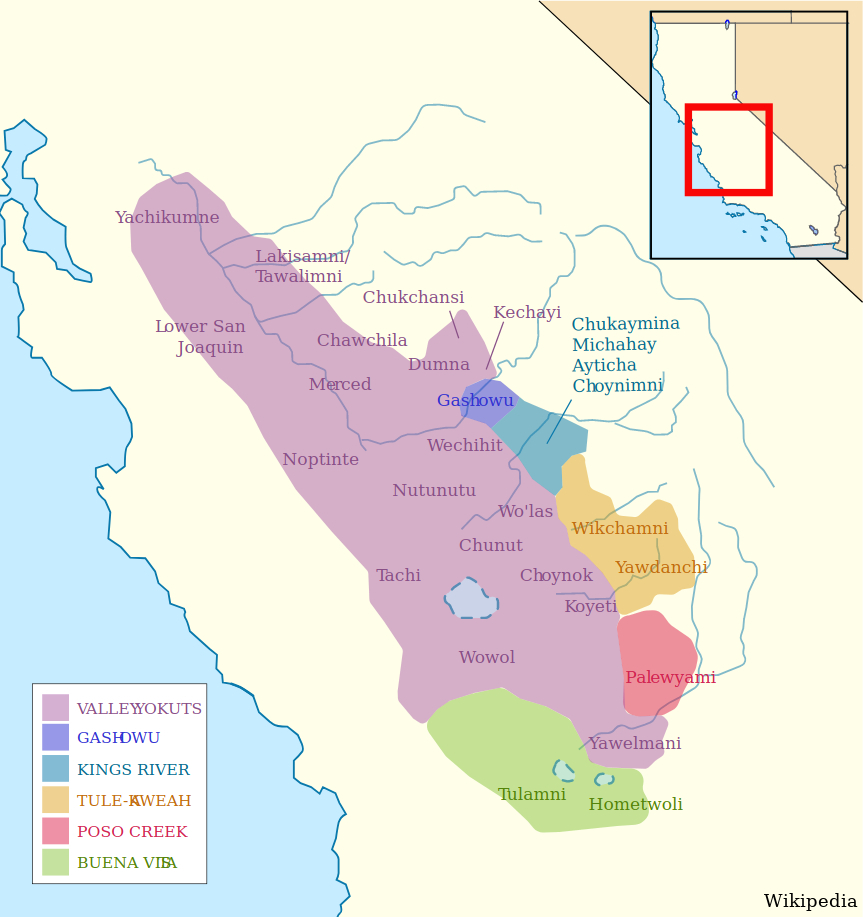
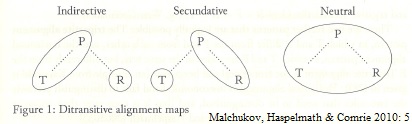
1/12 @wieldorg #FragmentedLanguageWorkshop2020

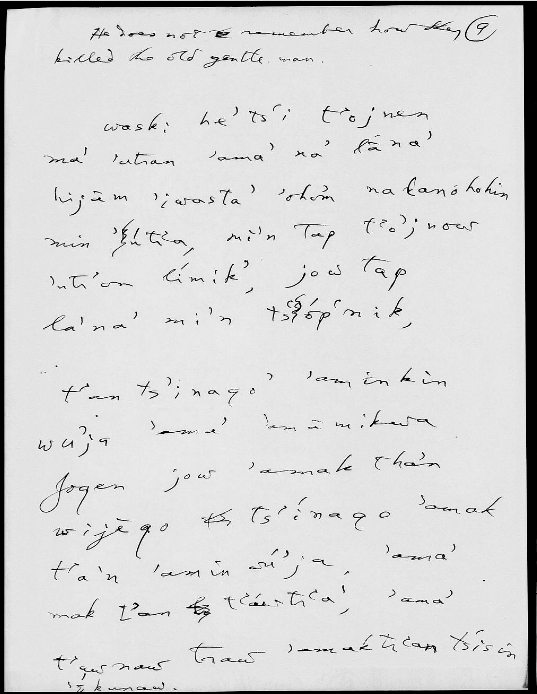
Yowlumne (aka Yawelmani) is a Yokuts language of the Valley Yokuts branch native to California. It appears near the bottom right of the large purple area of the map.



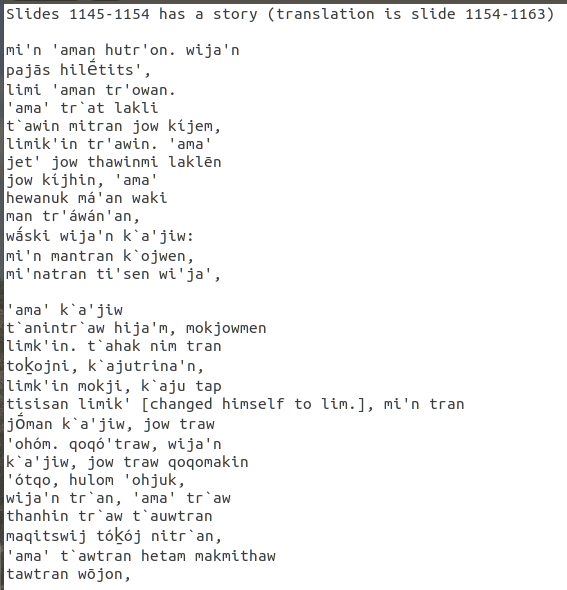
2/12 Yowlumne has a complicated system of secundative alignment in ditransitive constructions (Weigel 2005). Secundative alignment is where the recipient-like argument (R) of a ditransitive verb patterns like the patient-like argument (P) of a transitive verb (MHC 2010).



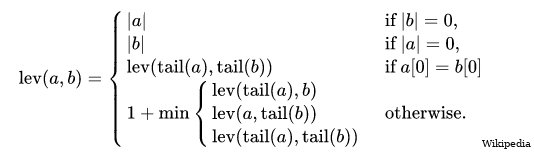
3/12 Large amounts of available data that document the alignment system in Yowlumne are found in messy, handwritten archival documents in the archives of John Peabody Harrington and Stanley Newman, recorded between 1910 and 1940. This data is not readily searchable in this form.

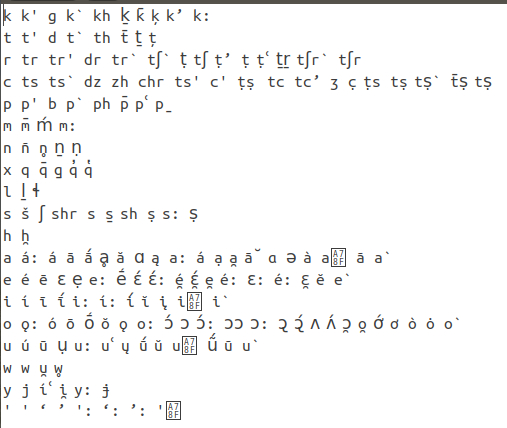


4/12 Digitization and accessibility to these documents takes place in several stages. The first is digital transcription of 17 texts (so far).

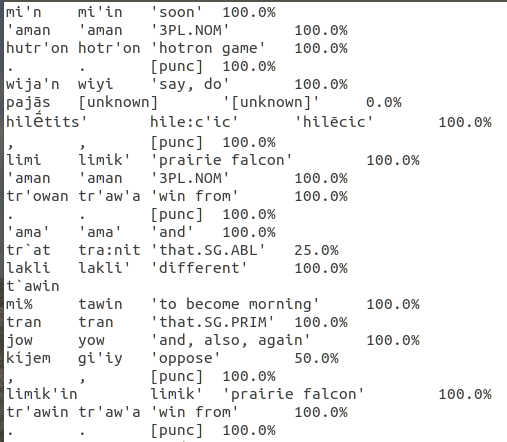


5/12 The second stage is a preliminary text normalization process using an algorithm involving Levenshtein distance (image 1). Harrington confused many sounds in his transcriptions, so I use sound classes (image 2) to group these sounds in the Levenshtein distance calculations.

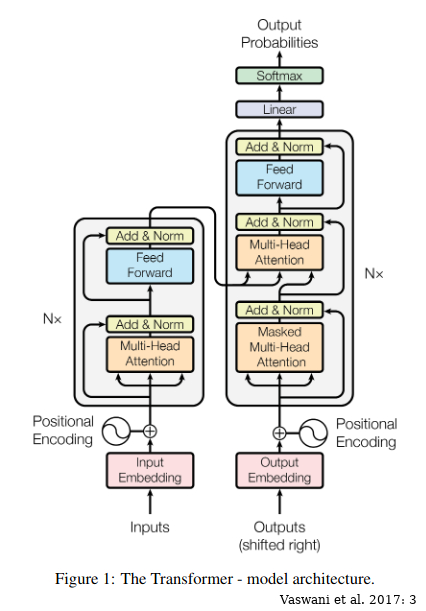


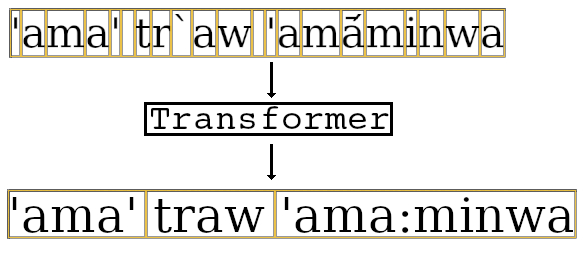


6/12 This first pass at normalization is then hand-checked and corrected for accuracy, with a percentage given to indicate subjective certainty for the word identified.

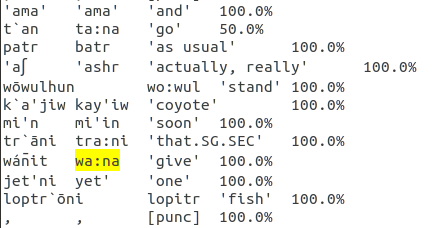


7/12 This first-pass normalized data is used to train a Transformer model (Vaswani et al. 2017) that maps from characters to lexemes. The Transformer model can then be used to normalize other data obtained from the archives.

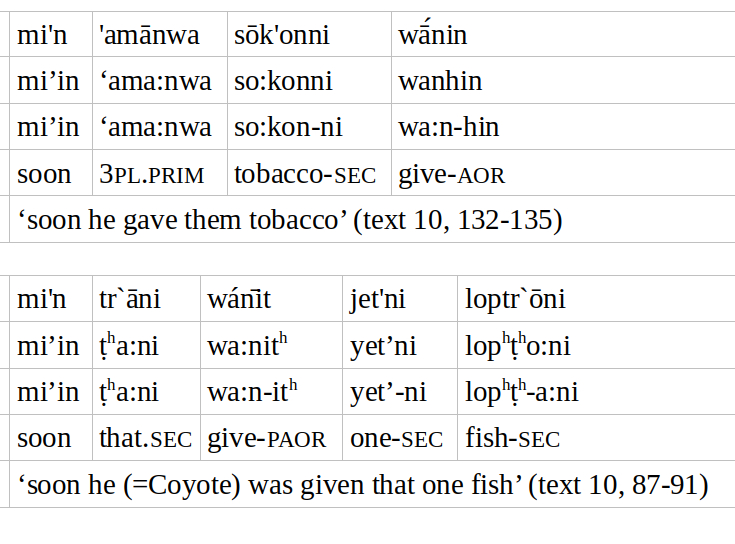




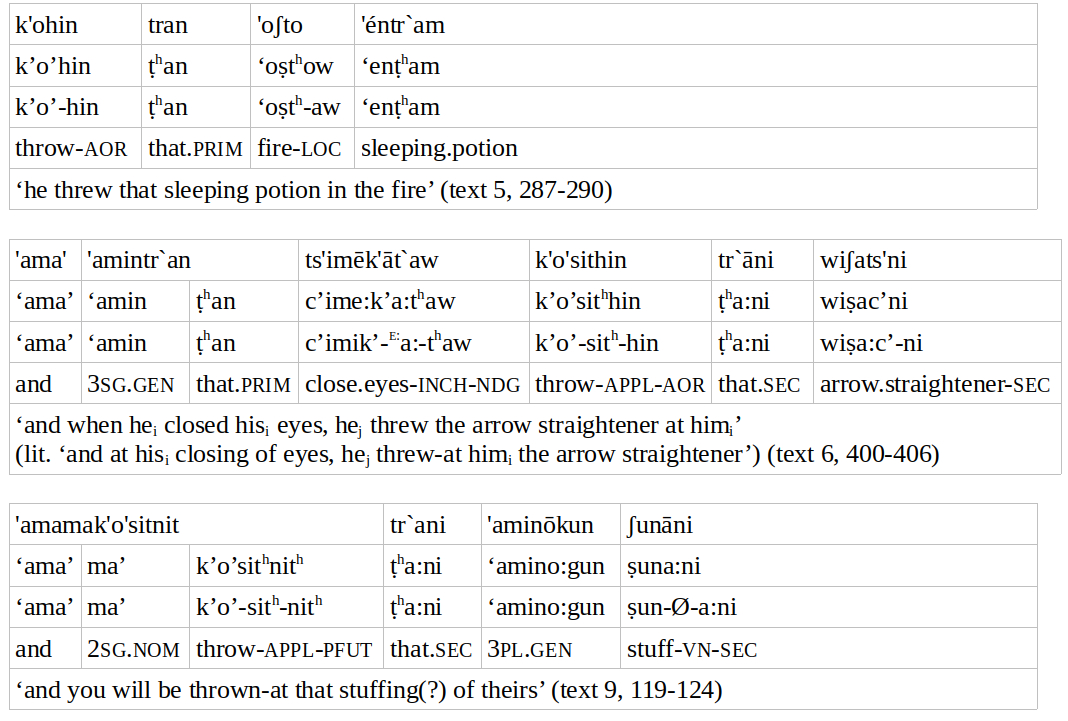
8/12 The data normalized by the Levenshtein distance calculations and the Transformer model can be searched by lexeme. This enables selecting individual ditransitive verbs (e.g. \*wa:na ‘give’) to chart out their attested argument structures, including in passive environments.



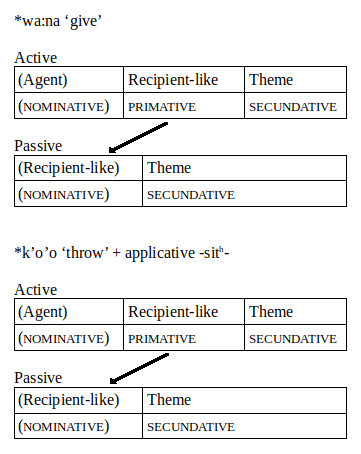
9/12 The data obtained provide preliminary results for ditransitive alignment with passivization: the examples show \*wa:na ‘give’. The primative-marked argument in the active sentence is the recipient, while the secundative-marked argument is the theme in active and passive.



10/12 Applicativized verbs work the same way, as with \*k’o’o ‘throw’. For the base verb, the theme (T) is primative-marked. The applicative demotes T to secundative, and the recipient-like argument (R) is primative. In the passive, R is subject, and T remains secundative.



11/12 In either case, the preliminary results show that passivization applies to the recipient-like argument (R): R becomes subject. The theme remains a secundative-marked argument regardless. This is true for both ditransitive verbs and applicative-marked verbs.



12/12 In conclusion, the computational approach used to access the data will be useful to those working with archival data. This approach enabled analysis of the ditransitive alignment in Yowlumne and found that passivization promotes the primative-marked argument to subject.

