

Comparison of various writing characteristics of Hindi and Marathi languages of Devanagari origin

Priyanka Sorate,^{*} Navjot Kaur Kanwal^{**} and Manish Malhotra^{***}

^{*}Student- M.Sc. Forensic Science, Department of Criminology & Forensic Science, Dr. Harisingh Gour University (Central University), Sagar-MP 470003, India

^{**}Corresponding author, Assistant Professor Department of Criminology & Forensic Science, Dr. Harisingh Gour University (Central University), Sagar-MP 470003, India

^{***}Research Scholar, JRF-UGC, Department of Criminology & Forensic Science, Dr. Harisingh Gour University (Central University), Sagar-MP 470003, India

Abstract

The examination of handwriting analysis has always been a challenging task and it becomes more challenging in India due to various languages and scripts with different styles of writing. The Hindi language which is one of the official languages of India is written in Devanagari script. Some other languages also written in Devanagari script and resemble with each other. The analysis of these languages which belongs to the same script is problematic. An attempt has been made in this paper to compare handwritings of two different languages of same script written by same persons. A total of 50 samples were taken for present study. Manual examination is done to compare the class and individual characteristics in the handwriting of same scripts but different language. This study is helpful in understanding the basic concepts of handwriting analysis and also helpful in determining the authorship of the handwriting if written in different language of same script.

Keywords: Devanagari script, handwriting, examination, Hindi, Marathi languages.

1. Introduction

1.1. Devanagari script

India is a multi-lingual as well as multi-script country and there are twenty two languages. Eleven scripts are used to write these languages and Devanagari is the most widespread script in India. First research report on handwritten Devanagari characters was published in 1977 (Sethi et al; 1997). In India, around 300 million people use Devanagari script for writing languages like Hindi, Marathi, Sanskrit, Nepali, Sindhi, and Konkani, where Hindi is the official language of the country (Jayadevan et al; 2011). Hindi is also third most popular language in the world (Pal et al; 2004). Hindi and Marathi are the most popular languages written in Devanagari script. As the national language, Hindi is accepted all over India and is used for documentation particularly in the Indian states of Bihar, Chhattisgarh, Haryana, Himachal Pradesh, Jharkhand, Madhya Pradesh, New Delhi, Rajasthan, Uttar Pradesh and Uttarakhand. Marathi language is the official language of the Indian state of Maharashtra, which is one of the largest states in the country. To fill-up several paper documents like bank cheques, envelopes, application forms, railway reservation forms, answer sheets etc. people use Devanagari script. Devanagari initiated from ancient Brahmi script through a number of alterations (Pal et al;

2004). The script is written from left to right and has 13 vowels, 34 consonants and 14 modifiers. [Fig.2] In Devanagari script a vowel following a consonant takes a modified shape. Depending on the vowel, its modified shape is located at the left, right (or both) or bottom of the consonant. These modified shapes are called modified characters. A consonant or vowel following a consonant sometimes takes a compound orthographic shape, which we call as compound character. (Pal et al; 2004)

Every Indian script has its own specified composition rules for combining vowels, consonants, and modifiers. Some of them can be combined with their type, as shown in Fig.3; a modifier can be associated with vowel or with consonant. Consonants may have a half form when they are combined with other consonants as depicted in [Fig.3]. Excluding for some characters, the half forms of consonants are the left part of original consonants with the right part detached. Some special combinations are also shown in Fig. 2; where a new character or the half forms of consonants may look like in the lower half of the new composite forms. Another individual feature of Devanagari is the existence of a horizontal line on the top of all characters. This line is known as header line or “shirorekha” shown in Fig 1.



Fig. 1: Shirorekha in Devanagari script

अ	आ	इ	ई	क	ख	ग	घ	ङ
उ	ऊ	ऋ	ॠ	च	छ	ज	झ	ञ
ए	ओ	औ	अं	ट	ठ	ड	ढ	ण
अः				त	थ	द	ध	न
				प	फ	ब	भ	म
				य	र	ल	व	श
				ऌ	ॡ	ह	ळ	ॠ
				ॡ	ॢ	ॣ	।	॥
				॥	०	१	२	३
				४	५	६	७	८
				९	१०	११	१२	१३
				१४	१५	१६	१७	१८
				१९	२०	२१	२२	२३
				२४	२५	२६	२७	२८
				२९	३०	३१	३२	३३
				३४	३५	३६	३७	३८
				३९	४०	४१	४२	४३
				४४	४५	४६	४७	४८
				४९	५०	५१	५२	५३
				५४	५५	५६	५७	५८
				५९	६०	६१	६२	६३
				६४	६५	६६	६७	६८
				६९	७०	७१	७२	७३
				७४	७५	७६	७७	७८
				७९	८०	८१	८२	८३
				८४	८५	८६	८७	८८
				८९	९०	९१	९२	९३
				९४	९५	९६	९७	९८
				९९	१००	१०१	१०२	१०३
				१०४	१०५	१०६	१०७	१०८
				१०९	११०	१११	११२	११३
				११४	११५	११६	११७	११८
				११९	१२०	१२१	१२२	१२३
				१२४	१२५	१२६	१२७	१२८
				१२९	१३०	१३१	१३२	१३३
				१३४	१३५	१३६	१३७	१३८
				१३९	१४०	१४१	१४२	१४३
				१४४	१४५	१४६	१४७	१४८
				१४९	१५०	१५१	१५२	१५३
				१५४	१५५	१५६	१५७	१५८
				१५९	१६०	१६१	१६२	१६३
				१६४	१६५	१६६	१६७	१६८
				१६९	१७०	१७१	१७२	१७३
				१७४	१७५	१७६	१७७	१७८
				१७९	१८०	१८१	१८२	१८३
				१८४	१८५	१८६	१८७	१८८
				१८९	१९०	१९१	१९२	१९३
				१९४	१९५	१९६	१९७	१९८
				१९९	२००	२०१	२०२	२०३
				२०४	२०५	२०६	२०७	२०८
				२०९	२१०	२११	२१२	२१३
				२१४	२१५	२१६	२१७	२१८
				२१९	२२०	२२१	२२२	२२३
				२२४	२२५	२२६	२२७	२२८
				२२९	२३०	२३१	२३२	२३३
				२३४	२३५	२३६	२३७	२३८
				२३९	२४०	२४१	२४२	२४३
				२४४	२४५	२४६	२४७	२४८
				२४९	२५०	२५१	२५२	२५३
				२५४	२५५	२५६	२५७	२५८
				२५९	२६०	२६१	२६२	२६३
				२६४	२६५	२६६	२६७	२६८
				२६९	२७०	२७१	२७२	२७३
				२७४	२७५	२७६	२७७	२७८
				२७९	२८०	२८१	२८२	२८३
				२८४	२८५	२८६	२८७	२८८
				२८९	२९०	२९१	२९२	२९३
				२९४	२९५	२९६	२९७	२९८
				२९९	३००	३०१	३०२	३०३
				३०४	३०५	३०६	३०७	३०८
				३०९	३१०	३११	३१२	३१३
				३१४	३१५	३१६	३१७	३१८
				३१९	३२०	३२१	३२२	३२३
				३२४	३२५	३२६	३२७	३२८
				३२९	३३०	३३१	३३२	३३३
				३३४	३३५	३३६	३३७	३३८
				३३९	३४०	३४१	३४२	३४३
				३४४	३४५	३४६	३४७	३४८
				३४९	३५०	३५१	३५२	३५३
				३५४	३५५	३५६	३५७	३५८
				३५९	३६०	३६१	३६२	३६३
				३६४	३६५	३६६	३६७	३६८
				३६९	३७०	३७१	३७२	३७३
				३७४	३७५	३७६	३७७	३७८
				३७९	३८०	३८१	३८२	३८३
				३८४	३८५	३८६	३८७	३८८
				३८९	३९०	३९१	३९२	३९३
				३९४	३९५	३९६	३९७	३९८
				३९९	४००	४०१	४०२	४०३
				४०४	४०५	४०६	४०७	४०८
				४०९	४१०	४११	४१२	४१३
				४१४	४१५	४१६	४१७	४१८
				४१९	४२०	४२१	४२२	४२३
				४२४	४२५	४२६	४२७	४२८
				४२९	४३०	४३१	४३२	४३३
				४३४	४३५	४३६	४३७	४३८
				४३९	४४०	४४१	४४२	४४३
				४४४	४४५	४४६	४४७	४४८
				४४९	४५०	४५१	४५२	४५३
				४५४	४५५	४५६	४५७	४५८
				४५९	४६०	४६१	४६२	४६३
				४६४	४६५	४६६	४६७	४६८
				४६९	४७०	४७१	४७२	४७३
				४७४	४७५	४७६	४७७	४७८
				४७९	४८०	४८१	४८२	४८३
				४८४	४८५	४८६	४८७	४८८
				४८९	४९०	४९१	४९२	४९३
				४९४	४९५	४९६	४९७	४९८
				४९९	५००	५०१	५०२	५०३
				५०४	५०५	५०६	५०७	५०८
				५०९	५१०	५११	५१२	५१३
				५१४	५१५	५१६	५१७	५१८
				५१९	५२०	५२१	५२२	५२३
				५२४	५२५	५२६	५२७	५२८
				५२९	५३०	५३१	५३२	५३३
				५३४	५३५	५३६	५३७	५३८
				५३९	५४०	५४१	५४२	५४३
				५४४	५४५	५४६	५४७	५४८
				५४९	५५०	५५१	५५२	५५३
				५५४	५५५	५५६	५५७	५५८
				५५९	५६०	५६१	५६२	५६३
				५६४	५६५	५६६	५६७	५६८
				५६९	५७०	५७१	५७२	५७३
				५७४	५७५	५७६	५७७	५७८
				५७९	५८०	५८१	५८२	५८३
				५८४	५८५	५८६	५८७	५८८
				५८९	५९०	५९१	५९२	५९३
				५९४	५९५	५९६	५९७	५९८
				५९९	६००	६०१	६०२	६०३
				६०४	६०५	६०६	६०७	६०८
				६०९	६१०	६११	६१२	६१३
				६१४	६१५	६१६	६१७	६१८
				६१९	६२०	६२१	६२२	६२३
				६२४	६२५	६२६	६२७	६२८
				६२९	६३०	६३१	६३२	६३३
				६३४	६३५	६३६	६३७	६३८
				६३९	६४०	६४१	६४२	६४३
				६४४	६४५	६४६	६४७	६४८

क + क = क्क	ल + ल = ल्ल
घ + न = घ्न	श + न = श्न
ब + व = ब्व	त + न = त्न
म + ल = म्ल	प + ल = प्ल
Combinations	
क + ष = क्ष	ज + झ = झ
द + व = द्व	ट + ट = ट्ट
श + र = श्र	ठ + ठ = ठ्ठ
त + र = त्र	द + द = द्द
द + य = द्य	द + ध = द्ध
Special combinations	

Fig. 3: Consonants and their corresponding half forms in Devanagari script

1.2. Key Features of Hindi Morphology

Hindi is spoken by nearly 500 million people (non-native included) and ranks third among the world's most spoken languages (Chinese is the first; English is second while Hindi is third language). The term "Hindi language" however does not mention to a well-defined and clearly consistent language but rather to a fairly large group of dialects where in inter lingual understanding is always possible (just as English is in the UK and the U.S.). Hindi is written using the Devanagari script, including 11 vowels and 33 Simple consonants and also nasal symbols such as anusvar and anurasik, plus a symbol for the weak aspiration visgar (although very rare in this language). Generally no dissimilarity is made between uppercase and lowercase letters. In Hindi grammar there are only two genders, masculine and Feminine, while the neuter found in Sanskrit has extinct. Feminine nouns are typically formed from the masculine, either by interchanging the final 'आ' by 'ई' (e.g., "घोड़ा" (horse), "घोड़ी" (mare) or by adding 'ई' for nouns ending with a consonant "बंदर" (monkey), "बंदरिया" (female monkey). Number is stated through distinctive singular and plural forms. This language does not have a definite article (the), and as a substitute of placing prepositions before the noun, it positions them after in the form of post positions (e.g., "on the table "→"table on"). These are used in definite Western European Languages such as German, as in the expression "den Fluss entlang" (along the river), while the use of this linguistic construction in other Indo-European languages is obviously the omission. Nouns and adjectives also have two distinct grammatical cases, direct and Oblique (Dolamic et al; 2010).

1.3. Key Features of Marathi Morphology

Marathi is spoken in western India by about 70 million people, and thus ranks fourth among the languages spoken there. As in other languages it may comprise various dialects along with certain spelling and phonological variations. Marathi is written in the Devanagari script as well as another variant, the Balbodh script. Marathi contains 52 letters, of which only 50 represent diverse sounds. These sounds are stated by 14 vowels having different initial leading forms and also different shapes when following consonants. There are 36 consonants in all, together with two compound consonants as

well as some nasal symbols. As in Sanskrit, Marathi nouns may have three possible genders (masculine, feminine, and neutral) and be either singular or plural in number. Masculine, feminine, or neutral noun forms are resulting through applying systematic and simple rules (for example, a child "मुलगा" (masculine), "मुलगी" (feminine), "मुलग" (neutral); or for a dog "कुत्रा" (masculine), "कुत्री" (feminine), "कुत्र," (neutral)). As in other languages there are certain exceptions, such as the noun "camel" which has two distinct forms ("उंट" (masculine), "मांड" (feminine)(Dolamic et al; 2012).

The analysis of Devanagari script is not a simple job because the style of writing varies greatly through the different regions of India. Identification of a handwriting and classification of handwriting characteristics need a definite set of parameters (Harrison; 1978, Osborn; 1929, Bates; 1970). Each division and subdivision involving the characteristics is first assessed and tabulated and finally defined clearly. The first divisions of handwriting components are the general elements such as style, slant, size, proportion and alignment. These are sub-divided according to the qualities they possess; for example, handwriting may be slanted to the left or to the right, or be perpendicular the line of writing. In practical examinations these qualities are further sub-divided into secondary divisions, defined by measurements which convey their physical attributes. Analysis of a set of handwriting samples thus starts as a class characterization, where a number of writers found to possess similar characteristics. On deeper analysis, each handwriting is found to have its own individual characteristics, making it possible for an examiner to identify the writer (Saxena et al; 1992).

2. Materials

- Writing sample
- Cam scanner
- Magnifying glass
- Pen
- Pencil
- Scale
- Light source

3. Methodology

- The handwritten samples in Hindi and Marathi were collected from same writer.
- A total of 50 samples were collected, each sample enclosed Hindi and Marathi handwritings from same writer.
- The 50 writers knowing the 2 language were carefully chosen to make available samples in their own language for analysis. They were given a standard manuscript in Hindi with a request to translate and write it in their own language i.e. Marathi.
- The content of this standard text diverse from language to language but were same of a particular language.
- The so collected sample were then scrutinized and evaluated by applying the principle of handwriting examination of their overall pictorial appearance, general writing characteristics, individual writing characteristics, similarities and closeness between two languages, the commonest form of letter, formation of

letter, combination of letters, start and finish, relative location of stroke and their combination, direction and curvature of stroke, formation of loop, its size, shape and inclination, uniformity and irregularity of the height of the letters in words.

4. Result and Discussions

Every sample of Hindi and Marathi languages were studied separately to find out its individual and class characteristics. Each sample was carefully examined and each letter and word was compared. Earlier studies were based on Devanagari character recognition by using various computational approaches. The study displays Devanagari character recognition system and the Devanagari script comparison with other script. In this paper we emphasized on resembling Marathi and Hindi which belongs to Devanagari script for the purpose of comparison.

4.1. Class characteristic

In Devanagari handwriting samples we examined class characteristics. Class characteristic are group characteristics. In case of Marathi and hind handwriting sample appearances almost same class characteristics. The characteristics we talk over here show correspondences in both handwriting. The general writing characteristics of the handwriting sample were verified in form of observation chart as shown in Table 1.

4.2. Individual characteristics

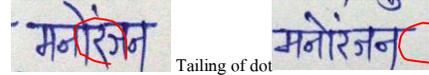
In the samples we considered 4 individual characteristics placement and formation of modifiers, spacing between words, shape and placement of dot, formation of loops and placement of letters. The sample revealed some individual writing habit of the writer which had same characteristic and displayed resemblances among both (Hindi and Marathi) handwriting.

In the establishment of modifier one stroke is vertical line and other stroke is rounded stroke above the shirorekha and the matra form small slant rightward line and these formation seen in both sample.

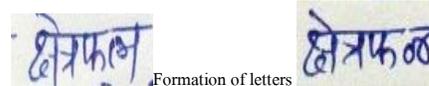


Spacing between words, letter and line to line were analogous in both.

Simple spherical dot with tailing above shirorekha with no loop formation.



Establishment of letter indicated one vertical stroke and other horizontal 'S' shape creation. Other letters also indicated analogous formation and placements.



Hindi Marathi

Sr. no.	General characteristics	Observation (Hindi)	Observation (Marathi)
1	Alignment	Along the base line	Along the base line
2	Shape	Moderate	Moderate
3	Skill	low	low
4	Slant	leftward	leftward
5	Movement	wrist	wrist

Table 1: General characteristics

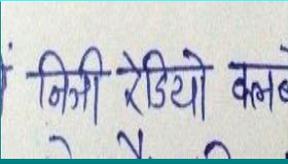
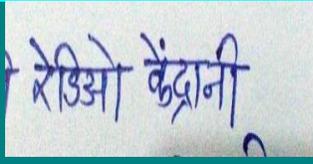
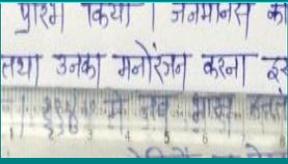
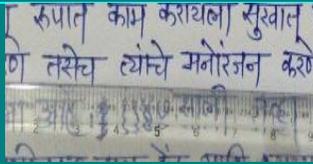
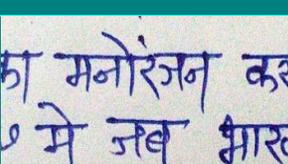
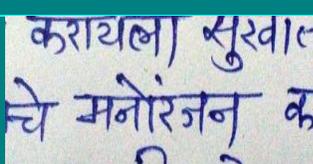
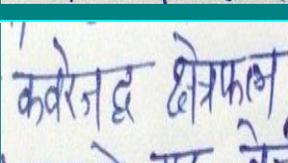
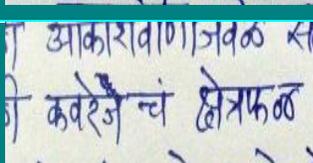
Sr. No.	Individual Characteristics	Hindi Sample	Marathi Sample
1	Placement and formation of modifiers.		
2	Spacing between words		
3	Shape and placement of dot.		
4	Formation of loop in letters, arrangement of letters and their formation		

Table 2: Individual characteristics

5. Conclusions:

Through the study undertaken on examination of the 50 handwritten samples, we have reached to the conclusion that if the principles of handwriting examination are carefully followed and applied by a forensic document examiner, It is very much possible for a document examiner to undertake the examination of a disputed document consisting of writings in Marathi language and to render opinion thereon because the Marathi language shows similarities with Hindi language and both are originated from Devanagari script.

In this study, the examination of handwritten samples of 50 writers was carried out by analyzing their general and individual writing characteristics for arriving at a definite opinion regarding their authorship. The same handwriting of same person in two different languages was (Marathi and Hindi) studied. Both languages are different but having same consonants and vowel therefore shows similarity and is comparable.

Such handwritings are comparable on the basis of placement of modifiers, spacing between words, placement of punctuation marks, formation of letters, formation of loop and respective size.

References:

1. I.K. Sethi and B. Chatterjee, "Machine Recognition of constrained Hand printed Devnagari", Pattern Recognition, Vol. 9, pp. 69-75, 1977.
2. R. Jayadevan, S. R. Kolhe, P.M. Patil and U. Pal, "Offline Recognition of Devanagari Script: A Survey", IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews, Vol.41, pp. 782-796, 2011.
3. U. Pal and B. B. Chaudhuri, "Indian script character recognition: A Survey", Pattern Recognition, Vol. 37, pp. 887-1899, 2004.
4. Dolamic, L. and Savoy, J. 2010. Comparative study of indexing and search strategies for the Hindi, Marathi, and Bengali languages. ACM Trans. Asian Lang. Inform. Process. 9, 3, Article 11, Vol.9, pp. 11-24, 2010.
5. Harrison WR. "Suspect Documents. London: Sweet & Maxwell, 1978.
6. Bates BP. "Identification Systems for Questioned Documents". Springfield, Illinois: Charles C Thomas Publishers, 1970.
7. Brewster, F. "Contested documents & forgeries". The book company ltd., 1932.

8. HM Saxena and mahipal Singh, "Classification of the writing elements in Devnagari script", Central Forensic Science Laboratory, Bureau of Police Research and Development, MHA(Government of India)Journal of the Forensic Science Society,Vol.32, pp.143-150, 1992.
9. Ellen D. Handwriting Examination of Unfamiliar Scripts. Proceedings of the Sixth International Meeting of Forensic Science, 1972.
10. Hilton O., "Scientific examination of questioned documents". Revised Ed., CRC press, 1993.
11. Kelly, J. S., Lindblom, B. S. Scientific examination of questioned documents. 2nd ed., CRC press, 2006.
12. Huber, R. A., Headrick, A. M. "Handwriting identification: facts and fundamentals", CRC Press, 1999.
13. Snehal S. Patwardhan and R.R. Deshmukh, "A Review on offline handwritten Recognition of Devnagari script", International Journal of Computer Applications, Vol. 117, pp. 7-10, May 2015.
14. Osborn, A. S. "Questioned documents". 2nd Ed., Boyd printing company 1929.

About the authors

1. Priyanka Sorate- Student, Post-graduation (Forensic Science) from Department of Criminology and Forensic Science, Dr. Harisingh Gour Central University, Sagar, Madhya Pradesh, India.
2. Dr. Navjot Kaur Kanwal- Assistant Professor, Department of Criminology and Forensic Science, Dr. Harisingh Gour Central University, Sagar, Madhya Pradesh, India. Dr. Navjot Kaur Kanwal, M.Sc, Ph.D (Forensic Science), PGDCA, M.Sc.-IT, holds a vast teaching, administrative and research experience in field of Forensic science. Her field of interest are Questioned documents and Fingerprints, Cyber Forensics, Photo Forensics and Narcotic drugs. She has guided many Master's degree students and PhDstudents for their dissertation and thesis work. She has also published one book and presented/published many research papers in various National and International conferences and has organized many conferences/symposium for the development of the subject.
3. Manish Malhotra- Research Scholar (UGC NET-JRF) from Department of Criminology and Forensic Science, Dr. Harisingh Gour Central University, Sagar, Madhya Pradesh, India has been actively working in the field of photoanthrometry/photogrammetry.