

Automatic intonational contour clustering in Patwin

Anna Björklund*

Abstract. This study uses automated methods from Contour Clustering (Kaland 2021) to identify seven common intonational patterns in Patwin, an understudied Wintuan language of Northern California that survives via archival recordings. Only two phonetic or phonological analyses currently exist for Patwin (Lawyer 2015, 2021). This study finds that all seven contours suggested by Contour Clustering are attested in word-list elicitation, demonstrating a remarkable diversity of intonational types. In so doing, this study challenges claims made in Shafer (1961) that Patwin has lexical tone. Though the results are generally successful, Contour Clustering is not robust to the effects of poor recording quality on pitch tracking and subsequent cluster assignment. In general, this study indicates that using automated methods in tandem with a more phonologically grounded method of analysis such as PaToBI (Silverman et al. 1992; Björklund 2024) is fruitful for facilitating working with large amounts of archival data. This study adds to our limited understanding of Wintuan intonation, suggesting new intonation types for future investigation.

Keywords. prosody; intonation; Californian languages; Wintuan languages; automated clustering

1. Introduction. The Wintuan languages are a sleeping language family spoken in the Sacramento Valley of California.¹ The languages – Wintu, Nomlaki, and Patwin – form a vertical chain along the valley, from Mount Shasta in the north to Suisun Bay in the south. The divergence and time depth of Wintuan is similar to Romance, at approximately 1,500-2,500 years (Whistler 1980; Shepherd 2005). Patwin belongs to the southern branch of the Wintuan family, which, based on linguistic reconstruction of regional flora and fauna, is thought to have diverged first from the Proto-Wintun homeland in southern Oregon (Whistler 1977). This is reflected by the relatively high degree of divergence between Patwin and the Northern Wintuan branch (Wintu and Nomlaki) in phonology, morphology, and syntax (Golla 2011). Southern Patwin, for which only a few word lists remain, is alternately analyzed as a dialect of Patwin (e.g. Kroeber 1932) or a separate member of the Southern Wintuan branch (e.g. Whistler 1980; Shepherd 2005), and is not further discussed in this paper.

Literature on Patwin phonetics and phonology is limited, as is true for the entire Wintuan family. Prior to the 21st century, the only description of Patwin phonetics was Shafer (1961). Shafer (1961) provided wordlists for ‘Southwest’ and ‘Southeast’ Patwin, identified by Lawyer (2021) as Hill and River Patwin, respectively. Shafer additionally claimed that these Patwin dialects had a system of five tones. This claim is doubtful: only a few languages of the region are tonal, and none with as complex a system as Shafer suggests (Golla 2011; Maddieson 2013). However, the aforementioned tonal languages are all located in northern California and Oregon, near the northern border of Wintuan. As tone is often areal (Southeast Asia being perhaps the

* University of California, Berkeley (aebjorklund@berkeley.edu)

¹ ‘Sleeping’ language is here used —instead of ‘extinct’ or ‘dead’— to indicate a language currently without living native speakers. As noted by scholars such as Perley (2012) and Leonard (2008), using death metaphors to describe language status (e.g. moribund, extinct, dead) implies permanent annihilation. This is not the outlook we wish to encourage for these languages or communities, especially as many undergo language revitalization efforts.

most famous example; see Kirby and Brunelle (2017)), it is not entirely out of the question that it might have spread to Wintuan from one of these nearby language families.

Lawyer (2015, 2021) provides the only other analyses of Patwin phonetics and phonology. As with sister language Wintu (Pitkin 1984), stress assignment in Patwin is based on syllable weight (Lawyer 2021), where stress is associated with increased loudness, raised pitch, and vowel lengthening. Stressed vowels do not change in quality. In discussing Patwin intonation, Lawyer (2021) notes that phrases in the middle of a connected passage may end in rising pitch, whereas those uttered in isolation generally fall, regardless of whether they are questions or statements. Little other description of Patwin intonation currently exists.

As Patwin is a sleeping language, several unique research challenges exist. Most phonetic work assumes that the researcher is situated either in a laboratory or a traditional field site (Whalen and McDonough 2015). Both situations assume greater control over the research program than a researcher working with a sleeping language likely has. In a laboratory, stimuli can be managed in accordance with the researcher's goals; in the field, the researcher can directly work with the speakers of the language. In the case of a sleeping language, the researcher cannot ask consultants for either clarification or judgement, nor gather new data. Rather, the research must be shaped by the available materials, which are often used for different purposes than originally intended. This may necessitate spending considerable amounts of time becoming familiar with existing audio and notes, much (or all) of which may not be digitized or searchable. This time requirement becomes a barrier for entry, and is one cause for detailed prosodic research to be rare for endangered and sleeping languages (Whalen et al. 2022).

Prosody is particularly difficult to investigate in sleeping languages. Among the challenges laid out by Himmelmann (2008) are the facts that 1) prosodic patterns are highly variable and contextual, and 2) relevant prosodic contrasts cannot be directly analyzed from the speech signal, but are perceptually bound, i.e. require a native speaker's judgement. However, much can be done to improve the situation if the researcher has a defined set of intonational contours whose uses can be investigated.

Automated prosodic clustering, such as that of Kaland (2021) therefore appeals for two reasons: 1) to streamline the researcher's process of becoming familiar with a new (and often large) body of data, and 2) to provide the researcher with a set of intonational contours present in the language, which can then be further investigated. Kaland (2021) additionally adds that automatic clustering is warranted to curtail researcher bias. Because most researchers are not native speakers of the language they work with, there is a high potential for the researcher to either miss important patterns that are not analogous in their native language, or to ascribe undue meaning to contours that are analogous to native patterns. Such bias is particularly problematic in the case of sleeping languages, where the consultant cannot provide insight into which patterns are meaningful. Using automated clustering is therefore not simply a matter of time efficiency, but also ensures a level of quantitative proof for the contours posited by the researcher. Previous uses of Kaland (2021) include Babinski and Bower (2022), who use CC to identify additional phrase types in Bardi, a Nyulnyulan language from Northern Australia.

2. Data. This study is based on two recordings between Cortina Hill Patwin speaker Nora Lowell and linguist Elizabeth Bright. The first was a 24.5-minute elicitation session of words and short phrases (Bright 1952c), intended to establish the language's phonemic inventory. The second was a 7.5-minute text elicitation, in which Lowell retells two Patwin myths (Bright 1952b). This study examines the second myth in this recording, entitled *The Creation of Heaven and Hell*. In total, the word list elicitation consisted of 333 intonational phrases, while the text consisted of

83. Both recordings were retrieved from the digitized collections of the California Language Archives (CLA), where they were preserved at the archive-standard sampling rate of 96 kHz.

Together, the two recordings provide examples both of single-word intonational phrases spoken in controlled speech and longer phrases in connected speech. This allows us to see how potential contours behave across a variety of phrase lengths and levels of articulatory care. The word list elicitation is also useful for attempting to diagnose what Shafer (1961) claims is lexical tone. The elicited text (Bright 1952b) was additionally chosen because it was one of the only Patwin recordings in the CLA’s collection with an accessible gloss (Bright 1952a). The gloss does not fully match the audio recording available in the CLA, but is nonetheless an important reference for understanding the meaning of the story and its constituent phrases. It is possible that Bright elicited this story from Lowell several times, and that the available recorded audio (Bright 1952b) and gloss (Bright 1952a) represent two separate retellings. The final gloss of the text used in this study (Bright 1952b) represents a combination of Bright’s glosses as well as the author’s own, supplied when Bright’s were either absent or lacking full analysis.

3. Methods. This study used the Contour Clustering (CC) toolkit (Kaland 2021) to automatically cluster Patwin intonation contours. In the first stage, contours were demarcated using Praat TextGrids (Boersma and Weenik 2024); F0 information was then extracted for each contour at 20 equidistant intervals. CC offers either an R or Praat script for this step. As there were issues running these scripts on the author’s computer, the Praat script was rewritten by the author in Python, using the audiolabel and parselmouth libraries; all settings were copied from CC’s default settings.

In the second stage, the contours’ F0 information was inputted into the CC GUI (written in R). Before cleaning, the data consisted of 416 contours. 376 contours remained after selecting the application’s ‘clean data’ option, which removed contours containing missing values and/or F0 errors. F0 errors are defined by Kaland (2021) as instances where the mean F0 value of any point in the contour exceeds a ratio of 0.01 before and after octave jump handling. Additional anomalous contours were removed by expanding the number of desired clusters beyond the expected number, then removing those flagged by the application (per the recommendations of Kaland 2021). Flagged clusters included those containing only one contour, and those whose mean standard error (MSE) was two or more times the median MSE for all clusters. In this paper, the number of clusters was set to 25. After removing these flagged clusters, 364 clusters remained (90.3% of total data).

CC uses agglomerative clustering with complete linkage, where cluster similarity is determined through Euclidean distance (1). Each contour contains 20 equally-spaced F0 measurements. Clusters with smaller calculated distances are considered more similar than those with larger distances. The most similar clusters (or cluster groupings) are merged iteratively until only one cluster remains.

$$(1) \quad d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}$$

To compute the distance between clusters where at least one cluster contains more than one contour, CC uses complete linkage (also called farthest neighbor linkage), represented by (2). Two clusters are considered most similar when their *most-distant members* are still found to have the smallest distance when compared between all clusters.

$$(2) \quad D(X, Y) = \max_{x \in X, y \in Y} d(x, y)$$

The output of this clustering algorithm is a dendrogram, representing cluster groupings at each stage of the merging process. Determining the appropriate number of clusters amounts to deciding how many contour types are ultimately informative. Assuming many clusters may allow us to discover low-level, detailed patterns, while assuming few clusters may indicate general, high-level patterns.

Supplementary material from Kaland (2021) recommends identifying the appropriate number of clusters along two criteria: 1) the number of members per cluster should be relatively homogenous, and 2) clusters should avoid high MSE values, as Kaland (2021) notes that listeners can find deviations of as little as 10 Hz to be meaningful. More qualitatively, the researcher should cut the dendrogram when increasing the number of clusters does not seem to add new cluster ‘types.’ This paper follows Kaland’s (2021) recommendations in order to maximize comparability and reproducibility.

Following the process recommended by Kaland (2021) and the evaluation tool provided in the CC GUI, 7 clusters were created. With seven clusters, the largest MSE is 2.82 Hz; with eight clusters, the largest MSE increases to 3.68 Hz, which continues to increase as clusters do. Seven clusters also creates a relatively balanced number of contours per cluster-- with each having 70, 50, 34, 69, 45, 85, and 11 members, respectively. Cluster balance appears to degrade with the inclusion of 9 or more clusters, where the smallest clusters begin containing $n = 4$ contours. As this study is interested in the most common Patwin intonational patterns, it is undesirable to examine clusters of such small size. Seven clusters are also preferred by the CC evaluation tool, based on its calculation of information cost.

In order to guide analysis, phrases were tagged with general part of speech or function. These labels were intended to provide a basic starting point for keeping track of the general shape of the data, but are not meant to be suggestive of precise functions. Phrases were tagged as repetitions of a word (first repetition: 46.98%, second repetition: 31.87%), declaratives (5.77%), hortatives (2.20%), quotatives (1.65%), and declarative noun phrases (5.77%). Declarative verb phrases (1.37%), unclear phrase types, and imperatives each accounted for (1.37%); negative declaratives accounted for (1.10%). The following types comprised less than 1% each of the dataset: third and fourth members of a list, onomatopoeia, partitives, declarative prepositional phrases, interrogatives, causatives, concessives, and adverbial phrases.

4. Results.

4.1. GENERAL RESULTS. Figure 1 illustrates the dendrogram produced for the Patwin data before the removal of outliers.

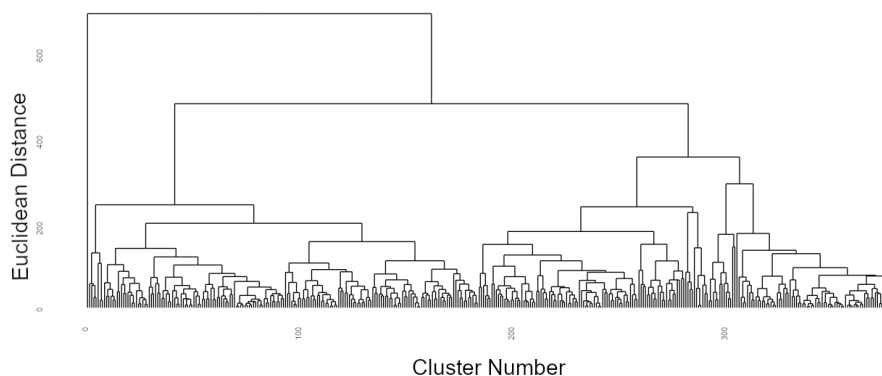


Figure 1. Clustering dendrogram for Patwin intonational contours

With the exception of the single outlier cluster on the far left, we can see seven clusters of similar size approximately midway down the graph. These correspond to the ‘cut’ of the dendrogram we will further examine in this section. The average intonational contours for $n=7$ clusters is shown in Figure 2, with the black line in each grouping illustrating the average contour for each cluster.

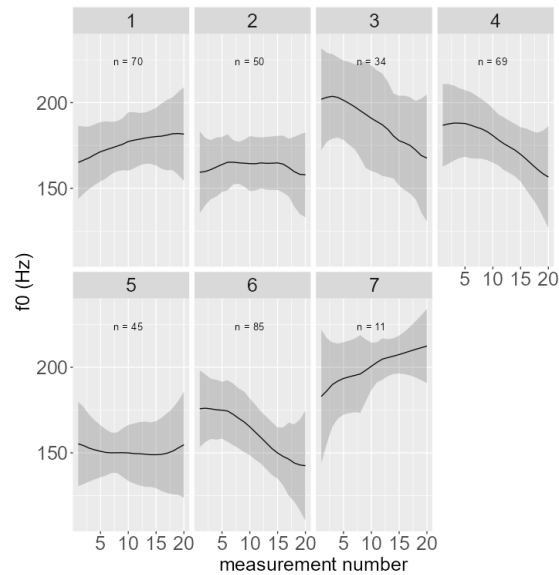


Figure 2. Average intonational contours when n clusters = 7

A few basic patterns emerge in these seven clusters. Three clusters appear to be falling (Clusters 3, 4, and 6) and two appear to be rising (Clusters 1 and 7). Two other clusters (2 and 5) appear relatively level. Word list repetition is the most represented utterance type across all clusters – as data from word lists is overrepresented overall, this is not surprising.

4.2. FALLING CLUSTERS. Clusters 3 ($n = 34$), 4 ($n = 69$), and 6 ($n = 85$) all show a general falling pattern. Small differences appear upon further examination. Cluster 3 begins at the highest pitch of the three (~ 200 Hz); Cluster 4 begins lower (~ 190 Hz); Cluster 6 begins lowest (~ 175 Hz). The slopes of Clusters 4 and 6 both begin relatively high and level before falling in pitch. However, in the case of Cluster 4, pitch increases slightly before falling; in Cluster 6, pitch remains flatter before falling. In Cluster 3, pitch appears to rise briefly before the contour falls, and the delay before the fall is shorter than in Clusters 4 and 6. Figure 3 illustrates three phrases that were assigned to Clusters 3, 4, and 6, respectively.

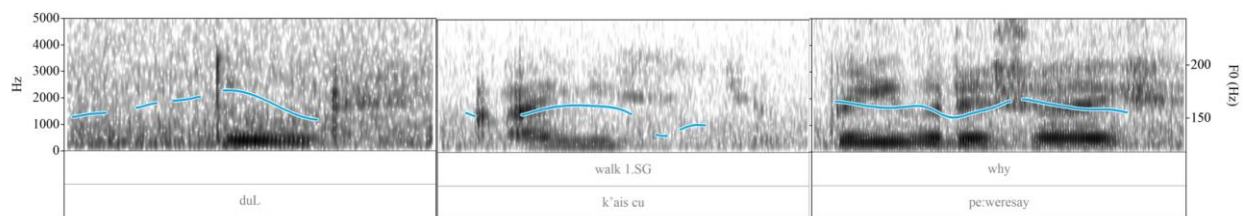


Figure 3. Left-to-right: phrases tagged as Clusters 3 (duł [unknown meaning]), 4 (k’ais cu ‘I walk’), and 6 (pe:weresay ‘why’)

An examination of the pitch tracks in Fig. 3 provides some context for the abstracted average contours shown in Fig. 2. Perhaps as expected, the contours assigned to Clusters 4 and 6 (Fig. 3a and 3c) are most similar, beginning with a relatively high, level pitch that falls later in the phrase. It is unclear how these two clusters might be meaningfully distinct. In a ToBI-style system for Patwin such as PaToBI (Björklund 2024), these contours would be annotated nearly identically: (%H) H* L-L%, where Cluster 4 (Fig. 3b) lacks a high boundary tone %H. Perhaps such a high initial boundary tone in Cluster 6 (Fig. 3c) accounts for its greater initial flatness in the Fig. 2 pitch track.

Ignoring the erroneous pitch tracking at the beginning of the phrase, the contour assigned to Cluster 3 (Fig. 3a) falls in pitch immediately after the release of the initial stop. In PaToBI, this phrase would likely be identified as H*L L-L% or H* L-L%. (As PaToBI is still in its initial stages, it is unclear if these analyses are distinguishable in a monosyllabic context). If the former is true, Cluster 3 (Fig. 3a) may simply be a compressed variant of the intonational pattern in Cluster 4 (Fig. 3b).

Tagged phrase types for Cluster 3 are overwhelmingly the first repetition of a word (70.59%), followed by the second repetition of a word (17.65%), and imperatives (5.88%). Third members of a list or hortatives are both 2.94%. For Cluster 4, the tagged phrase types are also mostly first repetitions of a word (55.07%), followed by second repetitions (39.13%), declaratives (2.90%), and declarative noun phrases (1.45%). First and second repetitions of a word are more equal in Cluster 6 (44.70% and 43.53%, respectively), followed by declaratives (3.53%). Hortatives, interrogatives, third and fourth repetitions of a word, onomatopoeia, declarative verb phrase, and unknown phrases each accounted for 1.18%.

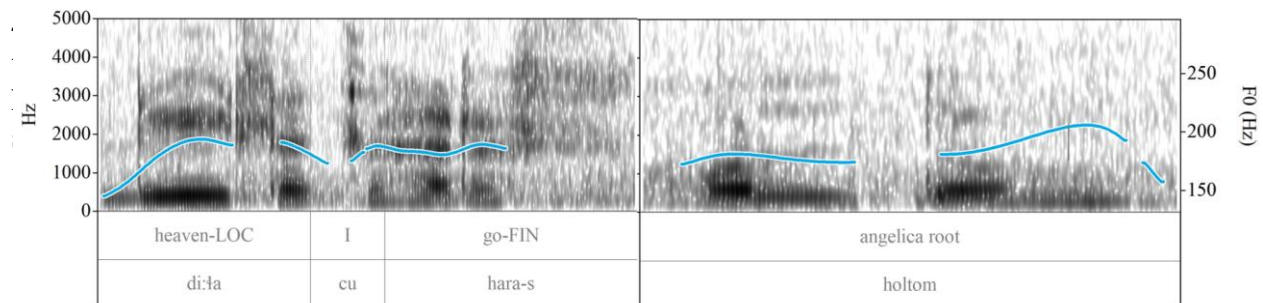


Figure 4. Left-to-right: Phrases tagged as Clusters 1 (di:la cu haras ‘I am going to heaven’), and 7 (holtom ‘angelica root’)

In examining Fig. 4, we see that Fig. 4a appears to be relatively level throughout, discounting what is likely a segmental effect of the initial voiced stop in di:la ‘to heaven.’ In contrast, Fig. 4b appears high-level until the last syllable, where the pitch is upstepped. Such a contrast may be generally captured in PaToBI as H* H-H% (Fig. 4a) versus H* H^H% (Fig. 4b). This seems to corroborate the notion in Fig. 2 that Cluster 7 (upstepped) is on average higher in pitch than Cluster 1.

The top three most common utterance types in Cluster 1 are first repetitions of a word (45.71%), second repetitions of a word (18.57%), and declaratives (8.57%). Declarative verb phrases and negative declarative are each 4.29%; declarative noun phrases and quotatives are each 2.86%, and adverbials, causatives, partatives, and third/fourth repetitions of a word are each 1.43%. Cluster 7 is overwhelmingly word repetition from the elicitation recording, with the first repetition of a word constituting 90.91% of phrases, followed by declaratives (9.10%).

4.4. LEVEL CLUSTERS. Clusters 2 and 5 are mostly level clusters: whereas Cluster 2 appears convex, Cluster 5 is concave. Cluster 2 appears to begin and end at around 155 Hz, while Cluster 5 begins slightly lower and ends a bit higher. Figure 5 illustrates three phrases from the dataset that were assigned to Clusters 2 and 5, respectively.

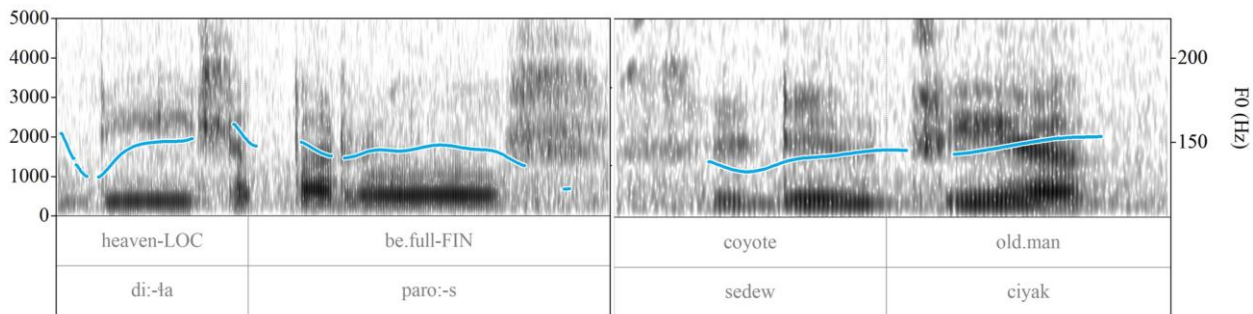


Figure 5. Left-to-right: Phrases tagged as Clusters 2 (*di:la paro:s* ‘heaven was full’) and 5 (*sedew ciyak* ‘Old Man Coyote’)

In Fig. 5, we find that the phrase assigned to Cluster 2 (Fig. 5a) maintains a high level pitch throughout, while the phrase assigned to Cluster 5 (Fig. 5b) appears to shallowly increase in pitch at a steady rate throughout the phrase. Though both are assigned to ‘level’ clusters by CC, Fig. 5b appears to actually be upstepped; in PaToBI it would likely be transcribed as %H H^H% (with no pitch accent), whereas Fig. 5a would be transcribed H* H* H-H%. The latter is identical to Cluster 1 in Fig. 4a.

The top most common utterance types in Cluster 2 are first repetitions of a word (50%), second repetitions of a word (26.0%), declaratives (6.0%), as well as hortatives and unclear phrase types (each 4.0%). Interrogatives, negative declaratives, declarative noun phrases, declarative prepositional phrases, and onomatopoeia each account for 2.0%. Cluster 5 is the only cluster where the second repetition of a word is the majority at 44.44%. This is followed by declaratives (13.33%), first repetitions of a word (11.11%), quotatives (8.89%), declarative noun phrases and unclear types (each 4.44%), and concessive phrases (2.22%).

5. Discussion. It is difficult to discover precise links between Patwin intonation and function given the data discussed in Section 4. Such a task is challenging in general without a deeper knowledge of Patwin discourse structure than is currently available (Xu 2011). However, this data nonetheless sheds light both on the question of Patwin tone (as raised in Shafer 1961) and the effectiveness of Contour Clustering on found archival data.

As most of the audio in this dataset came from a word list elicitation session, word repetitions are the most represented data type. The first repetition comprised the majority of all clusters with the exception of Cluster 5, where the majority was instead the second. This demonstrates a remarkable heterogeneity in possible intonational shapes for word elicitation contexts. Found in the data were a large amount of minimal tonal pairs, where a word was repeated multiple times each with differing intonation: most commonly, this consisted of a falling contour followed by a level or upstepped contour. This finding makes it unlikely that Patwin has lexical tone. However, the contours—an assortment of falling, rising, and level shapes—are similar to those found in lexical tone, which likely influenced Shafer’s (1961) conclusion. What this, or the general high level of diversity represented in word list elicitation, contributes to pragmatic meaning is a matter for future investigation.

That all seven cluster types were well attested with word elicitation alone suggests that these contours have a diverse range of functions. Declaratives appear most in high-level Cluster 5 (Fig. 5a, while interrogatives are most represented in rising Cluster 2 (Fig. 5b). Though CC suggests

that the latter cluster is level, an examination of the actual pitch track in Fig. 5b reveals a shallowly rising contour more in line with Lawyer's (2021) observation that Patwin questions tend to rise in pitch.

The connection between CC's averaged contours and the assigned clusters' raw pitch track is not always immediately apparent from the CC output alone. In some cases, different CC clusters seem to somewhat clearly map to a possible phonological distinction. For example, the sharpness of the initial fall in Cluster 4 versus the more high-level beginning to Cluster 6 seems to correspond to the presence of absence of %H in Figs 3a and 3c. In other cases, possibly meaningful in cluster differences are only clear upon examination of the raw pitch tracks, as in the high-level versus upstepped examples in Figs. 4a and 4b. Still in other cases, the averaged CC output does not appear to match the raw output, as with the upstepped phrase in Fig. 5b that was assigned to seemingly level Cluster 5. These impressions should be followed up with a more detailed comparison of CC cluster assignment to a manual phonological analysis such as PaToBI (Björklund 2024).

At present, a cursory examination of the phrases assigned to each cluster reveal mixture of types that would probably not have been classified together in a manual analysis. Use of PaToBI often clarified the existence of contours that were classed separately by CC, despite identical PaToBI analyses (e.g. Fig. 3a and 3b, Fig. 4a and 5a). In the case of Figs. 3a and 3b, this seems to be a case of the same contour 'stretching' or 'shrinking' to fit differing phrase lengths. In such cases, the surface form of these contours appears sufficiently different for CC to assign them separately. The underlying unity of these patterns may not have been discovered if not for the synergy between CC's automatic analysis and a careful manual re-examination. This is one example of a beneficial pipeline between these two modes of examining the data.

In other cases, discrepancies between phrases assigned to the same cluster appears to be a consequence of CC being overly sensitive to ultimately unimportant pitch changes derived from pitch tracking errors. These seem to mostly be a consequence of using older archival recordings, whose general quality was further degraded by background noise. In other instances, pitch was impacted by segmental effects, such as the dramatic pitch lowering before initial voiced stops in Figs 4a and 5a.

6. Conclusion. This study is the first detailed examination of intonational contours in Patwin. The seven clusters suggested by Contour Clustering generally corroborate the six clusters posited by Björklund (2024) in a traditional phonological analysis. However, the connection between CC's averaged clusters and the phonological significance between them is not always clear. In some instances, potential phonological distinctions between clusters were clear in both the CC output and the pitch tracks; in others, distinctions were not obvious from the CC output alone, but appeared in a closer examination of the raw pitch tracks. In other cases, the averaged contour provided by CC did not seem to match the raw pitch track at all. These differences were often illuminated by comparing CC output to a method grounded in traditional phonological analysis, such as PaToBI (Björklund 2024). Differences between the CC and manual analyses seemed to be caused by a variety of factors, including 1) CC's inability to recognize compressed or stretched instances of the same underlying pattern, and 2) pitch tracking errors caused by the poor sound quality of older recordings.

These results joins Babinski and Bower (2022) in extending Kaland's (2021) automated clustering methodology to new data and language families. While the results appear generally successful, care must be taken to mitigate the (sometimes dramatic) effects of poor recording quality (e.g. background noise) on pitch tracking and subsequent cluster assignment. In general, this study indicates the fruitfulness of using automated methods in tandem with traditional ones:

while an automated method like CC can help provide a general overview of the data (particularly useful in archival contexts where the data is not collected by the researcher), traditional analysis can then be used to turn a closer eye to the patterns suggested by CC. These findings also contribute to a better understanding of the usage of Patwin intonation, critical for teaching Patwin faithfully in ongoing language revitalization efforts.

References

- Björklund, Anna. 2024. Towards a transcription system of Patwin intonation (PaToBI). Talk given at the Society for the annual meeting of the Study of the Indigenous Languages of the Americas.
- Boersma, Paul & David Weenink. 2024. Praat: doing phonetics by computer [Computer program]. Version 6.4.06, retrieved 25 February 2024 from <http://www.praat.org/>.
- Bright, Elizabeth. (1952a). Materials for a grammar of Patwin, BrightE.004. In Elizabeth Bright papers on the Patwin language, California Language Archive, Survey of California and other Indian Languages, University of California, Berkeley. <https://doi.org/10.7297/X23R0QRM>.
- Bright, Elizabeth. (1952b). Texts: Coyote tests birds for new chief and creation of heaven and hell, LA 20.006. In the Elizabeth Bright collection of Hill Patwin sound recordings. California Language Archive, Survey of California and Other Indian Languages, UC Berkeley. <http://cla.berkeley.edu/item/18849>.
- Bright, Elizabeth. 1952c. Words and phrases: Phonetic contrasts, LA 20.001. In the Elizabeth Bright collection of Hill Patwin sound recordings. California Language Archive, Survey of California and Other Indian Languages, UC Berkeley. <http://cla.berkeley.edu/item/18844>.
- Dixon, Roland B. & Alfred L. Kroeber. 1913. New linguistic families in California. *American Anthropologist* 15(1). 647–655. <https://doi.org/10.1525/aa.1913.15.4.02a00050>.
- Dixon, Roland B. & Alfred L. Kroeber. 1919. Linguistic families of California. *University of California Publications in American Archaeology and Ethnology* 16(3). 47–118. <https://doi.org/10.1525/9780520340497-007>.
- Golla, Victor. 2011. *California Indian Languages*. University of California Press: Berkeley.
- Heizer, Robert (ed.). 1978. *Handbook of North American Indians Vol. 8: California*. Washington, DC: Smithsonian Institution.
- Himmelman, Nikolaus P. 2008. Prosody in language documentation. In Jost Gippert, Nikolaus P. Himmelman & Ulrike Mosel (eds.), *Essentials of language documentation*, 163–182. Berlin: De Gruyter.
- Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23(1). 225–257. <https://doi.org/10.1017/S0952675706000893>.
- Kaland, C. 2021. Contour clustering: A field-data-driven approach for documenting and analysing proto-typical f0 contours. *Journal of the International Phonetic Association* 7(1). 159–188. <https://doi.org/10.1017/S0025100321000049>.
- Kirby, James & Marc Brunelle. 2017. Southeast Asian tone in areal perspective. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 703–731. Cambridge: Cambridge University Press.
- Kroeber, Alfred. 1925. Handbook of the Indians of California. *Bureau of American Ethnology Bulletin* 78(1). 1–995.
- Kroeber, Alfred. 1932. The Patwin and their neighbors. *University of California Publications in Archaeology and Ethnology* 29(4). 253–423.

- Latham, Robert Gordon. 1856. On the Languages of Northern, Western, and Central America. *Transactions of the Philological Society of London* 3(1). 57–115.
- Lawyer, Lewis. Patwin phonemics, phonetics, and phonotactics. *International Journal of American Linguistics*. 81(2). 221-260. <https://doi.org/10.1086/680310>.
- Lawyer, Lewis. 2021. *A grammar of Patwin*. Lincoln: University of Nebraska Press.
- Leonard, W. Y. 2008. When is an “extinct language” not extinct. In Kendall A. King (ed.), *Sustaining linguistic diversity: Endangered and minority languages and language varieties*, 23-33. Washington, D.C.: Georgetown University Press.
- Maddieson, Ian. 2013. Tone. In M. S. Dryer & M. Haspelmath (eds.), *The World Atlas of Language Structures online*. Frankfurt: Max Planck Institute for Evolutionary Anthropology.
- Perley, Bernard C. 2012. Zombie Linguistics: Experts, Endangered Languages and the Curse of Undead Voices. *Anthropological Forum* 22(2). 133–149. <https://doi.org/10.1080/00664677.2012.694170>.
- Pitkin, Harvey. 1984. *Wintu Grammar*. Berkeley: University of California Press.
- Schoolcraft, Henry Rowe. 1847. *Historical and statistical information respecting the history, condition, and prospects of the Indian tribes of the United States*. Philadelphia: Lippincott, Grambo & Company.
- Sheperd, Alice. 2005. *Proto-Wintun*. Berkeley: University of California Press.
- Shafer, Robert. 1961. Tones in Wintun. *Anthropological Linguistics* 3(6). 17–30. <https://www.jstor.org/stable/30022318>
- Silverman, Kim, Mary Beckman, John Pitrelli, Mari Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert & Julia Hirschberg. 1992. TOBI: A Standard for Labeling English Prosody *ICSLP-ISCA* 1(1). 867–870.
- Whalen, D. H., & Joyce McDonough. 2015. Taking the laboratory into the field. *Annual Review of Linguistics* 1(1). 395-415. <https://doi.org/10.1146/annurev-linguist-030514-124915>.
- Whalen, D. H., Christian DiCanio, & Rikker Dockum. 2022. Phonetic documentation in three collections: Topics and evolution. *Journal of the International Phonetic Association* 52 (1). 95-121. <https://doi.org/10.1017/S0025100320000079>.
- Whistler, Kenneth W. 1977. Wintun prehistory: An interpretation based on plant and animal nomenclature. *Annual Meeting of the Berkeley Linguistics Society (BLS)* 3(1). 157–174 <https://doi.org/10.3765/bls.v3i0.3287> .
- Whistler, Kenneth. W. 1980. *Proto-Wintun Kin classification: A case study in reconstruction of a complex semantic system*. Berkeley, CA: UC Berkeley dissertation.
- Xu, Yi. 2011. Speech prosody: A methodological review. *Journal of Speech Sciences* 1(1). 85–115. <https://doi.org/10.20396/joss.v1i1.15014>.