



How Handwriting Evolves

An Initial Quantitative Analysis of Development of Indic Scripts



University of
St Andrews

FOUNDED
1413

vinodh Rajan vrs3@st-andrews.ac.uk

PhD Student, Computer Science
University of St Andrews, Scotland

sicsa* The Scottish Informatics & Computer Science Alliance



Indic Scripts

- ⦿ There are currently a myriad of Indic scripts – Devanagari, Tamil, Malayalam, etc.
- ⦿ They all evolved from Brahmi, which is the original source script
- ⦿ Various scripts emerged due to accumulation of handwriting variations (in a span of two millennia)



University of
St Andrews

FOUNDED
1413

Devanagari

अ	आ	इ	ई	उ	ऊ
a	ā	i	ī	u	ū
ऋ	ॠ	ऌ	ৡ		
ए	ऐ	ଓ	ঔ	অং	অঃ
ক	খ	গ	ঘ	ঢ়	
কা	খা	গা	ঘা	ঢ়া	
চ	ছ	জ	ঝ	ণ	
ট	ঠ	ঢ	ঢ়	ণ	
ত	থ	দ	ধ		
প	ফ	ব	ভ		
য	ৰ	ল	ৱ		
শ	ষ	স	হ		
শ	ষ	স	হ		

Brahmi

ક	ક̄	િ	િ	ઉ	ઉ
એ	એ	ઓ	ઓ		
કા	કા	ગા	ગા		
દ	દ	જા	જા		
ચ	ચ	ફા	ફા		
ત	ત	રા	રા		
પ	પ	લા	લા		
ય	ય	ષા	ષા		
શ	શ	સા	સા		
শ	ষ	হ	হ		



University
of
St Andrews

FOUNDED
1413

Grantha

କ	କୁ	ତ	ତୁ	ଏ
କା	କା	ଗ	ଗା	ନା
ଚା	ଚା	ଜା	ଜା	ନା
ତା	ତା	ଧା	ଧା	ନା
ପା	ପା	ବା	ବା	ନା
ଯା	ଯା	ଲା	ଲା	ନା
ଶା	ଶା	ଷା	ଷା	ନା
ହା	ହା	ହା	ହା	ନା

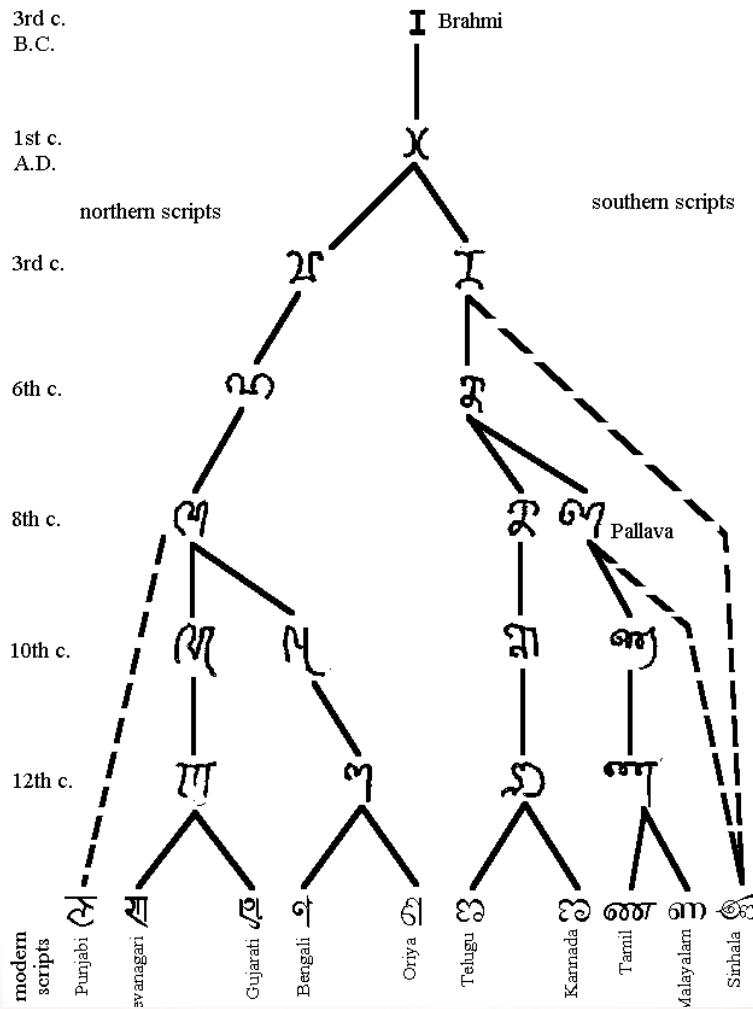
Kannada

ಅ	ಆ	ઇ	ಇ	ಉ	ಊ	ಈ	ಔ	ಏ	ಏಂ	ಏಃ
ಏ	ಏ	ಎ	ಎ							
ಕಾ	ಕಾ	ಗಾ	ಗಾ							
ಚಾ	ಚಾ	ಜಾ	ಜಾ							
ತಾ	ತಾ	ಧಾ	ಧಾ							
ಪಾ	ಪಾ	ಬಾ	ಬಾ							
ಯಾ	ಯಾ	ಲಾ	ಲಾ							
ಶಾ	ಶಾ	ಷಾ	ಷಾ							
ಹಾ	ಹಾ	ಹಾ	ಹಾ							
ಇ	ಇ	ಒ	ಒ	ಒ	ಒ	ಒ	ಒ	ಒ	ಒ	ಒ
ಉ	ಉ	ಔ	ಔ	ಔ	ಔ	ಔ	ಔ	ಔ	ಔ	ಔ
ಈ	ಈ	ಏ	ಏ	ಏ	ಏ	ಏ	ಏ	ಏ	ಏ	ಏ
ಔ	ಔ	ಏಂ								
ಏಃ										

sicsa*

The Scottish Informatics & Computer Science Alliance

Evolution of Indic Scripts



University of
St Andrews

FOUNDED

1413



Evolution of Indic Scripts

- ② unique opportunity to analyze script developments through handwriting behavior.
- ② investigate how different handwriting features have evolved
- ② Enables us to understand the variations occurring in handwriting



University of
St Andrews

FOUNDED
1413



Paleographic Dataset

- ⌚ Four major modern scripts were selected – Devanagari, Kannada, Grantha & Tamil
- ⌚ Their development was divided into ~6 distinct stages covering two millennia
- ⌚ The dataset was normalized for various factors
- ⌚ Totally ~730 distinct glyphs were included in the dataset



University of
St Andrews

FOUNDED
1413

Paleographic Dataset

- Evolution of character "KA" in various stages in different scripts

+ ತ ತ್ ತಂ ತ್ರ ತ್ರಂ + த் துக்க

Kannada

Tamil

+ ॐ ಹೃಹೃಹೃ + ತ್ತುಮ್ಮಾಮ್ಮಂ

Devanagari

Grantha

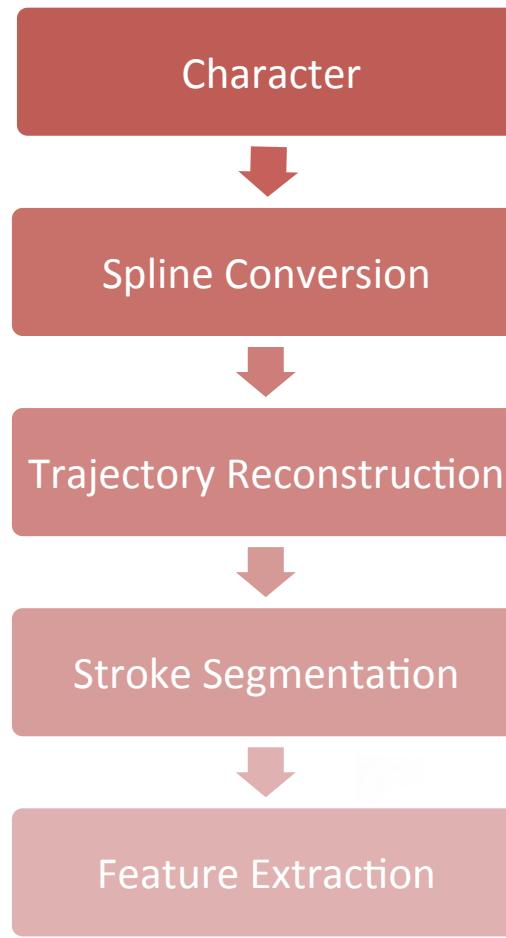


University of
St Andrews

FOUNDED
1413



Script Analysis Framework

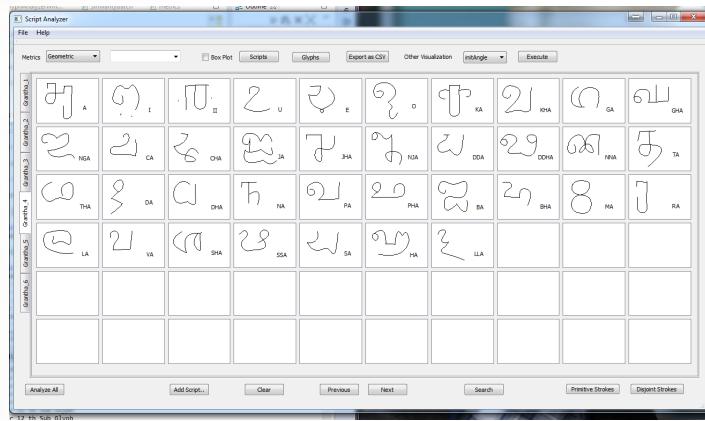


University of
St Andrews

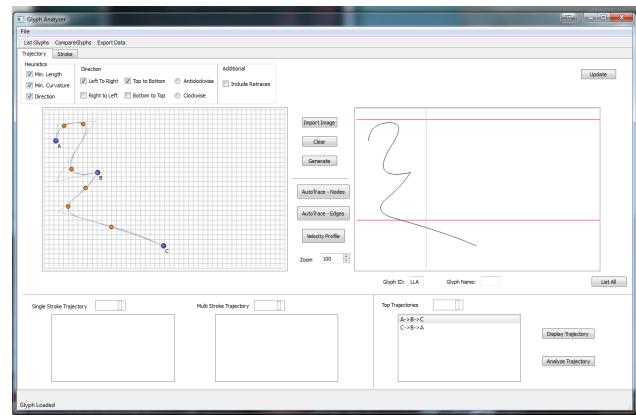
FOUNDED
1413



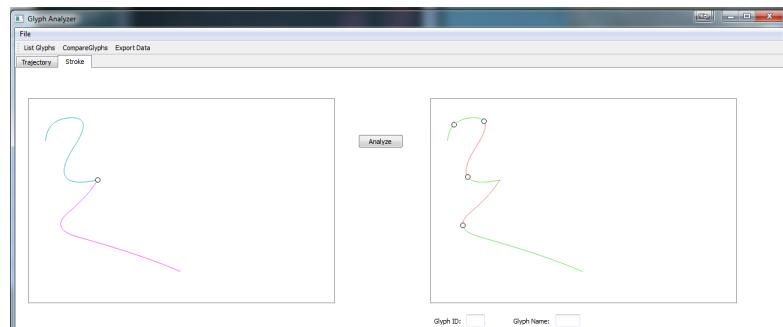
Script Analysis Framework



Script Repository



Digitized character



Decomposed character



University of
St Andrews

FOUNDED
1413

sicsa* The Scottish Informatics & Computer Science Alliance



Feature Sets

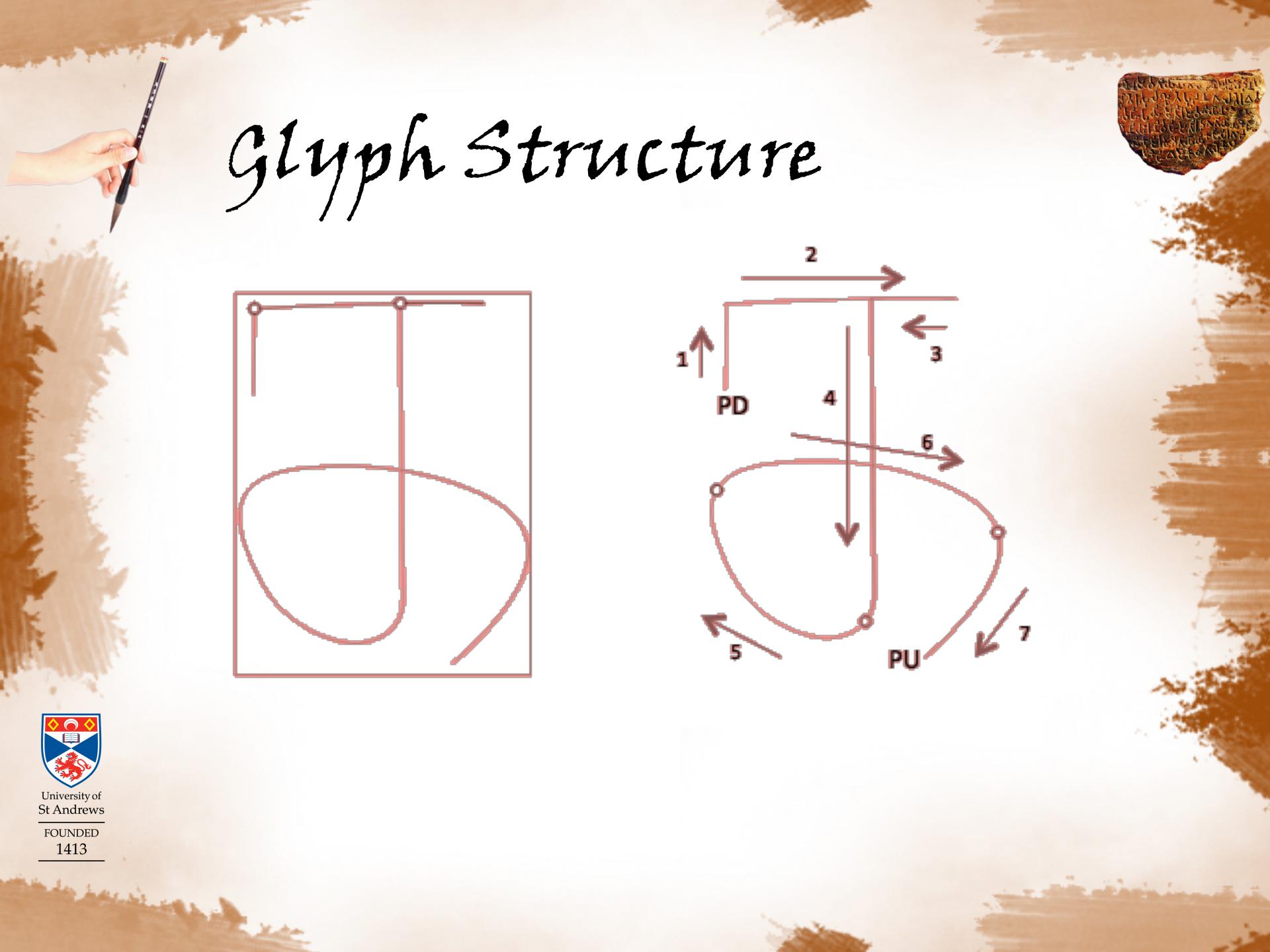
- ② Two sets of features are being extracted
 - ② 9 static features based on shape of the character
 - ② 12 Production features based on the written trajectory of the characters



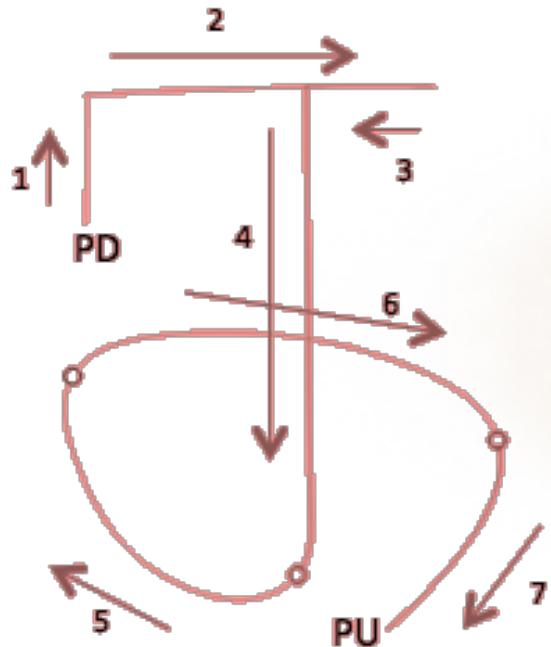
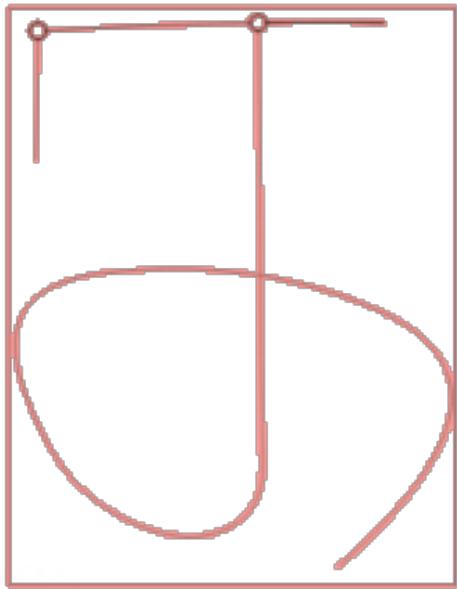
University of
St Andrews

FOUNDED
1413





Glyph Structure



University of
St Andrews

FOUNDED

1413

Geometric Features

- ⌚ Length
- ⌚ Divergence
- ⌚ Length-Breadth Index
- ⌚ Size
- ⌚ Openness
 - ⌚ (= Divergence/Length)
- ⌚ Compactness
 - ⌚ (= Length/Size)
- ⌚ Average Curvature
- ⌚ circularity
- ⌚ Rectangularity



University of
St Andrews

FOUNDED
1413



Production Features

⌚ Stroke Counts

- ⌚ Pen
- ⌚ Disjoint
- ⌚ Retraces
- ⌚ Up
- ⌚ Down

⌚ Disfluency

⌚ Changeability

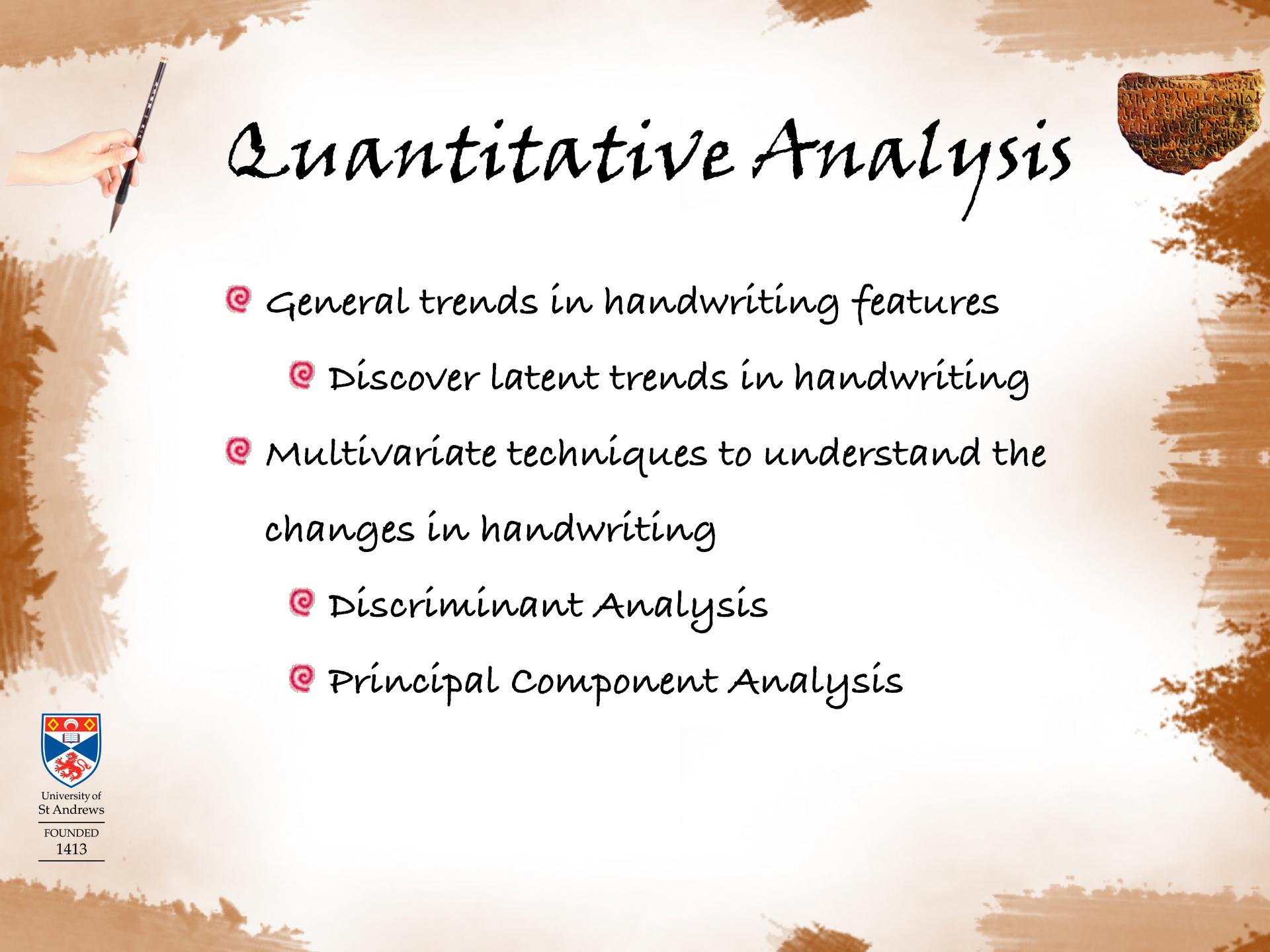
- ⌚ Entropy
- ⌚ Disjoint Angles
- ⌚ Stroke Lengths
- ⌚ Major Stroke Lengths



University of
St Andrews

FOUNDED
1413





Quantitative Analysis

- ⌚ General trends in handwriting features
 - ⌚ Discover latent trends in handwriting
- ⌚ Multivariate techniques to understand the changes in handwriting
 - ⌚ Discriminant Analysis
 - ⌚ Principal Component Analysis



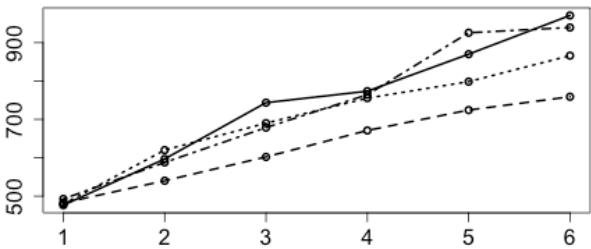
University of
St Andrews

FOUNDED
1413

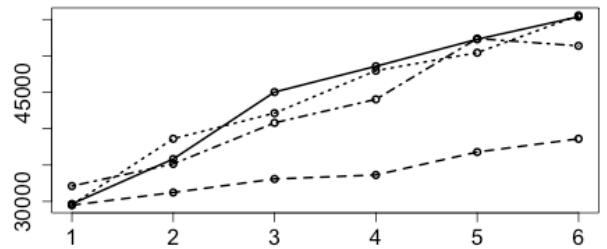
Trends: Geometric



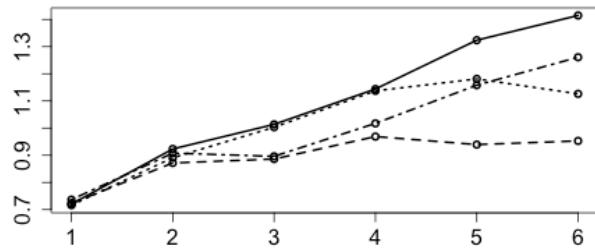
Length



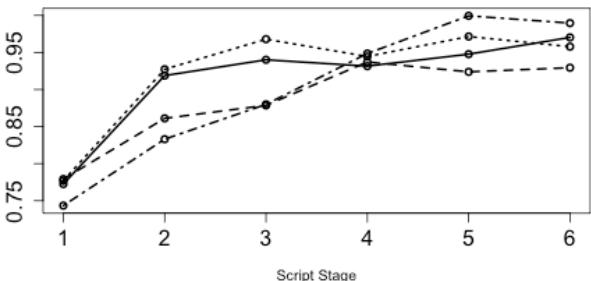
size



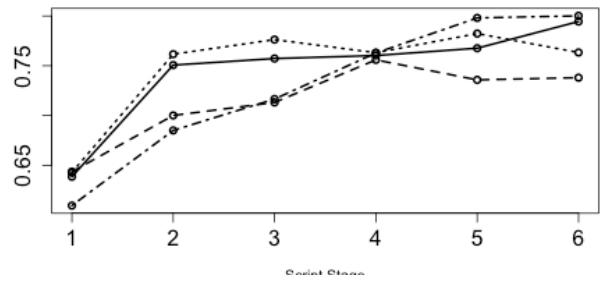
LBindex



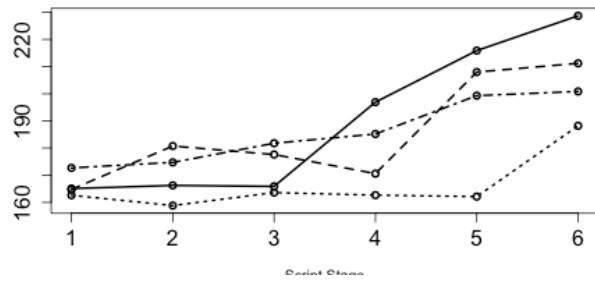
circularity



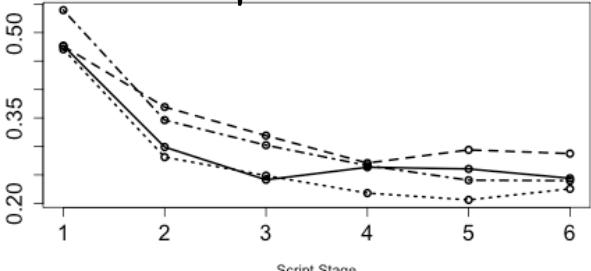
Rectangularity



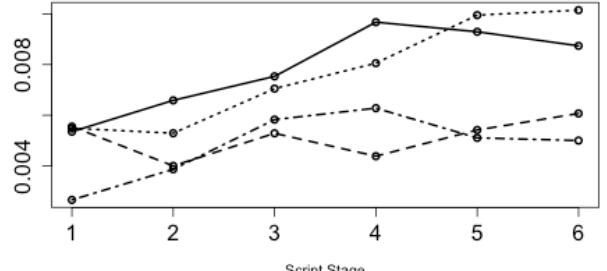
Divergence



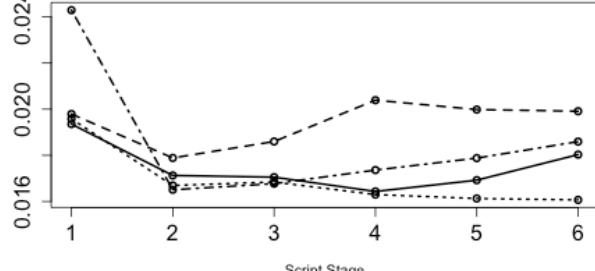
Openness



Avg. Curvature



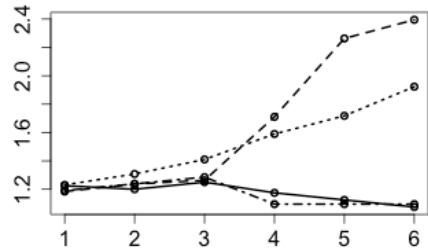
Compactness



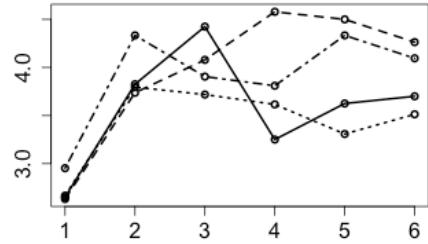
Trends: Production



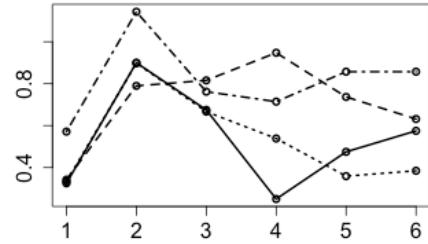
Pen Count



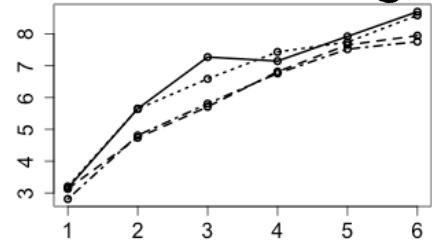
Disjoint Count



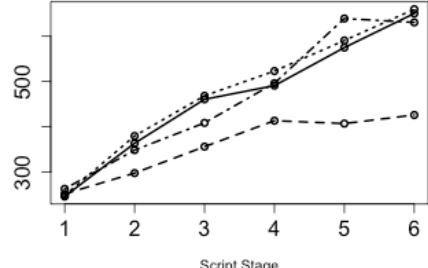
Retrace Count



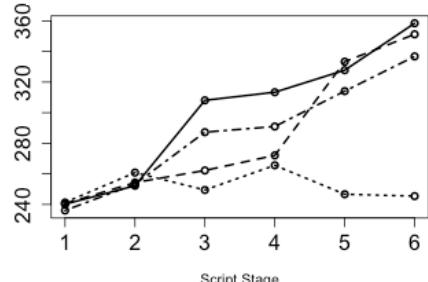
Disfluency



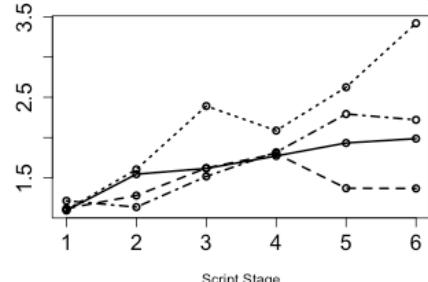
Up Count



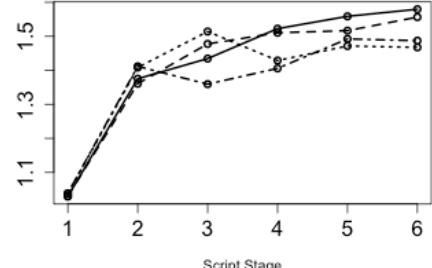
Down Count



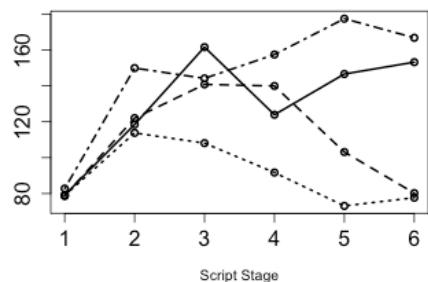
Changeability



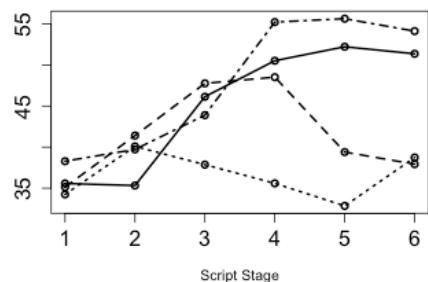
Entropy



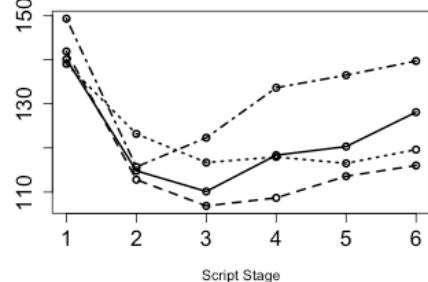
Disjoint Angles (Sum)



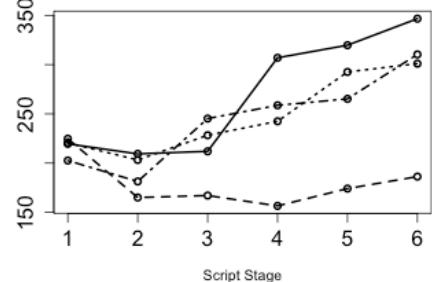
Disjoint Angles (Mean)



Stroke Lengths (Mean)



Major Lengths (Mean)





Diversification

- ② Apply discriminant analysis to find aggregate discriminants that “discriminate” different scripts
- ② In terms of script development these can be further elaborated as the factors, which caused diversification.



University of
St Andrews

FOUNDED
1413



Diversification: Geometric

- ② Scripts have diversified based on the following major geometric features:
 - ② compactness, average curvature, circularity and rectangularity.
- ② Characters' curvature and their shape outlines have together played a major role in diversification.



University of
St Andrews

FOUNDED
1413



Diversification: Production

- ② Scripts have diversified mostly based on:
 - ② entropy of writing and the number of major strokes in characters
- ② Entropy of characters appears to have been a major factor in diversification



University of
St Andrews

FOUNDED
1413



Spread of Variations

- ⌚ We discussed the general trends in various features of scripts
- ⌚ Now, we analyze the individual character variations that occurred.
- ⌚ We use principal component analysis to discover descriptive aggregate features
 - ⌚ Also reduces the feature set



University of
St Andrews

FOUNDED

1413



Principal Components



- ⌚ For PC1, characters that are “open”, “short” and “asymmetric” have positive scores
 - ⌚ Large negative scores are “closed”, “long” and “symmetric” characters
- ⌚ For PC2, characters that are “compact” and very “symmetric” have negative scores
 - ⌚ Positive scores indicate characters that are “large” “loose” “divergent”.



University of
St Andrews

FOUNDED
1413



Principal Components



- ② For PC3, high negative scores indicate “large” and highly “curved” characters.
- ③ For PC4, large negative scores point to highly “curved” and “symmetric” characters
 - ④ Positive scores pointing to characters that lack those characteristics.

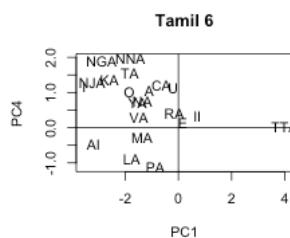
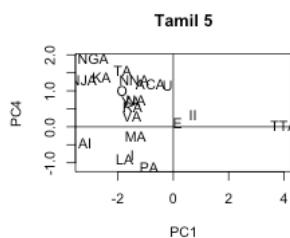
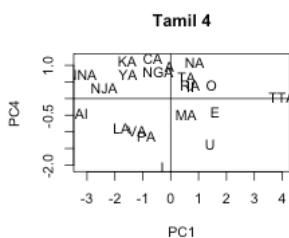
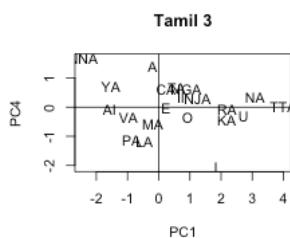
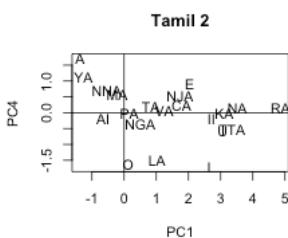
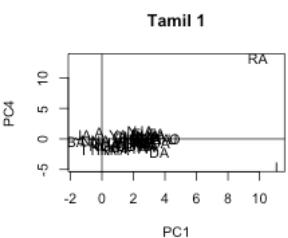
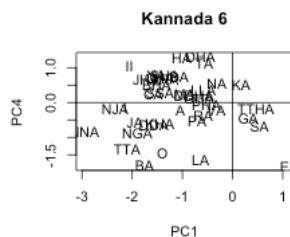
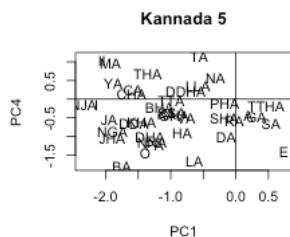
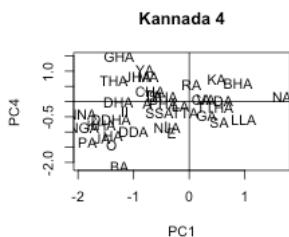
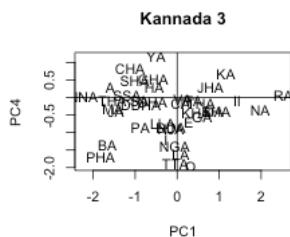
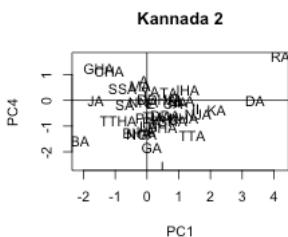
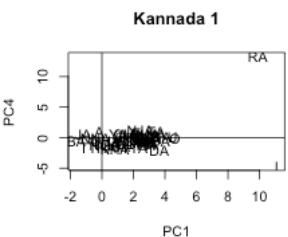
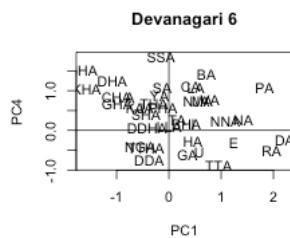
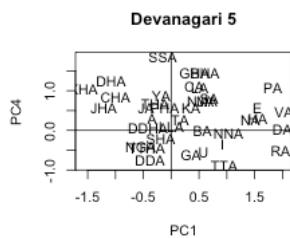
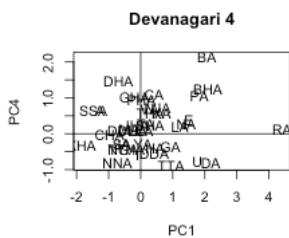
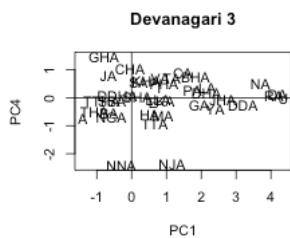
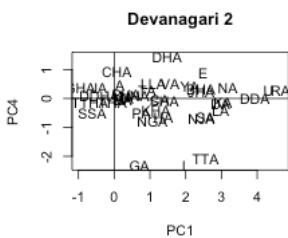
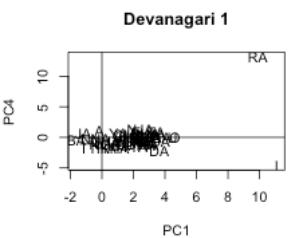
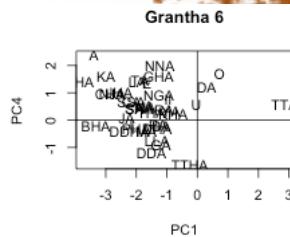
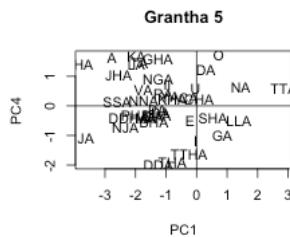
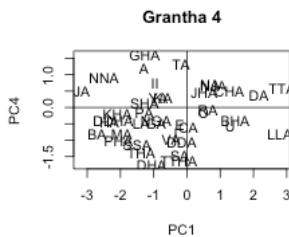
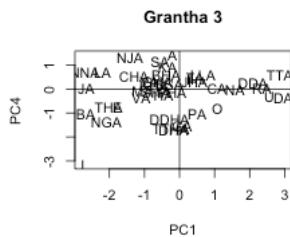
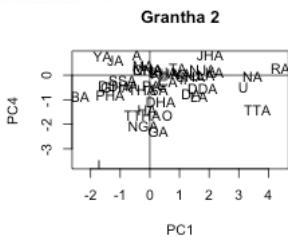
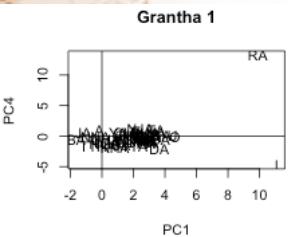


University of
St Andrews

FOUNDED
1413

Spread of Variations

PC1 VS PC4





Spread of Variations



- ⌚ In Brahmi, the characters are just “open” “short” and “asymmetric”
- ⌚ During the second stage characters gain more “symmetry” “closure” and “length” and ultimately many of the characters gain “curved symmetry” along with “lengthy closure”.
- ⌚ We can clearly see the interplay of features that cause the variations.



University of
St Andrews

FOUNDED

1413



Applications

- 🕒 Paleography has mostly been a subjective field.
- 🕒 The quantitative techniques and feature sets can contribute towards a more objective and quantitative paleographic analysis.
- 🕒 Results presented are specific, but the methodology can be applied to similar analysis



University of
St Andrews

FOUNDED
1413

Future Work

- ⌚ The nature of distribution of features and their corresponding changes
- ⌚ Influence of usage frequency on character properties is also to be studied in detail.
- ⌚ Analyzing the change in stroke inventory and their impacts
- ⌚ Specific feature sets that have produced fairly stable characters.
- ⌚ Interaction between the geometric and production features



University of
St Andrews

FOUNDED

1413



Conclusion

- ⌚ Presented the general trends in handwriting that occurred during script development
- ⌚ Found the major features on the basis of which Indic scripts diversified over the years.
- ⌚ Analyzed the variations acquired by individual characters using aggregate features.
- ⌚ Discussed applications and future work



University of
St Andrews

FOUNDED
1413



Thanks!

?

Questions?



University of
St Andrews

FOUNDED
1413