Surkhet Journal
[A Peer-Reviewed, Open-Access, Indexed in NepJOL, DOI, Multidisciplinary Journal]
ISSN 2362-1230 (Print)
Published by the Research Management Cell (RMC)
Surkhet Multiple Campus, Birendranagar, Surkhet
Tribhuvan University
DOI:https://doi. org/10. 3126/surkhetj. v3i1. 84733

कर्णालीका लोककथाहरूको पाठसङ्ग्रह निर्माण

Received on 18 July, 2025; Accepted on 31 August, 2025; Published on…September, 2025

बलराम प्रसाईं[१]
कृष्णप्रसाद घिमिरे[२]
दुर्गाबहादुर घर्ती[३*]
*Correspondence: durgagharti@gmail.com, https: orcid.org/0009-0009-0118-8344
[१]पिएचडि, प्राध्यापक, भाषाविज्ञान केन्द्रीय विभाग, [२]पिएचडि, प्राध्यापक, नेपाली केन्द्रीय विभाग
[३]पिएचडि, सहप्राध्यापक, नेपाली केन्द्रीय विभाग

लेखसार

कर्णाली क्षेत्रमा प्रचलित मौलिक तथा मौखिक रूपमा पुस्तान्तरण भइरहेका लोककथाहरूको कम्प्युटरले पढ्न सक्ने तथा प्रशोधन योग्य एउटा पाठसङ्ग्रह निर्माण गर्ने यस अनुसन्धानको उद्देश्य रहेको छ। लोककथाहरू कर्णालीका सबै जिल्लाहरूलाई प्रतिनिधित्व गर्ने गरी उद्देश्यमूलक तथा परिस्थितिजन्य विधिलाई उपयोग गर्दै ५७ ओटा लोककथाहरूको श्रव्य रेकर्ड गरियो। उक्त सङ्कलित लोककथाहरूलाई देवनागरी युनिकोडमा लिपीकरण गर्दै हरेक कथाको एउटा कम्प्युटर फाइल बनाइयो। यसरी तयार गरिएको पाठसङ्ग्रहमा कम्प्युटर प्रोग्रामिङ कोडको प्रयोग गरी हरेक कथामा शब्दसङ्ख्या, अद्वितीय शब्दसङ्ख्या तथा शाब्दिक घनत्व प्रतिशत निकालेर तथ्याङ्कसहित उक्त पाठसङ्ग्रहलाई कम्प्युटर सफ्टवेयर (AntConc) मा प्रविष्टि गरी केही भाषिक पक्षहरू सन्दर्भ शब्द, शब्दगुच्छ, शब्दशृङ्खला, सहप्रयोगी आदि कुराका नमुनाहरू प्रस्तुत गरिएको छ।

शब्दकुञ्जी : पाठसङ्कलन, पाठसङ्ग्रह, शब्दगुच्छ, शब्दशृङ्खला, सहप्रयोगी, लोककथा

विषयप्रवेश

नेपालको मध्यपश्चिमोत्तर भागमा अवस्थित कर्णाली क्षेत्र भौगोलिक, सांस्कृतिक र भाषिक दृष्टिले अति विविधतायुक्त क्षेत्र हो। हिमाली, पहाडी र भूपरिधीय भूभागहरू मिलेर बनेको यो क्षेत्र दुर्गम, प्राकृतिक सौन्दर्य र जैविक विविधताले भरिपूर्ण छ। यहाँ रहेका प्रमुख जिल्लाहरू जस्तै सुर्खेत, जुम्ला, हुम्ला, डोल्पा,

मुगु, कालिकोट, रुकुम पश्चिम, सल्यान र डोल्पा आफैँमा सांस्कृतिक र भाषिक विविधताको उदाहरण हुन्। यस क्षेत्रमा बसोबास गर्ने जातजातिहरू, जस्तै पहाडी ब्राह्मण, क्षेत्री, ठकुरी, राउटे, मगर, दलित तथा अन्य तिब्बती मूलका जनजातिहरूको आफ्ना मौलिक परम्परा, भाषा र मौखिक सम्पदाका धनी छन् । सामाजिक संरचना सामूहिक सहयोगमा आधारित छ भने धार्मिक विश्वासहरू मिश्रित रूपमा हिन्दू, बौद्ध र आदिवासी परम्परामा आधारित छन् । यो सम्पूर्ण क्षेत्र मौलिक मौखिक परम्परा, जातीय चाडपर्व, संस्कार र लोकसाहित्यको जगेर्ना गर्ने जीवनशैलीमा बाँधिएको छ ।

लोककथा भन्नाले पुस्तौँदेखि मौखिक रूपमा चल्दै तथा भनिँदै आएका कथाहरूलाई बुझिन्छ जसले कुनै समुदायको सामाजिक, सांस्कृतिक, धार्मिक तथा ऐतिहासिक जीवनलाई चित्रण गर्छन् । यिनीहरू प्रायः काल्पनिक प्रतीत भए पनि तिनको मूलमा जीवनका यथार्थ अनुभव, नैतिक शिक्षा र सांस्कृतिक मूल्य लुकेका हुन्छन् । लोककथाहरूको भाषा सहज, जनभाषामा आधारित, पुनरावृत्तिमूलक संरचनामा हुने गर्दछ । कर्णाली क्षेत्रका लोककथाहरूमा स्थानीय जनजीवन, देवीदेवताप्रति आस्था, प्रकृति र जनावरहरूप्रतिको दृष्टिकोण, पारिवारिक सम्बन्ध, वर्गीय संघर्ष, साहसिकता र बुद्धिमत्ताका विषयहरू केन्द्रमा देखिन्छन् । उदाहरणका लागि, ससुरा-बुहारीको सम्बन्धमा आधारित कथा, सौतेनी नाताका काथा, पशुहरूसँग संवाद गर्ने पात्र, अन्यायविरुद्धको न्याय प्राप्ति, प्राकृतिक प्रकोपपछि आउने जीवन परिवर्तन आदि विषयहरू प्रमुख छन् । यी कथाहरू अझै पनि गाउँका बुढापाकाहरू, देवीस्थानका पूजारीहरू तथा मेलापातका बेलामा विभिन्न व्यक्तिहरूले सुनाइने मौखिक सम्पदाका रूपमा जीवित छन् ।

पाठसङ्कलन भाषावैज्ञानिक विश्लेषण भन्नाले ठूलो सङ्ख्यामा सङ्कलित भाषिक पाठहरूमा तथ्याङ्कीय आधारमा भाषिक विशेषता पत्ता लगाउने प्रक्रिया हो । यसमा कम्प्युटर-आधारित प्रविधिको प्रयोग गरी शब्दको आवृत्ति, शब्दगुच्छ, सहसंबन्ध/सहप्रयोगी, शब्दशृङ्खला, प्रमुख शब्द, वाक्य संरचना, शैली, पुनरावृत्त संरचना आदिको विश्लेषण गरिन्छ । कर्णाली क्षेत्रका मौखिक रूपमा चल्दै आएका लोककथाहरूलाई सङ्कलन गरी लिपिबद्ध गरी अङ्कीय पाठसङ्कलन/पाठसङ्ग्रह तयार पारेर विश्लेषण गर्नुको उद्देश्य भनेको ती कथाहरूमा प्रयोग भएको यस किसिमको भाषा, शैली, भावात्मक प्रवृत्ति, सांस्कृतिक सन्दर्भ र स्थानीय भाषिक विशेषताहरूको पहिचान गर्नु हो । कर्णाली क्षेत्रको मौलिक भाषिक सम्पदा मात्र उजागर गर्दैन, नेपाली भाषाविज्ञान, सामाजिक मानवशास्त्र, मौखिक साहित्य तथा अङ्कीय मानविकी क्षेत्रमा नयाँ अनुसन्धानका आयामहरू पनि थप्दछ । साथै, यस कार्यले नेपाली भाषामा पाठसङ्कलन भाषावैज्ञानिक अनुसन्धानलाई स्थानीय तहसम्म विस्तार गर्न सहयोग पुऱ्याउनेछ, जुन प्राज्ञिक दृष्टिकोणले अत्यन्तै महत्त्वपूर्ण छ ।

अनुसन्धानका प्रश्नहरू

यस अनुसन्धानको मुख्य लक्ष्य भनेको कर्णाली क्षेत्रका प्रचलित मौलिक तथा मौखिक लोककथाहरूको एउटा पाठसङ्ग्रह तयार पार्ने हो र यसका लागि निम्नलिखित अनुसन्धान प्रश्नहरू तय गरिएका छन् :

- कर्णाली क्षेत्रमा के कस्ता मौखिक रूपमा रहेका लोककथाहरू पाइन्छन् ?
- लोककथाहरूलाई कम्प्युटरले पढ्ने र प्रशोधन गर्न सक्ने गरी कसरी तयार पार्न सकिन्छ ?
- सङ्कलनबाट तयार भएको पाठसङ्ग्रहमा प्राथमिक विशेषताहरू के कस्ता छन् ?

अनुसन्धानका उद्देश्यहरू

कम्प्युटरमा प्रशोधन तथा विश्लेषण गर्न सकिने गरी कर्णाली क्षेत्रका प्रचलित लोककथाहरूको एउटा पाठसङ्ग्रह तयार पार्न निम्न उद्देशहरू रहेका छन् :

- कर्णाली क्षेत्रमा पाइने मौलिक तथा मौखिक रूपमा रहेका लोककथाहरूको पहिचान गरी श्रव्यरूपमा सङ्कलन गर्ने ।
- लोककथाहरूलाई कम्प्युटरले पढ्ने र प्रशोधन गर्न सक्ने गरी देवनागरी युनिकोडमा टङ्कन गरी तयार गर्ने ।
- सङ्कलनबाट तयार भएको पाठसङ्ग्रहमा शब्दरूप, शब्दगुच्छ, शब्दशृङ्खला, सहप्रयोगी, आदिको विश्लेषण नमुना प्रस्तुत गर्ने ।

पूर्वकार्यको समीक्षा

लोककथाको अध्ययन मानवशास्त्रीय, भाषावैज्ञानिक, साहित्यिक तथा सांस्कृतिक दृष्टिकोणबाट अत्यन्तै मूल्यवान् मानिन्छ । २० औं शताब्दीको पूर्वार्धदेखि नै लोककथाको विश्वव्यापी वर्गीकरण र विश्लेषणमा ठूल्ठूला कार्यहरू हुँदै आएका छन् । नेपाली सन्दर्भमा लोककथाको अनुसन्धान तुलनात्मक रूपमा कम भए पनि लोकसंस्कृति विश्लेषणमा लामिछाने (२०१८), राकेश (२००५), बन्धु (२०५८) लगायतका विद्वान्हरूले लोकसाहित्यको अध्ययनमा प्रचुर मात्रामा योगदान पुर्‍याएका छन् ।

नेपालमा भाषाको पाठसङ्ग्रह बनाउने र यसमा आधारित अध्ययन गर्ने प्रचलन खासै देखिँदैन, यद्यपि सन् १९७० को दसकमा प्रा.डा. चूडामणि बन्धुले विभिन्न जिल्लाबाट नेपाली कथ्य भाषाको एउटा करिब ९६ पृष्ठको दस्तावेज तयार गरेको र उक्त दस्तावेजलाई सन् २००६ मा ब्रिटिस एकेडेमीको सहयोगमा अङ्कीकरण गरी नेपाली राष्ट्रिय पाठसङ्ग्रहको अंशको रूपमा विद्युतीय भण्डार गरेको पाइन्छ । नेपालमा आधिकारिक रूपमा पाठसङ्ग्रहको निर्माण कार्य भाषा सञ्चार परियोजना (सन् २००५–२००७) ले मदन पुरस्कार पुस्तकालय (मपुपु) मा युरोपेली सङ्घको आर्थिक सहायतामा मपुपु, भाषाविज्ञान केन्द्रीय विभाग, त्रिवि, ओपन विश्वविद्यालय र लाङ्कास्टर विश्वविद्यालय, संयुक्त अधिराज्य र गोथेनवर्ग विश्वविद्यालय, स्वीडेनको संलग्नमा नेपाली राष्ट्रिय पाठसङ्ग्रह तयार भयो । यसैगरी विभिन्न व्यक्तिगत वा संस्थागत प्रयासमा नेपाली भाषाका धेरै खाले पाठसङ्ग्रहहरू तयार गरी विद्युतीय प्रणालीमा भण्डारण भएको पनि पाइन्छ । तर लोककथाहरूको पाठसङ्ग्रह भने अहिलसम्म तयार भएको सूचना उपलब्ध छैन । तसर्थ सानै भए पनि यो कर्णाली क्षेत्रमा पाइने 'लोककथाहरूको पाठसङ्ग्रह' तयार गर्नु उपयुक्त नै ठहरिन्छ ।

यसरी हेर्दा नेपालका विभिन्न स्थानमा पाइने लोककथाहरूको विभिन्न परिपेक्ष्यबाट अध्ययन, अनुसन्धान तथा विश्लेषण भएको पाइन्छ, यद्यपि मौलिक तथा मौखिक रूपमा रहिरहेका लोककथाहरूलाई एकै ठाउँमा ल्याई पाठसङ्ग्रहको निर्माण गरी कम्प्युटरको प्रयोगबाट अध्ययन, अनुसन्धान तथा विश्लेषण भएको पाइँदैन, तसर्थ यस कार्यलाई अघि बढाउनु नै शोधप्रस्तावको केन्द्रबिन्दु रहेको छ ।

अनुसन्धान विधि

पाठ सङ्कलन : परम्परागत रूपमा पुस्तान्तरण हुँदै आएका मौखिक लोककथाहरूलाई नियन्त्रणमूखी विधिबाट सङ्कलन गर्दा हामीले लिएको लक्ष्य प्राप्ति गर्न त्यति न्यायोचित नहुने भएकाले उद्देश्यमूलक तथा परिस्थितिजन्य विधिलाई अवलम्बन गरी स्थान तथा समयको बन्देज नलगाई कथा वाचकलाई जुन कथा जस्तो अवस्थामा आउँछ सोही रूपमै वाचन गर्न लगाई प्रत्यक्ष श्रव्य रेकर्ड गरी कथाहरूको सङ्कलन क्षेत्रकार्य मार्फत् नै गरियो ।

पाठ-सङ्ख्या : अनुसन्धान परियोजनामा प्रस्ताव बमोजिम करिब ५० ओटा कथाहरू सङ्कलन गर्ने भए तापनि क्षेत्रकार्यको क्रममा अति छोटा, छोटा, मझौला तथा लामा गरी ५७ ओटा कथाहरू सङ्कलन भएका छन् ।

लिपीकरण : सङ्कलन गरिएका श्रव्य रेकर्डहरूलाई कम्प्युटरमा देवनागरी युनिकोडमा लिपीबद्ध गरी कम्प्युटरले पढ्नयोग्य तथा प्रयोगयोग्य बनाइयो । हरेक कथाको अलग-अलग फाइल तयार गरी कम्प्युटरको सहायताले हरेक कथामा कति शब्दहरू छन्, कति अद्वितीय शब्दहरू कति छन् र शाब्दिक धनत्व कति हुन्छ भन्ने कुराको गणना गरियो, उक्त विवरण निम्न तालिकामा देखाइएको छ ।

प्रस्तुति : कम्प्युटरमा प्रयोगयोग्य हुने गरी तयार गरिएको पाठसङ्ग्रहलाई कम्प्युटर सफ्टवेयर Anthony, L. (सन् २०२४), AntConc (Version 4.3.1) को प्रयोग गरी विभिन्न पक्षहरूको विश्लेषण गर्न सकिने केही दृष्टान्टहरूको प्रस्तुति ।

अनुसन्धानको महत्त्व

- विलुप्त हुँदै गइरहेका मौखिक परम्पराको दस्तावेजीकरण र संरक्षणमा योगदान पुग्नेछ ।
- स्थानीय भाषाहरू, संस्कृति र जनजीवनको अध्ययनका लागि उपयोगी श्रोत प्रदान गर्नेछ ।
- अङ्कीय मानविकी, भाषाविज्ञान, मानवशास्त्र, र लोकसाहित्य क्षेत्रहरूमा पाठसङ्ग्रह भाषावैज्ञानिक अनुसन्धानका लागि आधार तयार गर्नेछ ।
- साहित्यिक सम्पदा तथा स्थानीय ज्ञान प्रणालीको विकासमा योगदान गर्नेछ ।

पाठसङ्कलनको वितरण तथा तथ्याङ्क

कर्णाली प्रदेश भौगोलिक दृष्टिकोणले अत्यन्तै विविध र विस्तृत भए तापनि यस क्षेत्रभित्र मानव बसोबासको वितरण समान रूपमा देखिँदैन । यस असमान वितरणका पछाडि आधुनिकीकरणको

प्रक्रिया, भू-राजनीतिक कारकहरू, तथा आन्तरिक स्थानान्तरण जस्ता विविध कारणहरू रहेका छन् । विशेषतः सुर्खेत जिल्ला, कर्णाली प्रदेशको प्रशासनिक केन्द्रको रूपमा स्थापित भएपछिको अवस्थाले गर्दा, यस क्षेत्रमा अन्य जिल्लाहरूबाट जनसंख्या स्थानान्तरण हुने दर उच्च देखिन्छ । यिनै कारणले गर्दा अनुसन्धान प्रयोजनका लागि आवश्यक लोककथासम्बन्धी सामग्री (डेटा) सङ्कलन गर्दा सुर्खेत जिल्लालाई उच्च प्राथमिकतामा राखिएको छ । दोस्रो तहमा जुम्ला, जाजरकोट, दैलेख, तथा पश्चिम रुकुम जिल्लाहरूलाई समावेश गरिएको छ भने बाँकी जिल्लाहरूलाई तेस्रो तहको प्राथमिकतामा राखिएको छ । तय गरिएका प्राथमिकताका आधारमा विभिन्न जिल्लाबाट मौखिक रूपमै प्रचलित लोककथाहरूको श्रव्यसङ्कलन गरिएको छ । सङ्कलन गरिएका कथाहरूको परिमाण र तिनीहरूको जिल्लागत वितरण सम्बन्धी विवरण चित्र १ मा प्रस्तुत गरिएको छ ।

चित्र १

जिल्लाअनुसार लोककथा सङ्कलनको वितरण

माथि उल्लेख भएबमोजिमका कथाहरूको सङ्कलन पश्चात् देवनागरी युनिकोडमा लिपीकरण गरी कम्प्युटरले पढ्न र प्रशोधन गर्न योग्य बनाइयो । हरेक कथाको अलग अलग कम्प्युटर फाइल बनाइयो र ती फाइलहरूको नाम कथाकै नामबाट .तहत खाकामा राखियो । यसरी तयार भएका कथाका फाइलबाट हरेक कथामा भएका शब्दसङ्ख्या, अद्वितीय शब्द र कथा भएको शाब्दिक घनत्व प्रतिशत निकालियो जसलाई निम्न तालिकामा प्रस्तुत गरिएको छ। यो तथ्याङ्कबाट करिब १५ ओटा जति कथाहरू अति छोटा, २६ ओटा छोटा र १४ ओटा मध्यम खाल्का र २ ओटा लामा कथाहरू सङ्कलन भएको देखिन्छ । यो विभाजन कुनै सैद्धान्तिक नभएर तथ्याङ्कको अवलोकनमा आधारित तथा यादृच्छिक मात्र हो । यसलाई तालिका १ मा प्रस्तुत गरिएको छ ।

तालिका १

लोककथाहरूको सूची

कथा	शब्दसङ्ख्या	अद्वितीय शब्द	शाब्दिक घनत्व प्रतिशत
स्यार्पुतालको कथा	७०	५७	८१.४३%
लामो बुद्धि र छोटो बुद्धिको कथा	१११	६८	६१.२६%
छट्टु र सोझो मान्छेको कथा	११७	९०	७६.९२%
खुवाँरी बराहको कथा	११८	९५	८०.५१%
गुन नमान्नेको कथा	१२४	८५	६८.५५%
डाँडा छिवाङको मण्डलीको कथा	१३३	८३	६२.४१%
डिग्रे साइकुमारीको कथा	१४२	१०६	७४.६५%
जिमिदारको लोभको कथा	१४८	११०	७४.३२%
राजाको दुःखको कथा	१४९	१०८	७२.४८%
मन्त्र जान्ने मान्छेको कथा	१५३	११४	७४.५१%
कमल दहको कथा	१५५	१०१	६५.१६%
बिरालु र बाघको कथा	१७९	११९	६६.४८%
धामी र खुवाँरी बराहको कथा	१८५	१४४	७७.८४%
छायाँक्षेत्रको शिलाको कथा	१९०	१३१	६८.९५%
सूर्य गरो र आँख्या गरोको कथा	१९९	११४	५७.२९%
सही इन्साफको कथा	२२२	१४५	६५.३२%
अर्कासँग लागेकी आइमाईको कथा	२२८	१६७	७३.२५%
स्याल र भालुको कथा	२३६	१६१	६८.२२%
राम्रीकाँडाको घाँटुको कथा	२४७	१६६	६७.२१%
आमाछोराको कथा	२५८	१७५	६७.८३%
स्याल र बघेनीको कथा	२९९	१७३	५७.८६%
कपटी स्वास्नी	३२०	१९६	६१.२५%
न्याउली चरीको उत्पत्ति	३२२	१८५	५७.४५%
ससुराबुहारीको कथा	३४६	१९२	५५.४९%
देउती बज्यैको कथा	३४६	२४२	६९.९४%
रोङरोङे चरीको कथा	३५०	२०१	५७.४३%
वीरहरूका कथा	३५८	२२७	६३.४१%
बुढाबुढी र भालुको कथा	३५९	२१२	५९.०५%

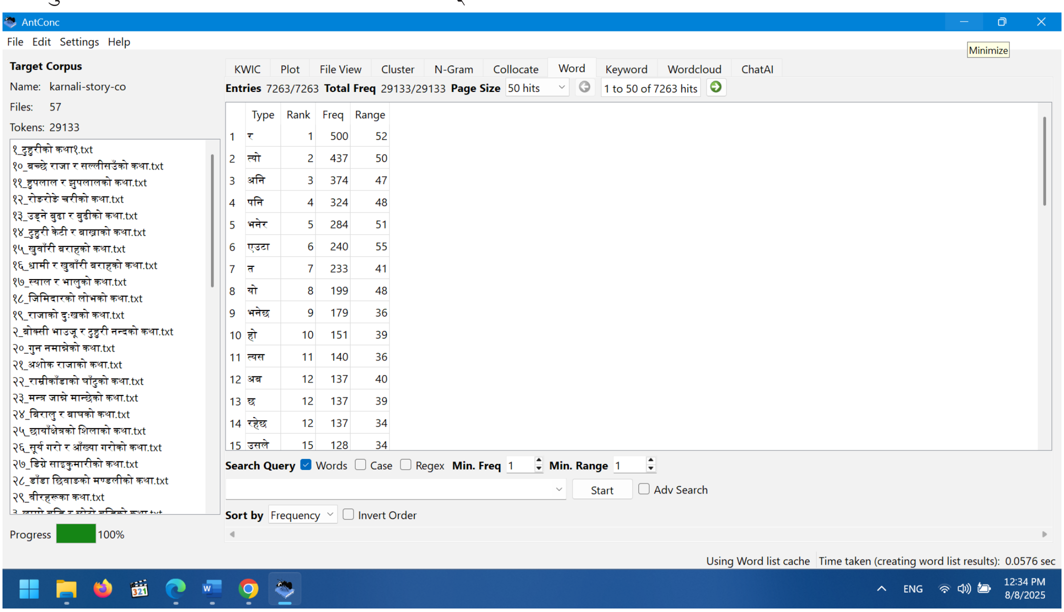
उड्ने बुढा र बुढीको कथा	४०३	२२१	५४.८४%
चिरबिरेले डाँफेलाई राजा स्वीकारे	४३१	२८६	६६.३६%
अभागी बहिनीको कथा	४४१	२२६	५१.२५%
बोक्सी भाउजू र टुहुरी नन्दको कथा	४४४	२६०	५८.५६%
आफ्नै कर्म खाने छोरोको कथा	४४८	२४१	५३.७९%
राजाको कथा	४६३	२८९	६२.४२%
अल्छी गोठालाको कथा	४७३	२५२	५३.२८%
अशोक राजाको कथा	४९७	३०८	६१.९७%
बुद्धिमान् केटाको कथा	५३३	३१२	५८.५४%
किचकन्नीको कथा	५३३	३६८	६९.०४%
भावीले लेखेको भाग्यको कथा	५३६	३१५	५८.७७%
भुल्टुङ चराको कथा	५४०	३१०	५७.४१%
किचकन्नीको मृत्यु	५६०	३६६	६५.३६%
टुहुरीको कथा१	६००	३२७	५४.५०%
केटी र मुरलीको कथा	६३०	३७४	५९.३७%
राजकुमारीको कथा	६४५	३५६	५५.१९%
मुसो, बिरालो र कुकुरको कथा	६८८	३६४	५२.९१%
इच्छामणिको कथा	६९२	३६८	५३.१८%
लछलाकृति पैकेलाको कथा	७०५	४५२	६४.११%
टुहुरी केटी र बाख्राको कथा	७२१	३६५	५०.६२%
राम्रो कामको राम्रो फलको कथा	७२६	४१०	५६.४७%
मुसेटोकोको कथा	७७३	४५५	५८.८६%
टुहुरीको कथा२	९११	५०६	५५.५४%
लोखर्केको कथा	९२०	४८८	५३.०४%
भावीको लेखा मेटिँदैन	९४२	४९०	५२.०२%
लिखेटीको कथा	१२७७	६८८	५३.८८%
भूतको कथा	१२९८	६८१	५२.४७%
बच्छे राजा र सल्लीसउँको कथा	१९१२	९९५	५२.०४%
दुई भाईको प्रेमको कथा	३२४५	१४२६	४३.९४%
जम्मा (५७ कथाहरू)	२९१३३	७२६३	४२.९३%

कथामा भएका शब्दसङ्ख्याले कथा कति लामो छ भन्ने देखाउँछ भने अद्वितीय शब्दसङ्ख्याले कथामा प्रयोग भएका भिन्न-भिन्न शब्दहरू जनाउँछ। शाब्दिक घनत्व प्रतिशतले वाचकले कथा भन्दा अर्थ सम्प्रेषण गर्न प्रयोग हुने शब्दहरूको स्थिति जनाउँछ। घनत्त्व बढी भएका कथाहरूमा अर्थ सम्प्रेषण हुँदा धेरै अद्वितीय शब्दहरूको प्रयोग भएको छ र दोहोरिने शब्दसङ्ख्या कम छ भन्ने बुझिन्छ भने घनत्त्व कम भएका कथाहरूमा सोही कार्यका लागि थोरै अद्वितीय शब्द प्रयोग भएका छन् र दोहरिने शब्दहरूको सङ्ख्या बढी छ भन्ने हो । माथि तालिकामा भएको तथ्याङ्कलाई हेर्दा छोटा कथाहरूमा बढी अद्वितीय शब्दहरू प्रयोग भएका छन् भने लामा कथाहरूमा दोहरिने शब्दसङ्ख्या बढी छ भन्ने देखिन्छ ।

माथि देखाइएका कथाहरूलाई कम्प्युटरमा प्रविष्ट गरेपछि एकमुस्ट रूपमा एउटा पाठसङ्ग्रहको रूपमा प्रयोग गर्न सकिन्छ, जसलाई चित्र २ मा देखाइएको छ।

चित्र २

कम्प्युटर सफ्टवेयर 'AntConc' मा पाठसङ्ग्रह



छलफल तथा विश्लेषण

कर्णाली क्षेत्रबाट सङ्कलन गरी विभिन्न प्रक्रिया पुरा गरी तयार भएको पाठसङ्ग्रहमा विभिन्न पक्षहरूको अध्ययन तथा विश्लेषण गर्न सकिन्छ, यहाँ विशेष गरी शब्द तथा आवृत्तिसँग सम्बन्धी के कस्ता विश्लेषण गर्न सकिन्छ भन्ने कुराको नमुना प्रस्तुति गरिएको छ।

सामान्य शब्द : शब्द भनेको भाषाको सबैभन्दा आधारभूत एकाइ हो। कुनै पाठमा प्रयोग भएका सम्पूर्ण शब्दहरूको गणना र वितरण संरचना अध्ययन यसको हिस्सा हो। कथामा प्रयोग भएका शब्दहरूको विविधता, दोहोरिने दर वा शब्दहरूको लम्बाइ आदिको विश्लेषणले भाषाको जटिलता, सरलता वा मौखिक परम्पराबाट आएका विशेषता बताउँछ। यो विश्लेषणले कुन शब्द कति प्रमुख छ भन्ने पनि देखाउँछ।

विशेष शब्द : विशेष शब्द भनेको कुनै विशिष्ट पाठमा अन्य समान प्रकारका पाठहरूमा भन्दा बढी प्रबलताका साथ देखा पर्ने शब्दहरू हुन्। उदाहरणका लागि कर्णालीका कथामा "जाँतो", "मागी", "जाडो", "छाक" आदि शब्दहरू यदि प्रायः देखा परे भने ती विशेष शब्द हुन सक्छन्। ती शब्दहरूले कर्णालीको भौगोलिक, आर्थिक र जीवन-शैलीको संकेत गर्छन्। यस्ता विशेष शब्दहरूको विश्लेषणले कर्णालीको भाषिक पहिचान छुट्याउन सहयोग पुऱ्याउँछ।

उच्च आवृत्ति भएका शब्दहरू

लोककथाहरूको पाठसङ्ग्रहमा भएका उच्च आवृत्ति भएका ५० ओटा शब्दहरूको सूची तालिका २ मा देखाइएको छ।

तालिका २

उच्च आवृत्ति भएका शब्दहरू

शब्द	Rank	Freq	Range
र	१	५००	५२
त्यो	२	४३७	५०
अनि	३	३७४	४७
पनि	४	३२४	४८
भनेर	५	२८४	५१
एउटा	६	२४०	५५
त	७	२३३	४१
यो	८	१९९	४८
भनेछ	९	१७९	३६
हो	१०	१५१	३९
त्यस	११	१४०	३६
अब	१२	१३७	४०
छ	१२	१३७	३९
रहेछ	१२	१३७	३४
उसले	१४	१२८	३४
म	१६	१२७	३७
कुरा	१७	१२६	३८

सबै	१८	१२४	३६
के	१९	१२३	४२
घर	२०	११९	३४
भने	२१	११३	३९
रे	२२	११०	१९
केही	२३	९५	३५
राजाले	२४	९०	१७
ऊ	२५	८९	३०
गएर	२५	८९	३७
भन्ने	२७	८८	४०
राजा	२७	८८	१८
भनी	२९	८६	२७
पानी	३०	८५	१८
भनिछ	३१	८४	२३
त्यहाँ	३२	८१	३३
त्यही	३२	८१	३२
थियो	३४	८०	२९
गएछ	३५	७६	३२
मलाई	३५	७६	२७
लिएर	३५	७६	२५
घरमा	३८	७४	२९
मेरो	३८	७४	२१
खान	४०	७३	२३
गर्न	४१	७१	२७
तर	४१	७१	२७
दिन	४३	७०	२९
छोरी	४४	६८	१५
आफ्नो	४५	६७	२५
चैँ	४५	६७	९
भएछ	४७	६५	२६
उसलाई	४८	६४	२२
रैछ	४८	६४	१६
उसको	५०	६१	२४

सन्दर्भ शब्द

सन्दर्भ शब्द भन्नाले कुनै शब्द 'भूत' पाठसङ्ग्रहमा के कस्तो वातावरणमा रहेको छ भन्ने देखाउँछ। उक्त शब्दभन्दा अगाडि अर्थात् बायाँ सन्दर्भ र पछाडि अर्थात् दायाँ सन्दर्भमा कसरी व्यवस्थित भएको छ भन्ने बुझिन्छ। यसलाई तालिका ३ मा देखाइएको छ।

तालिका ३

'भूत' शब्दको सन्दर्भ

बायाँ सन्दर्भ	खोज शब्द	दायाँ सन्दर्भ
किन फर्की आउनुभो त' सोद्धा धामीले 'तलको खुवाँरी खोलामा ठुलो	भूत	रहेछ । त्यसले मलाई ठ्याक्क बाटो छेकेर आफ्नो घर लग्यो ।
र सम्पत्ति भरिभराउ पारिदिएछ । दुबीचौरको महाजनको मामा रिठेचौरे पनि खतरनाक	भूत	रहेछ । त्यसले रुख ढालेर सल्लीसउँलाई बाटो छेकेछ । फागुने दादा
बस्ने बाइसपान्य राजा धेरै धनी रहेछ। उसको सहयोगी फागुनेदादा भन्ने	भूत	रहेछ । त्यसले सबै कुरामा सहयोग गर्दोरहेछ । बाइसपान्य राजाको छोरो
एकादेशमा एउटा	भूत	रहेछ । गाउँकी आइमाई दाउरा काट्न वनमा गएकी रहिछ । दाउरा
ठाउँमा फर्केछ । बाटोमा बाबुको सहयोगी फागुने दादालाई भेटेछ । त्यो भयङ्कर	भूत	रहेछ । पहिला त त्यस भूतले रोकेछ तर सबै बेलीबिस्तार
नतातेको तावा कालो थियो । अनि रिसले चुर भएर खान्छु भनी	भूत	आइपुगेछ । केटोले पहिले भूतलाई जदौ भनेछ । भूतले 'तँलाई खान
भूतले ऐरीको घारी फाँड्दासम्म दिदीभाइ घरमा पुगेछन् । घर पुगे पनि	भूत	आउँछ र खान्छ भन्ने पीर छँदै थियो । उनीहरू घरभित्र
आइमाई बच्चा पाउने भइछ । त्यो आइमाई पहिलोचोटि कर लागेपछि त्यो	भूत	आएछ । त्यो भूत बाहिर आएर 'अत्तानि कत्तानी, मेरी माग्यै
लाग्यो, अत्तानी कत्तानी चिनो गेडो भई' भनिछ । एक महिनापछि फेरि	भूत	आएर त्यसैगरी भनेछ र आमाचाहिँले भट्ट गेडी भई भनिछ ।
हात्ती आएछ । हात्ती रुखका हाँगा भाँच्दै घाँस खान थालेछ । राजाले	भूत	आयो भनी डराएछ र वाण हानेछ । अघिपछि हानेको वाण

शब्दगुच्छ

शब्दगुच्छको विश्लेषणले पाठमा प्रयोग भएका कुनै एक शब्दसँग अन्य शब्दहरूको संयोजन वा समूह (जस्तै दुई वा सोभन्दा बढी शब्दहरूको मिलन) पहिचान गर्छ । उदाहरणका लागि, 'अनि' शब्द अन्य कस्ता कस्ता शब्दहरूसँग आउँछ र तिनीहरूको आवृत्ति कस्तो छ भनि विश्लेषण गर्न सकिन्छ । यस्ता शब्दगुच्छहरूले समाजको संरचना, विश्वास प्रणाली र व्यवहारिक मान्यता केन्द्रीत रहेको देखाउँछ । यसले सांस्कृतिक ढाँचाको चित्रण गर्छ । 'अनि' शब्दसँग दुईओटा र तीनओटा शब्दगुच्छ कस्ता छन् र तिनीहरूको आवृत्ति कति छ भन्ने कुरा तालिका ४ र तालिका ५ मा राखिएको छ ।

तालिका ४

दुई शब्दको शब्दगुच्छा

शब्दगुच्छ	Rank	Freq	Range
अनि त्यो	१	२९	१२
अनि उसले	२	१८	१३
अनि त्यस	३	१०	१०
अनि ऊ	४	७	५
अनि बुढीले	४	७	३
अनि त्यसपछि	६	६	४
अनि त्यसले	६	६	६
अनि उनीहरूले	८	५	५
अनि राजाले	८	५	४
अनि उनीहरू	१०	४	४

तालिका ५

तीन शब्दको शब्दगुच्छ

शब्दगुच्छ	Rank	Freq	Range
अनि त्यस आइमाईले	१	३	३
अनि त्यो दमाई	१	३	१
अनि त्यसपछि गएर	३	३	२
अनि त्यो गरामा	३	२	१
अनि त्यो बाख्राको	३	२	१
अनि त्यो बाघ	३	२	१
अनि त्यो भूत	३	३	१
अनि भोजपातको लुगा	३	२	१
अनि लट्टी बजेन	३	२	१
अनि अन्तिममा एउटा	१०	१	१

शब्द शृङ्खला

शब्द शृङ्खला भन्नाले शब्दहरू सङ्ख्यात्मक रूपमा लगातार आउने शब्दहरूको समूह हो । द्विशब्द, त्रिशब्द आदि यसको उदाहरण हुन् । उदाहरणका लागि यदि कुनै पाठसङ्ग्रहमा भएका कथामा बारम्बार आउने द्विशब्द जस्तैः “बुढा बुढी”, “देउता आए”, “राजा बोले” को अध्ययनले सामान्य वाक्य विन्यास, पात्रहरूको भूमिका, वा घटना विकासको नियमितता देखाउन सक्छ । यस्ता शब्द शृङ्खलाहरू

शैलीगत विश्लेषणमा उपयोगी हुन्छन् । दुई शब्द शृङ्खला तथा तीन शब्द शृङ्खला तालिका ६ र तालिका ७ मा प्रस्तुत गरिएको छ ।

तालिका ६

दुई शब्द शृङ्खला

शब्दशृङ्खला	Rank	Freq	Range
भनेछ अनि	१	३४	२०
अनि त्यो	२	२९	१२
सबै कुरा	२	२९	१५
रे अनि	४	२७	१२
सुनको माला	५	२६	२६
सुन्नेलाई सुनको	५	२६	२६
जाला भन्ने	७	२५	२५
फूलको माला	७	२५	२५
भन्नेलाई फूलको	७	२५	२५
यो कथा	७	२५	२५

तालिका ७

तीन शब्द शृङ्खला

शब्द शृङ्खला	Rank	Freq	Range
सुन्नेलाई सुनको माला	१	२६	२६
भन्नेलाई फूलको माला	२	२५	२५
माला यो कथा	३	२४	२४
जाला भन्ने बेला	४	२३	२३
कथा वैकुण्ठ जाला	५	२०	२०
फूलको माला सुन्नेलाई	६	१९	१९
बेला टुपुक्क आइहाल्ला	६	१९	१९
भन्ने बेला टुपुक्क	६	१९	१९
माला सुन्नेलाई सुनको	६	१९	१९
यो कथा वैकुण्ठ	६	१९	१९

सहप्रयोगी शब्द

सहप्रयोगी शब्द भनेको कुनै शब्दसँग बारम्बार सहप्रयोगमा आउने अन्य शब्दहरू हुन् । उदाहरणका लागि “देउता” शब्दसँग “आए”, “बोले”, “पुजियो” जस्ता शब्दहरू आउने गर्छन् भने ती सहप्रयोगी

शब्दहरू हुन् । कथाका केन्द्रविन्दु शब्दहरू कुन अन्य शब्दहरूसँग बढी आउँछन् भन्ने विश्लेषणले सांस्कृतिक क्रिया, सम्बन्ध, र कार्य-सञ्चालन बुझ्न मद्दत गर्छ । यसले सामाजिक संरचना, धर्मिक विश्वास वा भावनात्मक तात्त्विक पक्षको व्याख्या दिन सक्छ । यस पाठसङ्ग्रहमा पाइने 'घाँटु' शब्दभन्दा अगाडि वा पछाडिको वातावरणमा के कस्ता अन्य शब्दहरू बढी देखापर्छन् र आवृत्ति कति भन्ने कुरा तालिका ८ प्रस्तुत गरिएको छ ।

तालिका ८

'घाँटु'को सहप्रयोगी शब्दहरू

सहप्रयोगी	Rank	FreqLR	FreqL	FreqR	Range
सुल्ट्याउन	१	३	०	३	१
नाच्ने	२	३	१	२	१
गीत	२	३	०	३	१
थालेछ	४	३	१	२	१
चलन	५	२	१	१	१
सोरठी	६	१	१	०	१
देउराली	६	१	१	०	१
ढुङ्गालाई	६	१	१	०	१
घर्काएछ	६	१	१	०	१
घर्काएपछि	६	१	१	०	१

निष्कर्ष

कर्णाली क्षेत्रमा मौलिक तथा मौखिक रूपमा प्रचलनमा रहेका लोककथालाई उद्देश्यमूलक तथा परिस्थितिजन्य विधिको प्रयोग गरी ५७ ओटा श्रव्यरूपको रेकर्डबाट प्राप्त सामग्रीलाई देवनागरी युनिकोडमा लिपीकरण गरी कम्प्युटरले पढ्न सक्ने र प्रयोगयोग्य बनाइ एउटा लोककथाहरूको पाठसङ्ग्रह तयार भएको छ । यो पाठसङ्ग्रलाई कम्प्युटर सफ्टवेयरको प्रयोग गरी ती कथाहरूमा के कस्ता शब्दहरू छन्, तिनीहरूसम्बन्धी विभिन्न तथ्याङ्कीय विवरण कसरी हेर्न र विश्लेषण गर्न सकिन्छ भन्ने बारे वर्णन गरिएको छ । शब्दसङ्ख्या, शब्दगुच्छ, शब्दशृङ्खला, सहप्रयोगी शब्द आदिको सहायताले कथाहरूलाई पाठसङ्कलनभाषावैज्ञानिक विधिबाट भाषाको प्रयोगशैली, सामाजिक अर्थ, सांस्कृतिक ढाँचा, पदावली तथा वाक्य संरचना साथै अर्थ संरचनाबारे महत्त्वपूर्ण अन्तर्दृष्टि तथा अन्तर्वितरण गर्न सकिन्छ भन्ने बाटो खुलेको छ । यस्ता विश्लेषणले कर्णाली क्षेत्रको मौलिक परम्परा, लोकमान्यता र सामाजिक आस्था भाषिक तहबाट उजागर गर्न सकिन्छ । साथै, यी विधिहरू अङ्कीय भाषिक स्रोत निर्माण, भाषा प्रविधिमा स्थानीयकरण र विभिन्न भाषाहरूको संसाधन विकासमा पनि उपयोगी सावित बन्न सक्ने देखिन्छ ।

कृतज्ञता

यो अनुसन्धान कार्यपत्र अनुसन्धान निर्देशनालय, त्रिविबाट प्राप्त अनुसन्धान परियोजना शीर्षक "कर्णाली प्रदेशमा प्रचलित लोककथामा संस्कृति" कोड नं. TU–NPAR–०७९/८०–EFRG–१२ को सहयोगमा तयार भएको हो र यस सन्दर्भमा हामी त्रिवि अनुसन्धान निर्देशनालयलाई हार्दिक धन्यवाद ज्ञापन गर्दछौँ ।

सन्दर्भसूची

बन्धु, चूडामणि (२०५८). *नेपाली लोकसाहित्य.* काठमाण्डौं : एकता बुक्स ।

राकेश, रामदयाल (२००५). *नेपाली लोककथा विश्लेषण.* काठमाडौँ : साहित्य संगम ।

लामिछाने, शंकर (२०१८). *नेपाली लोककथाको स्वरूप र संरचना.* काठमाडौँ : नेपाल प्रज्ञा प्रतिष्ठान ।

Anthony, L . (2024). AntConc (Version 4 .3 .1) [Computer Software] . Tokyo, Japan: Waseda University . https://www .laurenceanthony .net/software/AntConc