्रियोलीकृत भाषा की उत्पत्ति-

मॉरिशस एक प्राकृतिक टापू नही है इसलिए यहां अपनी कोई आबादी नही है, जो भी लोग हैं, वह सभी बाहर से आए हैं। चौदहवीं शताब्दी अरब लोग यहाँ पर आए, लेकिन यहां पर बसे नहीं। उसके बाद डचों ने इस टापू को आबाद किया। अठारहवीं शताब्दी की शुरुआत में फ्रांसीसियों ने इस टापू को डचों से छीन लिया। फ्रांसीसियों ने यहां गन्ने की खेती शुरू की, जिसके लिए अफ्रीका एवं मेडागास्कर से गुलाम मजदूरों को यहाँ पर लाया गया। यहीं से क्रियोल भाषा के इतिहास की शुरुआत हुई। फ्रांसीसी मालिक से बात करने के लिए जल्द ही गुलामों ने अपनी और मालिकों के समझ की आसान व्याकरण वाली भाषा का विकास किया जो आगे जाकर क्रियोल कहलाई। 1810 में गुलाम प्रथा का बहिष्कार हुआ साथ ही अंग्रेजों ने मॉरिशस को फ्रांसीसियों से जीत लिया। अब 1840 से मुक्त किए गए गुलामों को भारत से लाए मजदूरों से प्रतिस्थापित किया गया। भारत के विभिन्न भाषाई प्रांतों बिहार, कलकत्ता, तमिलनाडु, महाराष्ट्र, उत्तरप्रदेश से बंधुआ मजदूरों को अंग्रेज़ यहाँ पर लेकर आए। अब इन मजदूरों के पास भी कॉमन भाषा का अभाव था, यही कारण था कि इन्होंने भी जल्द ही क्रियोल भाषा को अपना लिया।

वर्तमान स्थिति-

90 प्रतिशत मॉरिशस वासियों की आम बोलचाल की भाषा क्रियोली है, बावजूद इसके यहां की कामकाजी भाषा अंग्रेजी और फ्रेंच है। क्रियोल को राजकीय भाषा बनाने के कई प्रयास विफल हुए हैं। 1982 में सरकार ने क्रियोल को मीडिया एवं स्कूलों में लागू करने का भी प्रयास किया जिसका आम जनता में बहुत विरोध हुआ और ऐसा हो न पाया। 2013 में मॉरिशस के स्कूलों में क्रियोल को वैकल्पिक भाषा के रूप में प्रस्तुत किया गया, जिसमें नामांकन ना के बराबर हैं।

संकट का कारण-

आज भी इसे अफ्रीकी मूल के लोंगों की भाषा समझा जाता है। इसे केवल बोलने की भाषा का दर्जा प्राप्त है, न कि लिखने का। यही कारण है कि इस भाषा का मानकीकरण नहीं हुआ है। एक ही शब्द को भिन्न-भिन्न तरीकों से बोला एवं लिखा जाता है। इसे फ्रेंच के बिगड़े रूप के तौर पर देखा जाता है। माता- पिता सोचते है कि बच्चों को क्रियोली सिखाने से उनकी फ्रेंच बिगड़ जाएगी जो एक अंतरराष्ट्रीय भाषा है। क्रियोली भाषा ना तो रोजगार से जुड़ी है और ना ही इसमें साहित्य रचना होती है ना ही मीडिया में इसकी जगह है।

निष्कर्ष-

क्रियोली भाषा पर संकट की जड़ें इसके जन्म में ही निहित हैं। निर्मूलन की आबादी से बसे इस टापू का हर व्यक्ति आज भी मानसिक रूप से अपनी जडों से जुड़ा है। चाहे वह अंग्रेज हो या फ्रांसीसी, भारतीय हो या चीनी सब अपनी पहचान अपनी भाषा में खोजते हैं। यहां का हर व्यक्ति सेगा संगीत पर नाचते हुए अपनी सुध-बुध खो देता है, पर मंच पर गाने की बारी आए तो वह अपने समुदाय के गीतों को ही चुनता है। यह बात सामुदायिक तौर पर अच्छी लगती है पर जब देश की बात आती है तो एकता वाली भाषा क्रियोली के साथ सौतेला व्यवहार लगता है। पूरे देश की अभिव्यक्ति की भाषा क्रियोली है पर मातृभाषा केवल 10 प्रतिशत लोगों की है, अब फ्रेंच के बढ़ते वर्चस्व के कारण यह आंकड़ा भविष्य में और भी कम होता नजर आ रहा है।

Language Maintenance and Shift: A Study of Patterns of Language Use among the Sirajis of Doda District of Jammu and Kashmir

Sabba Mushtaq
PhD Scholar
Department of Linguistics
University of Kashmir
Srinagar

Javaid Aziz Bhat
PhD Scholar
Department of Linguistics
University of Kashmir
Srinagar

Introduction

Language Maintenance is a situation where there is a continued use, preservation and support for a language or the traditional form of a language. Language Shift on the other hand is a gradual or sudden move from the use of one language to another and the speakers appear to have made a choice in favour of the dominant language. Babito (2005) defined language maintenance as a situation in which a language maintains its vitality even under pressure. It implies that degree of resistance is strong enough to contain any pressures that may be coming from the dominant language. Fishman (1972) points out that the basic pre-requisite of language maintenance and shift is a contact situation. The contact situation may give rise to either bilingualism or language shift.

Fasold (1989) noted that the concept of domain was first proposed by Fishman as a way of looking at language choices. According to Fishman, domains are institutional contexts in which one language variety is more likely to be appropriate than another. Domains are to be taken as constellations of factors such as location, topic and participants. 'Topic' refers to the subject one is talking about. 'Location' refers to the place where the speaker talks and the 'Participants' refers to the persons with whom one is talking. By looking at language use in different domains one can get an idea whether there are any particular linguistic preferences for different domains and different interlocutors. Then what remains to be seen is whether this multilingual community is stable or unstable(Fishman,1972).In the unstable case the gradual relocation of different languages to different domains occurs so radically that over a time people may give up one language in favour of the other. In this backdrop the present paper aims to look into the patterns of language use of Siraji community of Doda district of Jammu and Kashmir.

Siraji has been classified as a dialect of Kashmiri (Grierson,1919) spoken in Doda district of Jammu and Kashmir. In the 2001 census, Siraji has been reported to be spoken by around 87,179 speakers. The word 'Siraj' means 'The Kingdom of Shiva' and hence any 'wild mountainous country' and the speech variety became Siraji.Grierson(1919) regarded Siraji

as a mixed language because it contains borrowings from other languages with which it is in contact like Bhaderwahi, Kishtwari, Poguli and Rambani. Siraji has two major dialects namely Siraji of Doda and Siraji of Ramban. The differences in the two dialects have developed during the course of time due to different language contact environments. Both the Muslim and the Hindu population of the area speak Siraji. The linguistic repertoire of the Sirajis of Doda district of Jammu and Kashmir is Siraji, Kashmiri, Urdu and English.

Methodology

This paper deals with language use in different domains of Sirajis of Doda district of Jammu and Kashmir and its relevance in understanding sociolinguistic setup of that community. The main purpose of the elicitation of this data was to observe and understand if there is a shift away from Siraji(S)

a) Across different domains
b) Across the interlocutors in each domain
c) From the informal to the formal contexts

The study is based on the analysis of speech of 81 Muslim Siraji Respondents from Doda district. All the respondents are native speakers of Siraji. The study of the language use of Sirajis was carried out by means of a sociolinguistic questionnaire. The criteria for selecting the sub groups were as follows:

a. **Age :** The respondents were divided into 3 different age groups. In the table below (O) stands for Old, (M) stands for Middle-aged and (Y) stands for Young.

Table 1: Distribution of respondents across age groups

Group	**Age Range**	**Number of respondents**
Old (O)	51 and above	**27**
Middle-aged (M)	26-50 years	**27**
Young (Y)	Up to 25 years	**27**

b. **Gender:** These are the male and female sub groups. The male/female subdivision in each group is shown below where old males are denoted as OM, Old females as OF, Middle-aged Males as MM, Middle aged females as MF, Young Males as YM and Young females as YF.

O( )		**M()**		**Y()**	
Old Males (OM)	Old Females (OF)	Middle-aged Males (MM)	Middle-aged Females (MF)	Young Males (YM)	Young Females (YF)
13	**14**	**13**	**14**	**13**	**14**

Table 2: Breakup of the three age groups of the respondents on the basis of gender

Language use within the Home domain

Language use in the Home domain is investigated in almost every research concerning maintenance and vitality. Home is the domain where the interaction is mainly with the members of the family and it is expected that mother tongue will be the language of home. Home domain is an important domain for communication changes in languages taking place in this domain may reflect changing happening elsewhere. Fishman (1972) has taken into account two different approaches (one of Braunshausen and Mackey, and the other of Gross) in studying language use patterns in the home domain. Braunshausen and Mackey (1962) has specified specified-father, mother, child, domestic help etc as family members while as Gross (1951) specified it as dyads: grandfather to grandmother etc i.e. language of interaction between speaker and hearer within the home domain. In this study the respondents returned their use of language with 5 role relationships with in the family such as grandfather, grandmother, parents, spouse and children.

Table 3 shows the summary of language use in the Home domain.

Q#	Question	n=	Siraji	%	Kashmiri	%	Urdu	%	English	%
1a	With Grand Father	52	50	96%			2	4%		
1b	With Grand Mother	62	61	98%			1	2%		
1c	With Parents	76	70	92%			6	8%		
1d	With Spouse	40	36	90%			4	10%		
1e	With Children	50	42	84%			8	16%		
Mean		280	259	93%			21	7%		

Table 3: Language use in the Home Domain

As we see in Table 3 on an average over 90% of the Subjects reported that they speak Siraji in the Home domain. The % drops when the children are the interlocutors. The reason for this shift is that some Sirajis feel that learning other languages will fetch them bright future and better job opportunities in life. It can be see that 96% of the total subjects speak Siraji with their grandfathers, 98% with their grandmothers, 92% with parents, 90% with their spouses and 84% with their children.8% of the total respondents use Urdu in this domain.

It is quite evident from Table 3 that Sirajis do not use Kashmiri or English in Home domain. When enquired by the researcher the common response was that Kashmiri and English can be learnt from other sources such as interaction with the Kashmiri Speakers, schools, and exposure to mass and print media.

Overall, Siraji is the dominant language in home domain as can be seen in Fig 1 This indicates strong Siraji vitality.

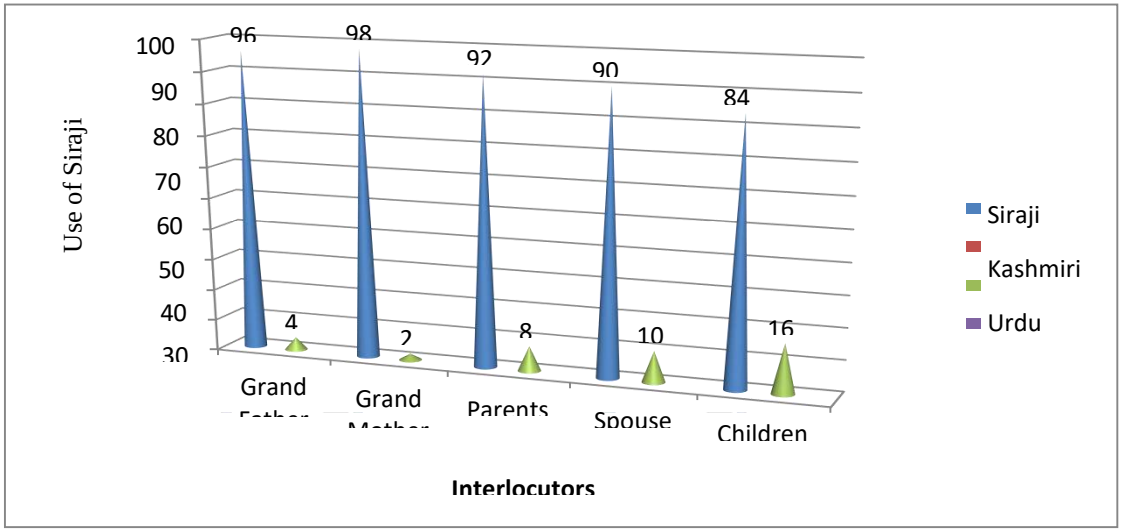

Fig.1: Language use in the Home Domain

Language use in Office domain

In the office domain the status of the office in particular and the interlocutors in general determine the language used by the people in office. The interlocutors with whom the patterns of language use in office domain are observed are:

a. Male Subordinate (MS)
b. Female Subordinate (FS)
c. Male Colleagues (MC)
d. Female Colleagues (FC)
e. Boss (B)

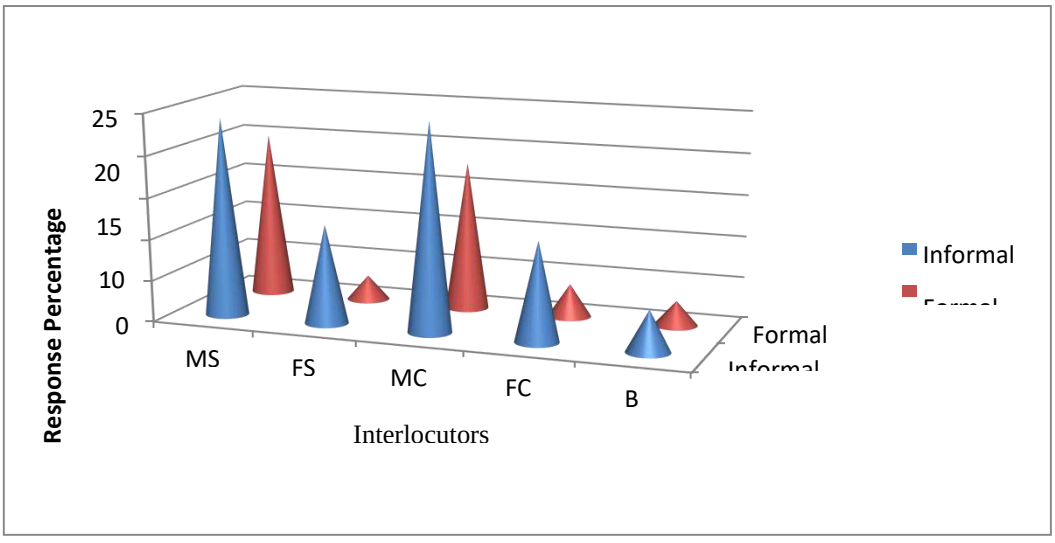

Fig 2: (%) Use of Siraji in the Office domain

As can be seen in Fig.2 that in both the formal and the informal contexts the use of Siraji is quite low (<25%) thereby indicating greater use of Urdu, English, and Kashmiri in the offices. Also, it is evident that the Use of Siraji shows decline across all interlocutors as the conversation shifts from informal to formal contexts. A fast decline in the use of Siraji is observed in Female subordinates and colleagues as compared to their Male subordinates and Male colleagues. The reason for such tremendous decrease is attributed to the fact that the females are non-Sirajis and are working within the Siraji community.

	Subordinates	Colleagues	Superior
OM	20	16	5
MM	13	12	3
YM	6	6	2

Table 4: Use of Siraji in the Office domain (Both Formal and Informal Context)

From Table 4 it is observed that there is a decrease in the use of Siraji from the Subordinates to the Superiors in all groups. In the case of Subordinates and Colleagues, Old males and Middle aged males show slight difference in scores, where as the young males show equal scores for both Subordinates and Colleagues. It is also evident from the above analysis that the Siraji language usage is negligible and the employed males prefer Urdu, English and Kashmiri in the office domain. Siraji does not go beyond 34% as observed in the interaction of old males with male colleagues and goes as low as 2% as observed in the case of interaction of Young Males with Superiors.

Language use in Market domain

Table 5 shows the summary of the language use patterns for the Market domain.

Q#	Question	n=	Siraji	%	Kashmiri	%	Urdu	%	English	%
2a	With Siraji Speakers	81	65	80			16	20		
2b	With Non - Siraji Speaker	81	81	100						
2c	With acquaintances	81	44	54			37	46		

Table 5: Language Use in Market domain

In the Market domain only two languages i.e. Siraji and Urdu as can be seen in fig 3, were found to be used by the respondents. With the Siraji Speakers in the market domain the respondents claimed to use Siraji 80% and Urdu 20%.Similarly with Non-Sirajis only Urdu (100%) is used. In the case of acquaintances use of Siraji (54%) exceeds that of Urdu (46%).

Thus from the above data we can conclude that Urdu is used with all the three groups of interlocutors.

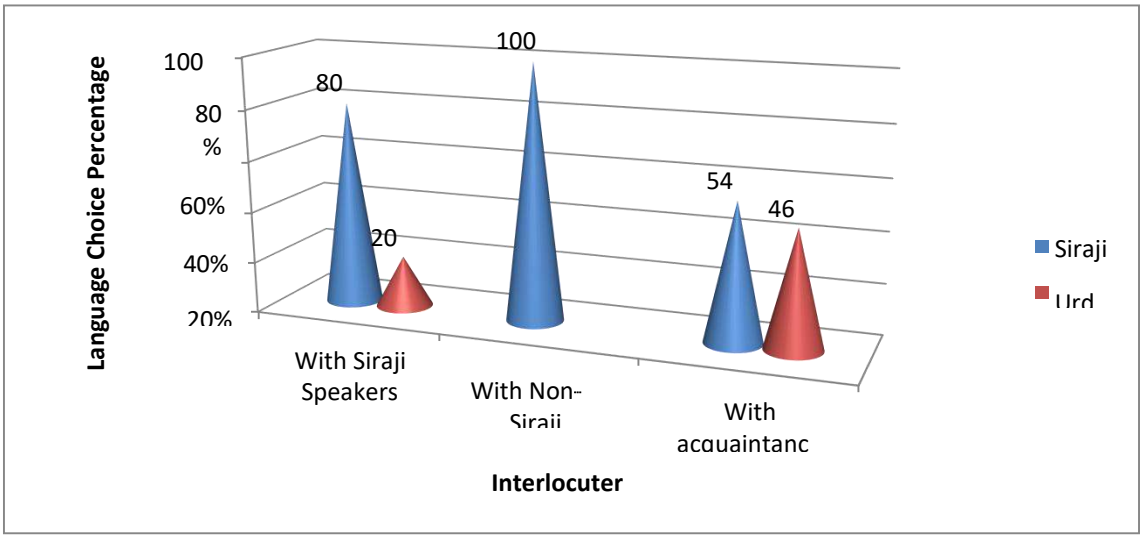

Fig.3: Language Use in Market domain

Language Use in Religious Domain

Table 6 is the summary of the language use in religious domain

Q#	Question	n=	Siraji	%	Kashmiri	%	Urdu	%	English	%
3a	Which language do you use while praying to God	81	73	90			8	10		
3b	Talking to other Worshippers	81	68	84			13	16		

Table 6: Language use in Mosque domain

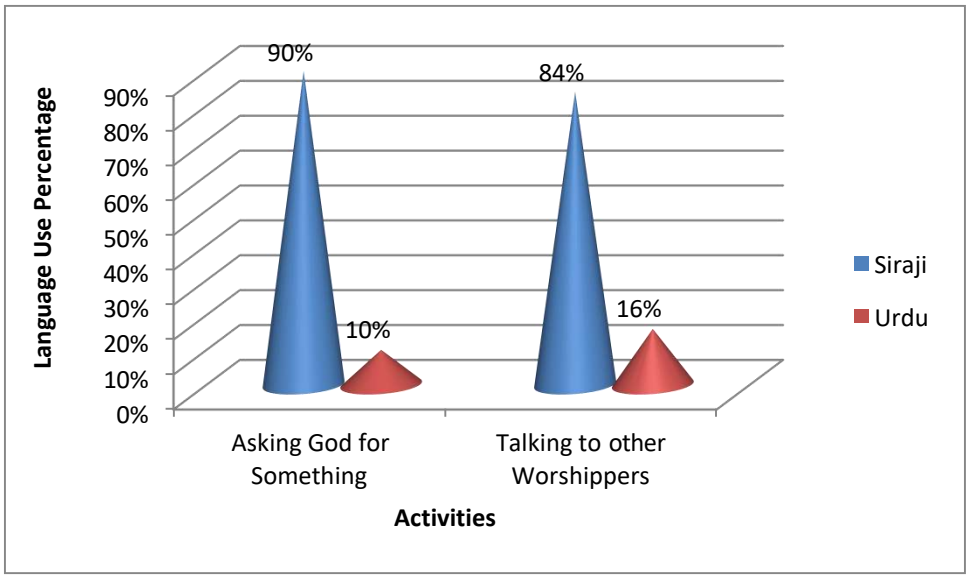

Fig.4: Language use in religious domain

In this domain as can be seen from fig.4 an average of more than 80% of the respondents chose Siraji as the only language for communication. Less than 20% of the respondents choose Urdu in their communication while Kashmiri and English are not used at all. This indicates a strong vitality of Siraji language.

Conclusion

From the above discussion of domains of language use in Siraji community of the Doda district of Jammu and Kashmir, it becomes evident that the Sirajis have maintained their mother tongue in the home domain and the use of other languages is negligible at present. On the other hand in the Office, Market, and religious domain Siraji is being replaced by languages like Urdu, English etc. The tendency of replacing Siraji with Urdu and English is more evident among the employed and educated young males in the office domain. However, language use in all domains points towards one conclusion that Siraji language vitality is strong among the Sirajis of Doda district of Jammu and Kashmir.

References:

- Batibo, H. M.2005. *Language decline and death in Africa. Causes, Consequences and Challenges*, Toronto: Multilingual Matters Ltd.
- Braunshausen and Mackey 1962:In Fishman ,J.A.1972, .*Language in Socio-Cultural Change.* Cambridge: Stranford University Press.
- Fishman, J. A. 1972 .*Language in Socio-Cultural Change.* Cambridge: Stranford University Press.
- ---1972. *The Sociology of Language* .Massachusetts: Newbury House.
- Fasold, R. 1989. *The Sociolinguistics of Society*. New York: Basil Blackwell.
- Grierson, G. A. 1919. *The Linguistic Survey of India Vol. VIII Part II.* Calcutta: Royal Asiatic Society. Reprinted Delhi: Motilal Banarsidas, 1968.

The documentation of Baram: A 'nearly extinct' language of Nepal

Dubi Nanda Dhakal
dubidhakal@yahoo.com

Abstract

This paper starts with the processes, and methodology adopted while documenting the Baram language. Some issues related to the documentation of Baram and some challanges faced during the documentation period have also been taken into account. We briefly outline the major outputs of the documentation project. As Baram is a severely endangered language, the later part of this article will try to highlight how a seriously language is influenced by a dominant language in its lexicon including morphological and syntactic features. This part will concentrate on some contact-induced changes.

Introduction

Baram (ISO'brd') is a severely endangered language spoken mainly in one village named Dandagaun (literally, "hill village") of the Takukot Village Development Committee (VDC) in the Gorkha district of western Nepal (van Driem 2007). In addition to Dandagaun, Mailung of the Takukot VDC is also a village where speakers with very low proficiency use this language. According to CBS (2002), there are 342 Baram people, and a more recent census (CBS 2011) reports that only 155 people speak the language; however, the number of speakers may actually be even lower than this figure. Agriculture is the main profession of the people in the village where they live today.

Bradley's (1997) classification of Tibeto-Burman (TB) languages includes Baram in the Eastern sub-branch of West Himalayish, which itself belongs to the Bodic branch of TB. In the same way, in Noonan's classification (2006) Baram and Thangmi form a single group (Thangmi-Baram) of central Himalayish of Bodish group of TB languages. Both Bradley (1997) and a genetic classification by Noonan (2006: 14) suggest that Baram and Thangmi are close neighbours.

The documentation of Baram entitled Linguistic and Ethnographic Documentation of the Baram Language (LEDBL) began in 2007. The documentation program was hosted by Central Department of Linguistics Tribhuvan University Nepal[1]. As the name suggests, the documentation aimed to document the Baram language by collecting the data from every possible communicative domain existing in its speech community. In addition, it aimed to develop the resources such as lexicon, sketch grammar, ethnographic profile, orthography, and primer on the basis of the database. Moreover, it also aimed to develop primer for use in

[1] I acknowledge thanks to Tej Ratna Kansakar, Yogendra Prasad Yadava, Krishna Prasad Chalise, Balaram Prasain and Krishna Poudel for their contribution while building the corpus for this project and the members of the Baram community for the data they provided. This article is based on the major documentation project funded by Hans Rausing Endangered Languages Documentation, University of London, School of Oriental and African Studies (grant MDP0158)

basic education for Baram pre-literates and to archive the materials including digital sound and video recordings of the Baram language.

Current sociolinguistic situation

The use of Nepali is gaining ground not only as a lingua franca but also as a mother tongue among the Baram people (Kansakar et al. 2011a). As Baram community constitutes a small ethnic group, they often marry speakers of other languages, which is one of the reasons for the interruption of inter-generational language transmission.

All Barams speak Nepali, and there is not a single monolingual Baram speaker. The young generation has altogether stopped speaking Baram. In interviews conducted as part of a sociolinguistic study (Kansakar et al. 2011a), some speakers who are now above 55 years responded that Baram is the first language they learnt at home. The language use of Baram is characterized by the following features: (a) since Nepali is used predominantly, the use of Baram has shrunk in most domains. In fact, it is difficult to meet a fluent Baram speaker; (b) the Baram speaking area is surrounded by other ethnic groups (such as Brahmins, Chhetris etc.), which increases the interaction of Baram with non-Baram speakers. Keeping in view the sociolinguistic situation, Eppele at al. (2012) note that Baram is a 'nearly-extinct' language.

Nowadays, Baram is hardly used in everyday life. Speakers report that the language was used when they were young. Mina Baram (one of the fluent speakers) comments that she is happy to talk to her sisters in her mother tongue at regular intervals of several months (see Kansakar et al. 2011a for details). Prolonged bilingualism of the Baram has thus given rise to contact-induced language change. Thus, Baram has a typical position in terms of its endangerment and contact-induced changes.

Linguistic profile of some speakers

We worked mainly with less than 12 speakers and even within this we spent most of the time with about half of them. Although there are some passive speakers, we did not work with them because it was not easy to obtain the linguistic data from them.

Tokman Baram (now 69 years old) is a fluent speaker in the Baram community. He possesses good retention of the langauge ever possessed by the Baram speakers. He may be considered a model of traditional speaker of Baram. He could judge the grammaticality of each sentences during the documentation period. Although he is equally fluent in Nepali, he has knowledge of some traditional folk talkes, and is well travelled. His language shows some kind of archaic forms (expressions) among Baram speakers. His texts are coherent, and uses some native verb morphemes lacking in some other speakers. He is also well respected in the community. He contributed texts which were short but well-structured. He says his children understand some Baram but don't speak it. His pronunciation is clear and speaks typical Baram.

Dammar Bahadur Baram (now in his late fifties) is one of the youngest speakers of Baram. We find massive borrowing from Nepali at lexicon and at syntactic level in his speech. During the elicitation sessions, he would sometimes ask Tokman to be sure whether a specific construciton is correct. However, he never gets opportunity to speak Baram to his wife and children as they don't speak in their mother tongue. Although his texts lack coherence with a number of loanwords and loan structures, he is the speaker who is best for elicitation of the structures. Because he is young, he is liked by other speakers in eliciting data. During the text-collection sessions, we taught him how he may ask questions to other speakers so that other speakers could answer them. He tells the stories of many kinds. His texts are full of Nepali calques. Here is an example of double marking, in both the native

perfect prefix *gi-* and the borrowed Nepali perfect suffix *-ko*[2] have been used at the same time.

(1) *hai gidako ni bal bhasa ni gida kã*

hai	gi-da-ko	ni	bal	bhasa	ni
what	PRF-say-PRF(<NPL)	also(<NPL)	Baam	language(<NPL)	also

gida	kã
PST-say	COP.PST

'What (have you) said, (they) also said in the Baram language.'

Not only the vocabulary but also the lexical items are borrowed in the texts. We see a number of double marking in the texts contributed by Dammar Baram. As we see in (1), the perfect aspect is formed by the prefix *ki-* but the speaker also makes use of the Nepali perfect marker *-ko*.

Mina Baram is the oldest living female speaker (now in her late seventies) of Baram. However, she does not have opportunites to speak in Baram to her family members. She is a fine and knowledgeable storyteller, and has good remembrance of her life. She can speak in her mother tongue for hours. Her speech contains some archic forms. When we visited her this time (August 2013), we found that she has become forgetful and relaxed about her speech habits. We tried to record a text by showing her a stimulus (, viz. Frog Story) but she could not narrate the tale by looking at it.

Panmati Baram (now 70) has a good command at Baram but her pronunciation is not clear because she does not have her front teech. She contributed some texts and her speech contains archaic forms. Because her speech is difficult to transcribe, we did not gather much texts from her during documentation.

Ram Bahadur Baram (in mid fifties) is least confident about the language but we worked with him. We found that even the possible structures in Baram are replaced by borrowed structures in his speech. He often shifts to the borrowed structures of Nepali because he does not have command over the language. He is the nearest of terminal speakers. He feels embarassed to speak in Baram. During our field work, he would prepare before he contributed the texts.

Data collection

Since Baram is a seriously endangered language, it was difficult to obtain the kinds of data we required without serious preparation. For example, all our speakers could contribute to some extent to the 'personal narratives' or 'reminiscences' but we found it difficult to obtain data on some specific topics in grammar. It was even more complicated to obtain specific paradigms unless some necessary arrangements were made. We therefore used some methods for collecting the data (a) We obtained data through direct elicitation for preparing glossaries, making paradigms, and data for specific topics while writing a grammar. Some speakers were capable of this whereas others were not. We realized that Dammar Bahadur Baram was very useful in eliciting paradigms and eliciting the grammatical structures, but his speech contains more borrowed structures. He could not be sure of the grammaticality (or acceptability) of some utterances (b) Everybody tells stories and thus narration is a skill all people are familiar

[2] The loans from Nepali have been given as (<NPL).

with to some degree. This is the main method we used for collecting large amount of texts and discourses. In many cases the language consultants narrated their personal experiences or travelogues. During the data collection, they commenced their narration when we said to them "Could you tell us about …. " How is wine prepared?", "How did you travel from Dandagaun to Gorkha?...." Several texts were obtained using this technique. In the beginning of the data collection, we were not sure whether we could collect the data we intended. We were very successful in obtaining the data by making use of this technique. (c) A large number of speech genres cannot be obtained simply by narration. We therefore tried to obtain the corpus by role play and simulation. A large number of conversations were recorded using this technique. They were mainly semi-controlled. (d) It was necessary to use stimuli in the process of data collection. Some kinds of structures are related to certain kinds of activities, or objects. People can easily describe the objects they see. For example, we found that the pictures can be used to elicit structures like, 'This is ….on/in/back/forward' or 'There is a stone on ….'. In addition, we also showed them a documentary and asked them to tell what they saw in the documentary. Details can be found in Dhakal et al. (2011).

Major outputs

As mentioned in the objectives, the major activities of this documentation project were to collect and annotate the corpus, prepare grammar, etc. This section discusses the major outputs of the project.

Corpus

The recorded texts were transcribed in ELAN and imported to Toolbox for inter-linearizing the texts[3]. The corpus of Baram is a documentation corpus. A total of 33 hours of oral texts have been transcribed and inter-linearized. We tried to record the data of every possible speech genre of this linguistic community. Since the language exists only in the oral form, the monolingual texts comprise the major component of the corpus. A large number of texts were narratives (personal, historic and remembrances). All the texts collected were semi-controlled. Almost all texts were recorded in our field office located some kilometers away from the villages where the language is spoken. There are many semi-controlled dialogues.

Although the language is not recorded in the real speech community, a number of topics were included in the corpus so that different grammatical structures would be obtained. Therefore, there are some texts which tell us recipes (e.g. how liquor is made, or how the millet is planted and harvested). However, the corpus does not comprise the ritual language as the language is not used on the occasion of rituals. For the purpose of linguistic analysis we built a corpus containing 8577 types (words) and approximately 250000 tokens (total occurrences of the words) using the completely annotated files that have already been sent to Endangered Languages Archieve (ELAR) for archiving. From the corpus we extracted necessary grammatical information about the words and structures.

The Baram corpus is based on the six structural features of a good documentation corpus as suggested by Woodbury (2003:46-47). The Baram corpus along with the developed language resources has been archived in Endangered Languages Archive (ELAR), SOAS, University of London and it is on the website accessible to linguists, ethnographers and the Baram

[3] These computer software programs are used in documenting the Languages. . Eudico Linguistic Annotator (ELAN) was developed at the Max Planck Institute for Psycholinguistics in The Netherlands. Similarly, Toolbox is used for text inter-linearization and dictionary compiling.

speech community. The corpus consists of media files, annotation files, metadata files and elicitated data. The corpus consists of altogether 191 sessions from 13 different speakers of which 144 sessions are monological and 47 are conversational. In total there are 42146 utterances and 186214 words as produced by the speakers. We tried to include different kinds of text genres while building the corpus (cf. Lüpke 2005:97).

Orthographic analysis

From the analysis of the sound system of Baram, we found that Baram sound system has largely converged with the Nepali phonological system. It has lost the typical features of Tibeto-Burman languages, so its phonology can be well represented by the orthography based on Devanagari script used in Nepali. With the consensus of the ethnic community, we adopted Devanagari script for writing Baram.

Baram-Nepali-English Dictionary

This dictionary is based on about 200 annotated and analyzed sessions of 30 hours containing more than 300,000 tokens. All the unique words that occurred in the corpus were collected and divided into base forms and affixes, and they were included in the head entries. The selection of the head entries was made regardless of their origin, so a large number of loan words have been incorporated in this dictionary. The information about the derivation and inflection has been given in the outline of the grammar in the 'Introduction' part of the dictionary. So there are no sub-entries. The *to*-infinitive forms of the verbs and the basic forms of the other word classes have been taken as citation forms. This dictionary consists of three parts: Baram-Nepali-English Dictionary, Nepali-Baram Index, and English-Baram Index.

Grammar

This book shows how a seriously endangered language which was in a dormant state could be documented and analyzed. Baram is a language which has not been spoken for about fifteen years or so. However, the older speakers have good memory of their language. Data were largely drawn from those speakers. This grammar is based on the corpus as a part of the documentation of the Baram language. It deals basically with sociolinguistics, phonology, morphophonology, nominal morphology, adjectives, verb morphology, adverbs, closed word classes, noun phrase, simple sentences and their modifications and clause combining in the language.

Primers

The textbooks are based on the model curriculum designed by Curriculum Development Centre, Government of Nepal for mother tongue education. The textbooks are bilingual in Baram and Nepali because the Baram children exclusively speak Nepali. In collaboration with the language community and Nepal Baram Association, we planned to implement the books in 1, 2 and 3 classes in primary education in three schools. We hoped that this implementation would help the revitalization and promotion of the Baram language. However, the community has not been able to use them for multilingual education programs.

Difficulties during documentation

We experienced some difficulties during the documentation period. Firstly, we encountered some difficulties because of equipments. We tried to explore high quality equipments, such as microphones, audio recorder and video camera from the local market but they were not available. Secondly, getting the fluent speakers was not easy as there were few fluent speakers. Thirdly, we found that it is a really challenging job to build a balanced and natural

corpus in a seriously endangered language like Baram. As Baram is not used naturally in the community, we could not record the natural use of the language nor include some of the domains of language use. Power cut off was another difficulty we experienced. The problem of load shedding has worsened in Nepal over the last few years (up to 17 hours per day). The armed conflict continued till the end of the first year of the project. This created some difficulty in undertaking the fieldwork.

Documentation and language contact

Based on the corpus built for the Baram language, we found a number of facts about the contact-induced changes in Baram. We outline them in the following sections. We aim to show how the lexicon and morphosyntactic features of an endangered language gradually disappear. Moreover, we also show the lexical can morphosyntactic borrowing from a source language.

Lexicon and loanwords

It is obvious that an endangered language borrows a number of vocabulary items. The result of loanword analysis is determined by the size of vocabulary and the semantic domains it contains. The result is the bigger the wordlist, the higher the borrowed words from other languages and vice versa. The borrowability scale may also be different. The Baram dictionary by Kansakar et al.(2011b) suggests that there are only 1,022 native Baram words among a total of 3,652 words, i.e., less than one third. The other two-thirds constitute borrowings from Nepali, English, or other languages via Nepali. This dictionary also includes some commonly used affixes. However, for the loanword analysis, only a total of 3646 words were included leaving aside the affixes. The list of words of each word classes categorized into nouns, verbs, adjectives, adverbs and function words are given in Table 1. We see that 70.81% words are borrowed. And the native vocabulary comprises only 29.18%.

	Parts of Speech	**Total**	**Native**	**Borrowed**
1	Adjectives	478	101 (21.12%)	377 (78.87%)
2	Nouns	2162	481 (22.24%)	1681 (77.75%)
3	Adverbs	230	85 (36.95%)	145 (63.04%)
4	Verbs	697	356 (51.0%)	341 (48.92%)
5	Function words	79	41 (51.89%)	38 (48.10%)
		3646	1064 (29.18%)	2582 (70.71%)

Table 1: Loanwords by major word classes

A higher percentage of adjectives is borrowed compared to other word categories as can be seen in Table 2 and Figure 1. As expected, it is shown that function words are borrowed least. However, the percentage is slightly lower than the verbs. The borrowability scale is shown in following Figure 1.

Adjectives > Nouns > Adverbs > Verbs > Function words

Figure 1: Borrowing scale

We would like to present an additional table here in order to show the borrowing pattern among the basic words collected based on the word-list used in the Loanword Typology Project (LWP) wordlist. Moving to another wordlist, we find slightly different results in borrowing. We listed the words which are same as we find in Loanword Typology Project

(LTP) following Haspelmath and Tadmor (2009). Among 1460 words, only 1307 equivalents were found in Baram. When we look at this vocabulary list, we get the result as presented in Table 2. We see that only 64% words were borrowed.

	Parts of Speech	**Total**	**Native**	**Borrowed**
1	Nouns	708	178 (25.14%)	530 (74.85%)
2	Adjectives	176	152 (86.36%)	124 (70.45%)
3	Adverb	66	27 (40.90%)	39 (59.09%)
4	Verbs	326	186 (57.02%)	140 (42.94%)
5	Function words	31	22 (70.96%)	9 (29.03%)
		1307	425 (32.51%)	842 (64.42%)

Table 2: Loanwords by major word classes

When we compared the data based on the LWP wordlist, we obtained only 1307 words. They are categorized into five major word classes, viz. noun, verb, adjective, adverb and function words. We see slightly different results in this table. First, we see the highest percentage of borrowing is found in nouns (74.85%). This is followed by adjectives. In this specific word list, 70.45% adjectives are borrowed. Function words are the least borrowed category in Baram when we look at LWT wordlist.

The complete wordlist was divided into 24 semantic domains following Haspelmath and Tadmor (2009). On the basis of the semantic class, we see the following borrowing pattern: Percentage of loanwords by semantic domains in Baram-

> Religion and belief, modern world (over 90 %), > The house, quantity, time, emotion and values, cognition, social and political relations, warfare and hunting, law (over 80 %), > The physical world, possession, speech and language, miscelleneous function words (over 70%) > Kinship, The body, Food and drink, Clothing and grooming, Agriculture and vegetation, basic actions and technology, spatial relations (over 60%), > Sense perception, motion (over 50%) > Animals (over 40%).

We see that very little can be said about the loanwords in Baram. We can see that the least borrowings or great stability is seen in the areas of animals followed by sense perception and motion. This may be because Baram lexical items contain a number of vocabularies related to insects. By contrast, the highest percentage of borrowing is seen in religion and belief, and modern world.

Morphological borrowings

Nepali grammatical morphemes such as the plural suffix and numeral classifiers in nouns have been borrowed. In addition, there is evidence of double marking in the case of the evidential construction, where elements of the original Baram and the innovative Nepali constructions are fused. Speakers use both the native as well as borrowed constructions in Baram in a large number of cases. Major grammatical features are compatible between these two languages, i.e. Nepali and Baram.

(2) The plural suffix in nouns in Baram, *-həru*, is borrowed from Nepali, where *-həru* is used to form a plural noun (Acharya 1991: 78). Interestingly, the form of the plural suffix has also fused with the second person pronoun in Baram, e.g., *naŋ* '2(SG)', *naŋ-ru* '2-PL'. The pronoun *nuŋ* '2PL' is also used interchangeably. Grierson (1909) assumes that the native plural suffix is *-du* in the word *hu-du* 'they'. The same suffix is assumed to have been the plural suffix in Hunter (1978 [1868]). This suffix can not be obtained in elicitation, nor does it occur in the corpus at present. Instead, the plural suffix *-həru* is typically used. Consider (3).

(3) *nəsjo nəsjo balhəru testak*
nəsjo nəsjo bal-həru testak
last year last year man-PL(<NPL) like that
'Men (were) like that long long ago.' [Panmati]

Keeping Grierson's observations in mind, it appears that native *-du* has now been substituted by the plural suffix *-həru* from Nepali or else by *-baŋ*.
The morphological borrowings are presented in Table 3. It is problematic to determine the morphological borrowing. In some cases, determining the loanwords is far from simple because we do not know how we can prove that a certain morpheme is a borrowing. Despite the problem, borrowed morphemes are listed in Table 3.

Morphemes	**Native**	**Gloss**	**Borrowed status**	**Use**
-həru	*-du*	PL	definitely borrowed	replacement
-oṭa	*-wa*	CLF	definitely borrowed	coexistence
-ma	*-ga*	LOC	definitely borrowed	coexistence
-lai	*-gəi*	ACC	definitely borrowed	coexistence
-səŋ	*?*	COM	definitely borrowed	unclear
dzəna	*-eŋ*	CLF	perhaps borrowed	coexistence
-ko	*-ku*	GEN	perhaps borrowed	replacement
-chə/-cə	*-a/o*	EVID	definitely borrowed	double marked
-ko	*gi-*	PRF	definitely borrowed	double marked

Table 3: Borrowed morphemes in Baram

There are three cases when we see the morphological borrowings in Baram as shown in Table (3). First, some grammatical morphemes have completely replaced the native ones. Such cases are indicated as 'replacement'. The speakers of the language no longer use the native morphological features in this case. Second, there are some cases in which both the native and borrowed patterns are used in different discourse contexts by different speakers. Such cases are indicated as 'coexistence'. There are some morphemes which are definitely borrowed because there are evidences that they are borrowed from Nepali. Some affixes seem to be borrowed from Nepali but we don't have evidence to show whether they are the result of language contact or of internal change, such as the genitive marker *-ko*. In another case, the constructions are double-marked because both the native and borrowed structures are used at the same time. They are indicated as 'double-marked'. The details can be found in Dhakal (2014).

Syntactic borrowings

Other cases of grammatical borrowing include the copula *ho*, the comparative construction, desiderative construction, borrowing of some clause combining devices, such as coordinator *rə*, relative correlative, sentence conjuntion with *əni*, and some discourse particles. Here, I will illustrate only three of them, viz. coordinating conjunction *rə* 'and', disjunctive coordinator *ki* 'or' and directional constructions.
The native coordinating conjunction *to* 'and' as reported in Hunter (1978 [1868]) is no longer used. Instead, the borrowed coordinating conjunction *rə* is used. Examples such as (4) abound in the corpus. We do not find a single occurrence of the native conjunction *to* 'and' in the entire corpus data.

(4) *pəŋkak akja rə ucwa kepthul hola ni*

pəŋkak	akja	rə	ucwa	ki-pəthul	hola	ni
frog	dog	and(<NPL)	child	PST-rear	may be(<NPL)	PART(<NPL)

'Somebody has perhaps reared the frog, dog and the child.' [Panmati]

The coordinator *rə* 'and' is used in the texts of all speakers. Consider example (5).

(5) *tjo akja rə ucwa kina belage*

tjo	akja	rə	ucwa	ki-na	bela-ge
that(<NPL)	dog	and(<NPL)	child	PST-sleep	time(<NPL)-LOC

'While the dog and child were sleeping…' [Tokman]

We may cite another example of clause combining where the disjunctive coordinator *ki* is borrowed from Nepali (and other Indic languages). The disjunctive coordinator *ki* 'or' is borrowed from Nepali or Indic origin. The disjunctive coordinator *ki* 'or' is used to link the sentences and indicates a choice among several grammatical items in Hindi (Kachru 2006: 240) in Nepali and perhaps in other IA languges as well. The disjunctive coordinator *ki* in Baram is of Indic origin and borrowed from its direct neighbour Nepali.

(6) *wat ho ki hai ho ubleiŋ kihuk*

wat	ho	ki	hai	ho		
wasp	cop.NPST(<NPL)	or(<NPL)	what	cop.NPST(<NPL)		
uble-iŋ	ki-huk					
fly-PROG	pst-SIT					

' Something like wasp or what else (other insect) was flying (over them)' [Tokman]

(7) *tigaŋ chetəŋ əbə namge hila ki əbə hai tukko*

tigaŋ	chet-əŋ	əbə	nam-ge	hil-a	ki	
and then	carry-SEQ	now(<NPL)	house-LOC	return-EVID	or(<NPL)	
əbə	hai	tuk-ko				
now(<NPL)	what	do-INF				

'Then carrying it (he did not know) whether to return home or what to do.' [Dammar]

As pointed out by Palosaari and Campbell (2011:112), there is preference for analytic constructions over syntactic ones in endangered languages. Another case of contact-induced change is the use of non-morphological means to mark directionals in purposive clauses.
The native way of encoding directionals in purpose clauses involves two directional prefixes in Baram, i.e. *he-* andative (movement away from the speaker) and *hjuŋ-* venitive (movement towards the speaker) (cf. Bybee et al. 1994: 320). In (8), the movement of the speaker away from the deictic center is morphologically coded. Similarly, the movement is morphologically coded in (9) to show the speaker's movement towards the deictic center. These directionals also yield a purposive reading. This construction occurs frequently in the corpus.

(8) *kuni hjadaŋgo*

kuni	he-adaŋ-go
where	ANDA-search-INF

'Where to go to search?'

(9) *bal kjoŋkham.*

bal ki-hjuŋ-kham

man PST-VEN-say

'The man came to say (it).'

In sharp contrast to examples (8-9), the direction with purposive clauses is also expressed in another way in Baram. This kind of construction seems to be a borrowing and is assumed to be an influence of Nepali. First of all, let's consider a Nepali sentence in (10).

(10) *keṭo nuhaunə dharama gəjo*

keṭo nuhau-nə dharama dza-jo

boy bath-INF tap-LOC go-PST.3SG.M.NH

'The boy went to the tap to bathe.' (Nepali)

The constructions are alike to Nepali in the sense that in a Nepali purposive clause the infinitive *-nə* is followed by the main verb, such as *nuhau-nə gə-jo* 'went to bathe' in (10). As can be seen in this example, the purposive clause precedes the main clause. The infinitive is also used to mark the purposive clause in Nepali. Structurally, the Nepali example (10) differs from example (11) in that it does not have a morphological way of expressing directional. Examples (11-12) from Baram is similar to the construction in Nepali.

(11) *pheri asiŋgo jachə*

pheri asiŋ-go ja-chə

again take-INF go-EVID

'They went to search.'

Since Baram has a morphological mechanism of expresing purposive clauses, the example such as (11) indicates a structural borrowing from Nepali to Baram. While relatively younger speakers use the structure such as (11), the elderly speakers and the corpus contain the structure like (8-9). There is no difference in meaning of these two kinds of constructions. The speaker instead could have said *kjasaŋ*.

Modal expressions

Two modal expressions are widely borrowed from Nepali in texts, viz. *rəichə* and *hola*. They are discussed below. Regarding the borrowing of modality expressions, Matras (2009: 187) proposes:

(12) Obligation > necessity > **possibility** > ability > desire

We see borrowings in possibility and desire. In Nepali grammar *hola* shows the possibility. Similarly, we see the same form being used in the texts of speakers. We not only see the verb being borrowed, this is also followed by the Nepali particle *ni*.

(13) *ubaŋe pəŋkak kepthul hola ni*

ubaŋ-e pəŋkak ki-pəthul hola ni

they-ERG frog PST-keep may be(<NPL) PART(<NPL)

'They might have kept the frog.' [Panmati]

(14) *ləllumle hai tukko hola*
ləllum-le hai tuk-ko hola
fall-COND what do-inf may be(<NPL)
'What might happen if (he) falls ?' [Tek]

The finite forms of verbs expressing the modality are also used by the speakers. For example, in (14), Nepali grammar *hola* shows the possibility.

Another modal used widely in the texts is *rəichə*. In Nepali grammar *hola* shows the possibility. Similarly, we see the same form being used in the following sentence from Panmati. We not only see the verb being borrowed, it is also followed by the Nepali particle *ni*.

(15) *dzərajoko siŋ rəichə*
dzərajo-ko siŋ rəhechə
deer-GEN horn remain.EVID(<NPL)
'It might the horn of the deer.' [Panmati]

(16) *kəstak bidi pərak rəichə*
kəstak bidi pərak rəichə
of what kind kind cliff remain.EVID(<NPL)
'What kind of slope it was (It was very sloppy).' [Tek]

7.5 Discourse particles

We find a number of discoruse particles being borrowed into Baram discourses. Matras (2009: 193) mentions "discourse markers occupy a position at the very top of the borrowabilty hierarchy". Here is an illustration of the discourse particle *ni*. The particle *ni* in Nepali indicates, "you must realize that" or "be sure that" (Schdmidt 1993: 344).

(17) *pəŋkak ŋikham ni*
pəŋkak ŋi-kham ni
frog NPST-say PART(<NPL)
'It is called frog, you know' [Dammar]

(18) *ubaŋe pəŋkak kepthul hola ni*
ubaŋ-e pəŋkak ki-pəthul hola ni
they-ERG frog PST-keep may be(<NPL) PART(<NPL)
'They might have kept the frog.' [Panmati]

We see that the discourse particle *ni*, *ki*, *həi*, *rə*, *caĩ* are directly borrowed from Nepali. The further illustration can be found in Dhakal, Chalise and Gurung (2013).

Conclusion

The major outputs of the project were the corpus, dictionary, grammar, ethnographic sketch, primers, and an ethnographic documentary. Here we also propose an orthography for writing system of Baram. We find a massive borrowing of lexical items into Baram from Nepali. In addition to that, we also find a number of morphosyntactic borrowings into Baram. Before concluding the paper, it is obvious that unless some serious revitalization attempts are made, the language will not be spoken by new learners. The documentation of an endangered language leads to a number of interrelated linguistic findings.

References

Acharya, J.R. (1991). *A Descriptive Grammar of Nepali and an Analyzed Corpus.* Washington D.C., Georgetown University Press.

Bradley, D. (1997). "Tibeto-Burman Languages and Classification." In: *Papers in Southeast Asian Linguistics No. 14: Tibeto-Burman Languages of the Himalayas*. Canberra, Pacific Linguistics (series A-86):1-72.

Bybee, J, Perkins, R. & Pagliuca W. (1994). *The Evolution of Grammar*. London and Chicago: The University of the Chicago Press.

Central Bureau of Statistics (2002). *Population census 2001: National report.* Kathmandu, Central Buereau of Statistics.

___(2012). *National Population and Housing Census 2011(National report).* Kathmandu, Central Buereau of Statistics.

Dhakal, D. N (2013). 'Contact-induced changes in Baram', *NorthEast Indian Linguistics*. Vol. 6, 2014, 167-190. Canberra, Australian National University: Asia-Pacific Linguistics Open Access.

Dhakal, D. N; Chalise, K. P; & Gurung, E. (2013). *The Influence of Nepali in Minority Languages: A Case of Baram*. A Research Report submitted to University Grants Commission, Bhaktapur, Nepal.

Dhakal, D. N; Kansakar T. R; Yadava, Y. P.; Prasain; B. R; Chalise, K. P; Poudel, K. (2011). ELDP Data Collection: Some Baram experiences. In: *Proceedings of the Third Students' Conference of Linguistics in India*. (Eds) Choudhary, Narayan and Gibu Sabu. 23-34.

Eppele, John W.; M. P. Lewis, D. R. Regmi; Y. P. Yadava. (2012). *Ethnologue: Languages of Nepal*. Kathmandu: Linguistic Survey of Nepal and SIL International.

Grierson, G. A. (1909). *Linguistic Survey of India*, Vol.3. Tibeto-Burman Family, Part I. Delhi, Book Faith India, 405-406.

Haspelmath, M. (2008). Loanword Typology: Steps toward a systematic cross-linguistic study of lexical borrowability. In: *Aspects of Language Contact*. (Eds). Thomas Stolz, Dik Bekkar, Rosa Salas Palomo. Berlin/ New York: Mouton de Gruyter. pp. (43-62).

Haspelmath, M. & Tadmor, U. (2009). Loanword Typology Project and the World Loanword Database. In: *Loanword in the world's Languages: A Comparative Handbook*. (Eds) M. Haspelmath & U. Tadmor. Berlin/ New York: Mouton de Gruyter (1-34).

Hunter, W.W. (1978 [1868]). *A Comparative Dictionary of the Languages of India and High Asia.*New Delhi, Cosmo Publications.

Kachru, Y. (2006). *Hindi*. Amsterdam/ Philadelphia, John Benjamins Publishing Company.

Kansakar, T. R.; Yadava, Y. P. ; Chalise, K. P. ; Prasain, B.; Dhakal, D. N. ; and Paudel, K. (2011a). "A Sociolinguistic Study of the Baram Language." *Himalayan Linguistics*. 10.1: 187-226.

_____ (2011b). *Baram Nepali English Dictionary*. Kathmandu, LEDBL, and Central Department of Linguistics, TU.

_____ (2011c). *A Grammar of Baram.* Kathmandu, LEDBL and Central Department of Linguistics, TU.

Lüpke, F. (2005). Small is beautiful: Contributions of field-based corpora to different linguistic disciplines, illustrated by Jalonke. In Peter K. Austen (Ed.). *Language Documentation and Description Vol.3.* (pp. 75-105). London: Hans Rausing Endangered Languages Project.

Matras, Y. (2009). *Language Contact.* Cambridge, Cambridge University Press.

Noonan, M. (2006). "Contact-induceed change in the Himalayas: the case of the Tamangic Languages." downloaded from http://www.uwm.edu/~ noonan. Accessed 17/03/2010.

Palosaari, N. and L. Campbell (2011). Structural aspects of language endangerment. In P. K. Austen & J. Sallabank (Eds). *The Cambridge Handbook of Endangered Languages.* (pp. 100-119). Cambridge: Cambridge University Press.

Schmidt, R. L. (Ed.) (1993). *A Practical Dictionary of Modern Nepali.* New Delhi, Ratna Sagar.

van Driem, G. (2007). "South East and the Middle East." In C. Moseley Ed. *Encyclopedia of the world's endangered languages.* London and New York, Routledge: 283-348.

Woodbury, A. C. (2003). Defining documentary linguistics. In P. K. Austin (Ed.), *Language Documentation and Description* vol. 1 (pp. 35-72). London: Endangered Language Project.

A Phonological Study of Burushaski Spoken in Srinagar

Aejaz Mohammed Sheikh
&
Saima Jan
Department of Linguistics
University of Kashmir
Srinagar-190001
Email: sheikh.aejazm@gmail.com

Abstract:

Burushaski, a language isolate, is mainly spoken in the northern areas of Pakistan with more than 85,000 of estimated Burushos (Burushaski speakers). Burushaski is also known to its speakers by the names Mishaski (my language), Brugaski, Boorishki. The language comprises three divergent dialects, i.e., Hunza, Nagar and Yasin. The language is considered as language isolate due to the genetic differences with the surrounding language families like Dardic, Indo-Aryan, etc. More than 300 speakers of Burushaski are also found in Srinagar and Pulwama districts of Jammu and Kashmir state. Burushaski being one of the unexplored languages of the valley needs to be analyzed within the descriptive phonological framework.

In the backdrop of the above discussion the paper aims to undertake the phonological study of Burushaski spoken in Srinagar district of Jammu and Kashmir.

1. Area and Speakers

Burushaski, also known by the names of Boorishki, Brugaski, Kanjut, Werchikwar and Mishaski is regarded as a language isolate. It is mainly spoken in two separate areas in northern Pakistan. The major valleys of the eastern Burushaski spoken area are Hunza and Nagar which belong to the Karakoram mountains and to the Hunza-Nagar district of Gilgit-Baltistan (federal capital territory of Pakistan, formerly known as Northern areas); and the major valley of the western area is Yasin which belongs to the Hindukush mountains and to Ghizer district of Gilgit-Baltistan. Thus, the researchers usually differentiate the two varieties as Eastern Burushaski (Hunza-Nagar variety) and Western Burushaski (Yasin variety). Anderson (1997) estimated the total number of Burushos

residing in the areas of Pakistan to be 50-60,000. "The *Ethnologue* (2004) reports 96, 800 speakers of Burushaski in Pakistan". Moreover, based on the government census figures, the total number of Burushos found in the areas of Pakistan is around 85,000-87,000.

The Burushaski speakers found in the valley of Jammu and Kashmir are settled in and around a small locality at the foothills of Hari Parbat in Srinagar, the summer capital of the Jammu and Kashmir state. They constitute a small linguistic community situated in the heart of the city and the locality is known as Mohalla Azur Khan, named after Raja Azur Khan. Burusho speakers are called by the name of Bot Raj by Kashmiris. In Kashmiri "Raj" means "king" and is used perhaps because most members of the community are the descendants of a tribal king who was originally from Nagar in Pakistan. It is for this reason that some members of the community claim to ascribe to a "higher" social status in terms of lineage, while the rest are considered of a non-royal descent by them. Jammu and Kashmir Burushos of the present day include some members who were originally from Hunza and probably migrated at a later stage. Burushos enjoy a state domicile and have recently been offered a Scheduled Tribe status by the Government of India. The migration of Burushos to Kashmir valley from Hunza and Nagar took place in 1891 due to various political upheavals at these places. Certain historical accounts believe that this group migrated in intervals from 1891 onwards. Burushos are also reported to live in Batamaloo and Bemina area of Srinagar which is away from the main Burusho group. As no specific information regarding the number of Burushaski speakers is available from any source, thus, based on the personal information gathered from the members of the speech community there are around 35-40 families residing in the area. Thus, an estimated average of 200-300 Burushos are living in Srinagar.

2. About the Language

Linguistically, Burushaski has been termed as a language isolate because it is said to have no genetic relationship with any of the surrounding language families like Indic, Sino-Tibetan, Dardic etc. Ruhlen (1987) classified Burushaski as a language isolate as its genetic affiliation remains a complete mystery, but Ruhlen (1994) reports on a possible classification of Burushaski as a separate branch of a newly proposed Dene-Caucasian super stock. More recently, Bengtson (1995) list a few etymologies relating Burushaski to the Yeniseian languages, spoken by the hundred people along the Yenisei River in Siberia. Burushaski reflects agglutinative characteristics, as many kinds of prefixes as well as suffixes are found in the language. The syllable structure is CCVCC and both onset and coda clusters (CC) are observed at the word initial and final position respectively. Burushaki has some Indian areal linguistic features, like, the echo-formation, expressives and onomatopoeia, but lack some of the characteristic features which are commonly observed in other languages, i.e. presence of double causatives and

classifiers. For example, Burushaski is surrounded by Kashmiri language which is having double causatives as well as classifiers.

There are three divergent dialects of Burushaski. They are named after the main valleys where they are spoken: Hunza, Nagar, and Yasin (also called Werchikwar). The dialect of Yasin is thought to be the least affected by contact with neighboring languages like Urdu, Shina, Balti, Wakhi and Pashto. All the three dialects of Burushaski are mutually intelligible. Until recently, Burushaski had no script of its own and written literature was very scarce. Texts available, if, any, would use modified Persio-Arabic script.

3. Burushaski in Contact with Neighboring Languages

By the end of the 19th century, Burushaski has been greatly influenced by contact with the neighboring languages. Urdu, Khowar, Shina, Wakhi, and Balti are some of the languages which were in contact with Burushaski. Having lost contact with their parent community in Pakistan over a century ago, the language of Buroshos has undergone several changes which make it systematically different from other dialects of Burushaski spoken in Pakistan. Thus, the study of Phonological aspect of Burushaski spoken in Srinagar city of Jammu and Kashmir is interesting from a linguistic point of view, as the language has been in isolation from the parent community for more than 120 years.

4. Methodology

The data used for the study was collected from the Burushaski speakers settled in and around the areas of Badamwari situated by the foothills of Hari Parbat in Srinagar, the summer capital of the Jammu and Kashmir state. The locality is called by the name Mohalla Azur Khan. The data was collected in the months of January and February, 2015. An extensive questionnaire consisting of words and sentences was prepared to elicit the data. The data was collected by using a highly sophisticated voice recorder. The collected data was later transcribed and subjected to analysis.

5. Phonology

Phonology is derived from an ancient Greek word *"phon"* which means *"voice"* or *"sound"* and *"logos"* which means *"subject of discussion".* Broadly speaking, phonology is the sub-discipline of linguistics concerned with the study of the patterns of the speech sounds of languages. It deals with, "description of the systems and patterns of speech sounds in a language" (Yule, 1997). It aims to identify the number and nature of the phonemes, allophones, the pattern of their distribution, phonotactics, phonological processes, etc. Below is provided a description of the phonology of Burushaski language spoken in Srinagar.

5.1 Segmentals

In linguistics, the term segment may be defined as any discrete unit that can be identified, either physically or auditorily in the stream of speech. Segments are called discrete because they are separate and individual, such as consonants and vowels. Below is provided the phonemic inventory of Burushaski language.

➢ *5.1.1. Consonants*

	Bilabial	**Labio-dental**	**Dental**	**Alveolar**	**Retroflex**	**Palatal**	**Velar**	**Uvular**	**Glottal Stops**
STOP									
vl. Unasp	**/ p /**		**/ t /**		ʈ		**/ k /**	**/ q /**	
vl. Asp	**/ p^{h} /**		**/ t^{h} /**		ʈ ^{h}		**/ k^{h} /**	**/ q^{h} /**	
vd. Unasp	**/ b /**		**/ d /**		ɖ		**/ g /**		
AFFRICATES									
vl. Unasp			**/ ts /**			**/ c /**			
vl. Asp			**/ ts^{h} /**			**/ c^{h} /**			
vd. Unasp						**/ ʤ /**			
NASAL	**/ m /**		**/ n /**				**/ ŋ /**		
TRILL									
LATERAL				**/ l /**		**/ r /**			
FRICATIVES									
vl.		**/ f /**		**/ s /**	**/ ʂ /**	**/ ʃ /**		**/ x /**	**/ h /**
vd.		**/ v /**		**/ z /**	**/ ʐ/**	**/ ʒ /**			
GLIDE	**/ w /**					**/ j /**			

➢ *5.1.2. Vowels*

	Front	**Central**	**Back**
High	**/ i: /**		**/ u: /**
Lower High	**/ i /**	**/ ɨ /**	**/ u /**
Mid	**/ e / / e: /**		**/ o / / o: /**
Lower Mid		**/ a /**	
Low		**/ a: /**	

5.2. Phonemic Contrast

A phoneme is the smallest contrastive unit in the sound system of a language. A study of the minimal pairs in the data exemplifies the phonemic contrast prevalent in the language. They are used to demonstrate that two phones constitute two separate phonemes in the language. Semi-minimal pairs are also found in Burushaski language, where the pair of

words differs by at least two continuous sounds. The examples of phonemic contrast are given below.

➢ *5.2.1. Contrastive Distribution of Burushaski Consonants*

Consonants	Example	Gloss
/p/ vs /b /		
/ p /	/pal/	Sleep
/ b /	/bal/	Wall
/ p / vs / p^h /		
/ p /	/pal/	Sleep
/ p^h /	/p^hal/	Fruits
/ t/ vs/ t^h /		
/ t /	/tal/	Ceiling
/ t^h /	/t^ha:l/	Plate
/ t / vs / d /		
/ t /	/tal/	Ceiling
/ d /	/da:l/	Elevated
/ ʈ / vs / $ʈ^h$ /		
/ ʈ /	/ʈak/	Button
/ $ʈ^h$ /	/$ʈ^h$ak/	Flap
/ ʈ / vs / ɖ /		
/ ʈ /	/ʈak/	Button
/ ɖ /	/ɖak/	Knock
/ k / vs /k^h /		
/ k /	/kar/	Wander
/ k^h /	/k^har/	Insect
/ k / vs/ g /		
/ k /	/kar/	Wander
/ g /	/gar/	Wedding
/ q / vs /q^h /		
/ q /	/qam/	Pit
/ q^h /	/q^ham/	Curry
/ ts / vs / ts^h /		
/ ts /	/tsar/	Tear
/ ts^h /	/ts^har/	Splash

/ c / vs / cʰ /		
/ c /	/cak/	Hit
/ cʰ /	/cʰak/	Hunger
/ c / vs / ʒ /		
/ c /	/cal/	Fight
/ ʒ /	/ʒal/	Scatter
/ f / vs / v /		
/ f /	/fan/	Fan
/ v /	/van/	Van
/ s / vs / z /		
/ s /	/sar/	Thread
/ z /	/zar/	Jolt
/ ʂ / vs / s /		
/ ʂ /	/bupuʂ/	Squash
/ s /	/gus/	Woman
/ ʂ / vs/ ʐ /		
/ ʂ /	/ʂaw/	Hit
/ ʐ /	/ʐaw/	Dislike
	/ ʃ / vs / ʂ /	
/ ʃ /	/buʃ/	Cat
/ ʂ /	/bupuʂ/	Squash
/ h / vs / z /		
/ h /	/har/	Ox
/ z /	/zar/	Jolt
/ m / vs / n /		
/ m /	/dam/	Steam
/ n /	/dan/	Stone
/ m / vs / ŋ /		
/ m /	/dam/	Steam
/ ŋ /	daŋ/	Sleep
/ r / vs / l /		
/ r /	/pʰur/	Flight
/ l /	/pʰul/	Cup
/ w / vs / j /		
/ w /	waʂ/	Bend
/ j /	/ajaʂ/	Sister

5.2.2. Contrastive Distribution of Burushaski Short Vowels

Vowel	Example	Gloss	Vowel	Example	Gloss
/i/	/bis/	Fat	**/e/**	/bes/	why
/a/	/han/	One	**/i/**	/hin/	door
/o/	/gos/	Heart	**/u/**	/gus/	women
/e/	/mel/	Wine	**/a/**	/mal/	/field/
/a/	/tal/	pigeon	**/o/**	/tol/	snake
/ɨ/	/gaʃkɨ/	Rope	**/i/**	/galgi/	wing
/u/	/ulpureŋ/	They	**/i/**	/ilpureŋ/	he

5.2.3. Contrastive Distribution of Burushaski Long Vowels

Vowel	Example	Gloss	Vowel	Example	Gloss
/i/	/hir/	Man	**/i:/**	/asi:r/	near
/u/	/ʒuk/	Touch	**/u:/**	/ʒu:k/	grief
/e/	/ʃe/	Eat	**/e:/**	/ʃe:/	wool
/o/	/mos/	Flood	**/o:/**	/mo:s/	wife
/a/	/gar/	wedding	**/a:/**	/ga:r/	trance

5.*3. Description and Distribution of Burushaski Phonemes*

The description and distribution (in different positions, i. e., initial, medial and final, in a word) of Burushaski phonemes is provided below.

5.*3. 1. Description and Distribution of Burushaski Consonants*

Consonant	Description	Distribution	Gloss
/ p /	voiceless un-aspirated bilabial stop	/pal/, /tsapi:/, /c^{h}ap/	lengthy sleep, shoe, Meat
/ p^{h} /	voiceless aspirated bilabial stop	/p^{h}uŋi:/, /map^{h}e:r/	moustache, old
/ b /	voiced un-aspirated bilabial stop	/bats^{h}in/, /ibran/, /ʒura:b/	thigh, see, socks
/ t /	voiceless un-aspirated dental stop	/tol/, /le:lati:/, /ba:lt/	snake, know, apple
/ t^{h} /	voiceless aspirated dental stop	/t^{h}a:l/, /mat^{h}an/, /siridint^{h}/	plate, far away, woman
/ d /	voiced un-aspirated dental stop	/daldanum/	wide
/ ʈ /	voiceless un-aspirated alveolar stop	/ʈuʈaŋ/, /baʈ/	dark, skin

/ ʈʰ /	voiceless aspirated alveolar stop	/ʈʰak/, /ruʈʰ/	flap, sit
/ ɖ /	voiced un-aspirated alveolar stop	/ɖak/, /ɖaɖaŋ/	knock, drum
/ **k** /	voiceless un-aspirated velar stop	/kar/, /hukay/, /hik/	wonder, dogs, ones
/ **k**^{h} /	voiceless aspirated velar stop	/k^{h}ãru:/	lice,
/ **g** /	voiced un-aspirated velar stop	/gus/, /hurgas/	woman, thick,
/ **q** /	voiceless un-aspirated uvular stop	/qam/	pit
/ **q**^{h} /	voiceless aspirated uvular stop	/q^{h}am/	curry
/ **ts** /	voiceless un-aspirated dental affricate	/tsar/	tear
/ **ts**^{h} /	voiceless aspirated dental affricate	/ts^{h}ar/	splash
/ **c** /	voiceless unaspirated palatal affricate	/cak/	hit
/ **c**^{h} /	voiceless aspirated palatal affricate	/c^{h}ak/	hunger
/ ʒ /	voiced unaspirated palatal affricate	/ʒal/	scatter
/ **f** /	voiceless labio-dental fricative	/furdeli:/, /fa:lis/	fly, some
/ **v** /	voiced labio-dental fricative	/vaxt/, /ava/, /p^{h}iv/	time, yes, flies
/ **s** /	voiceless alveolar fricative	/sirindint^{h}/, /usa:num/, /sis/	woman, long, person
/ **z** /	voiced alveolar fricative	/zar/	jolt
/ ʂ /	voiceless retroflex fricative	/ʂaw/, /bupuʂ/	hit, squash
/ ʐ/	voiced retroflex fricative	/ʐaw/	dislike
/ ʃ /	voiceless palatal fricative	/ʃapik/, /guʃpur/, /jaʃ/	food, prince, sky
/ **x** /	voiceless uvular fricative	/xurxamaʃ/, /rax/	garbage, desire
/ **h** /	voiceless glottal fricative	/hurʈʰ/, /kunah/	live, stick

/ m /	voiced bilabial nasal	/mimi:/, /c^{h}umo:/, /usa:num/	mother, fish, long
/ n /	voiced alveolar nasal	/nimo/, /mugonoh/, /gaman/	go, seed, root
/ ŋ /	voiced velar nasal	/c^{h}aŋti/, /uts^{h}riŋ/	vomit, guts
/ r /	voiced palatal lateral	/ramad/, /isarkan/, /ʒatur/	dig, kill, quince
/ l /	voiced alveolar lateral	/le:l/, /galgi:/, /m^{j}ul/	knowledge, wing, belly
/ w /	voiced bilabial semi-vowel	/waʂ/	bend
/ j /	voiced palatal semi-vowel	/ja:re/, /ujum/, /loj/	under, big, fox

5.3.2. Description and Distribution of Burushaski Vowels

Vowels	Description	Burushaski Example	Gloss
/ i /	high front unrounded short vowel	/ibran/, /mindil/, /ts^{h}aŋti/	see, chest, vomit
/ i: /	high front unrounded long vowel	/mi:/, /diʃi:mi/, /galgi:/	our, open, wing
/ e /	mid front unrounded short vowel	/beʃal/, /mineh/, /ume/	where, drink, your
/ e: /	mid front unrounded long vowel	/le:lati:/, /map^{h}e:ri/, /map^{h}e:r/	know, aged woman, aged woman
/ ɨ /	high central unrounded short vowel	/gaʃkɨ/	rope
/ a /	low central unrounded short vowel	/ʒah/, /daldanum/, /balats^{h}/	I, wide, bird,
/ a: /	low central unrounded long vowel	/ʒa:h/, /bija:num/, /mana:s/	my, thin, happen
/ u /	high back rounded short vowel	/ujum/, /osk^{h}usas/, /qarqa:muts/	big, fear, hens
/ u: /	high back rounded long vowel	/u:/, /ququru:qu/, /k^{h}ãru:/	their, onam, lice
/ o /	mid back rounded short vowel	/k^{h}os/, /k^{h}algo/	this, insect

/ o: /	mid back rounded long vowel	/o:mus/, /alto:lum/, /ts^{h}umo:/	tongue, other, fish

5.4. Consonant Clusters

Different types of consonant clusters are found in Burushaski Combinations like stop + /r/ is found at the initial place. Different types of clusters involve combination of sibilant + stop, nasal + stop, lateral + stop, stop + stop, affricate + fricative. Some examples are provided below.

Phonetic Form	Gloss	Combination
/**tr**ap/	clapping	stop + lateral
/gu**nts**/	day	nasal + affricate
/ba:**lt**/	apple	lateral + stop
/ba**sk**/	more	sibilant + stop
/**br**end/	give	stop + lateral
/bre**nd**/	give	nasal + stop
/**ṭh**aŋ t^{h}i/	push	stop + fricative
/**ts**il/	water	stop + sibilant
/da**nɖ**/	stone	nasal + stop

5.5. Syllable Structure and Pattern

A syllable is typically made up of a syllable nucleus (most often a vowel) with optional initial and final margins (typically, consonants). Syllables are often considered the phonological "building blocks" of words. Syllables have internal structure: they can be divided into parts. The parts are onset and rhyme; within the rhyme we find the nucleus and coda. A syllable may or may not have an onset and a coda. The basic syllable structure of Burushaski is as under:

- CV
- VC
- CVC
- CCVC
- CVCC
- CCVCC

In Burushaski monosyllabic, disyllabic as well as polysyllabic words are found.

5.5.1. Monosyllabic Words

Example	Syllable Pattern	Gloss
/um/	CV	Thou
/oʃ	VC	Neck
/k^{h}us/	CVC	This
/breɳ/	CCVC	See
/danɖ/	CVCC	Stone
/brenɖ/	CCVCC	Give

5.5.2. Disyllabic Words

Example	Syllable Pattern	Gloss
/ele:/	V.CV	There
/ajo:n/	V.CVC	All
/amli:/	VC.CV	Where
/k^{h}ule:/	CV.CV	Here
/awjar/	VC.CVC	Husband
/menen/	CV.CVC	Who
/walto:/	CVC.CV	Four
/p^{h}urgas/	CVC.CVC	Thick

5.5.3. Polysllabic Words

Example	Syllable Pattern	Gloss
/inaka:/	V.CV.CV	With
/usanam/	V.CV.CVC	Long
/utsiriŋ/	VC.V.CVC	Guts
/bijanu:/	CV.CV.CV	Thin
/isarganɳ/	V.CVC.CVC	Kill
/aʒama t^{h}i:/	V.CV.CV.CV	Dig
/uskuza:s/	VC.CV.CVC	Fear
/tʃ^{h}ur dali:/	CVC.CV.CV	Stab
/medu:mus/	CV.CV.CVC	Knee
/haralt^{h}i/	CV.CVC.CV	Rain
/daldanum/	CVC.CV.CVC	Wide
/ʃika:r ati:/	CV.CV.CV.CV	Hunt
/tikiɖgɔl/	CV.CVC.CVC	Fall
/gurbaltiŋ/	CVC.CVC.CVC	Trousers

References

Anderson Gerg. 1997. Burushaski Phonology. Winona Lake, IN: Eisenbrauns.

Bengston, J.D. 1995. Caucasian and Sino-Tibetan: A Hypothesis of S.A. Starostin. *General Linguistics.* Volume 36.

Fasold, R. 1989. *The Sociolinguistics of Society*. New York: Basil Blackwell.

Grierson, G. A. 1919. *Linguistic Survey of India*. (Volume X). Delhi: Munshi Ram Monoharlal Publishers.

Hock, Hans Heinrich. 1986. *Principles of Historical Linguistics*. Berlin, New York: Mouton de Gruyter.

Hock, Hans Heinrich, Joseph, B. D. 1996. *Language History, Language Change and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin, New York: Mouton de Gruyter.

Munshi, Sadaf.2006. *Jammu and Kashmir Burushaski: Language, Language Contact and Change.* Ph.D Dissertation. Austin: University of Texas.

Ruhlen, Merritt. 1987. *A Guide to World's Languages*, Volume 1: Classification. Stanford: Stanford University Press.

Ruhlen, Merritt. 1994. *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.

Sheikh, Aejaz M., Kuchay, Sameer A. 2014. Noun Morphology of Kishtwari: A Brief Sketch. *Indian Linguistics*, Vol. 75, No. 3-4.

Srivastava, R.N. 1995. *Studies in Language and Linguistics.* Delhi: Kalinga Publications.

Starostin, Sergi A, Ruhlen, Merritt. *Proto-Yeniseian Reconstructions, with Extra- Yeniseian Comparisons*. www.merrittruhlen.com/files/yeniseian.pdf. Retrieved on 01-02-2016.

Wardhaugh, R. 2005.*An Introduction to Sociolinguistics, Fifth Edition*. London: Wiley-Blackwell Publishing.

Yule, George. 1997. *The Study of Language.* Cambridgwe:Cambridge University Press.

Websites:

www.http://ethnologue.com/language/bsk.
Retrieved on 24/01/2014.