

David R. W. Sears\*<sup>1</sup>

\*Johannes Kepler University Linz, Austria

<sup>1</sup>david.sears@jku.at

## Family Resemblance and the Classical Cadence Typology: Classification Using Phylogenetic Trees

### ABSTRACT

#### Background

In the *Formenlehre* tradition, contemporary accounts of the classical cadence typically identify the most common cadence categories according to essential characteristics relating to harmony and melody (e.g., Caplin 1998 and 2004). In the perfect authentic cadence, for example, the dominant and tonic harmonies of the cadential progression must be in root position, and the tonic must support  $\hat{1}$  in the soprano voice. I present an alternative view, one that exemplifies the probabilistic approach to category formation adopted by cognitive psychologists over the last half century (Posner 1986; Rosch 1973 and 1978; Rosch and Mervis 1975), in which a category is understood as a network of overlapping attributes, and members are prototypical to the extent that they bear a *family resemblance* to — have attributes in common with — other members of the category.

#### Aims and Repertoire Studied

To support this claim, this paper presents a corpus study of the classical cadence that re-examines the cadence typology presented in William E. Caplin's treatise, *Classical Form* (1998) — represented here by a collection of 245 exemplars selected from 50 sonata-form expositions in Haydn's string quartets (Op. 17–76) of the five cadence categories that achieve cadential arrival (perfect authentic, imperfect authentic, half, deceptive, and evaded) — using a family of techniques for similarity estimation and clustering pioneered by psychologist Amos Tversky.

#### Methods

Classifiers typically depend on some notion of similarity, so I adapt Müllensiefen and Pendzich's (2009) recent implementation of Tversky's *ratio* model (1977), which determines the similarity  $\delta$  between two cadences  $a$  and  $b$  according to the sets of contiguous and non-contiguous sub-sequences (or  $n$ -grams)  $A$  and  $B$  they share that terminate at the end of the sequence.

$$\delta(a, b) = \frac{f(A \cap B)}{f(A \cap B) + f(B \setminus A)}$$

The function  $f(A \cap B)$  measures the salience of the  $n$ -grams shared by  $a$  and  $b$ , and  $f(B \setminus A)$  measures the salience of the  $n$ -grams that are distinct to  $b$ , where salience refers to the prevalence — measured by a statistic called *inverted document frequency* — of each  $n$ -gram in the Haydn Corpus.

To classify the cadences, I then apply an additive clustering algorithm called the *neighbor-joining* (NJ) method (Saitou and

Nei 1987), which visualizes the obtained similarity estimates using phylogenetic trees. In short, the NJ method groups the pair of cadences (or clusters) with the highest similarity estimate and then calculates the similarity between the resulting cluster and every other cadence (or cluster) in the similarity matrix. The algorithm then repeats this procedure until it obtains a single cluster that includes all of the cadences. The resulting tree is thus a metric representation of the similarity matrix, in which the dissimilarity between cadences is represented by the length of the path that joins them.

Figure 1 presents the tree calculated with the NJ method for the cadences from the corpus. Each cadence was partitioned into the five categories from Caplin's typology. The PAC and IAC categories appear at the top of the tree in blue and green, the HC category appears at the bottom of the tree in red, and the DC and EV categories appear on the right side of the tree in magenta and yellow.

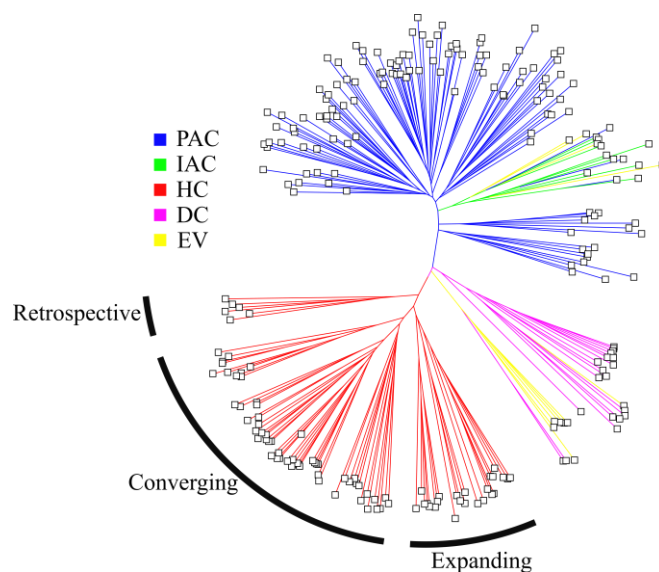


Fig. 1. Equal-angle dendrogram calculated with the NJ method for the cadences from the corpus.

Finally, to identify the most prototypical members in each branch of the tree, I borrow and extend a technique from Tversky and J. Wesley Hutchinson called *nearest neighbor analysis* (1984), which determines how frequently each cadence receives the highest similarity estimate — and thus, serves as the nearest neighbor — for each of the remaining cadences in that branch. This method exploits the fact that some cadences within each class will be more prototypical than others, which is to say that some cadences will share more features specific to the class than others.

## Results and Implications

Table 1 provides the confusion matrix comparing the model predictions with the annotations. Reading along the diagonal, the cluster analysis correctly classified 233 of the 245 cadences in the collection. What is more, for certain categories the NJ method discovered pertinent sub-types that have only recently been described in the *Formenlehre* tradition (Martin and Pedneault-Deslauriers 2015). The three subordinate branches of the half cadence sub-tree shown in Figure 1, for example, correspond quite closely with the *expanding*, *converging*, and *reinterpreted* sub-types.

		Annotation				
		PAC	IAC	HC	DC	EV
Prediction	PAC	120	3	0	0	3
	IAC	2	6	0	0	1
	HC	0	0	84	0	0
	DC	0	0	0	17	1
	EV	0	0	0	2	6

**Tab. 1.** Confusion matrix comparing the model predictions with the annotations.

Example 1 presents the most prototypical cadence from the lower-right branch of the half cadence sub-tree that exemplifies the expanding C–F#–G, in which the (b)6–5 bass clausula supports an upper register  $\hat{1}$  in the soprano that leaps down to  $\#4$  before resolving to  $\hat{5}$ .

**Ex. 1.** Expanding exemplar from the lower-right branch of the half cadence sub-tree (shown in red in Fig. 1). String Quartet in F Op. 17 No. 2, I, mm. 19–20.

Thus, this study provides evidence in support of the view that category systems for the classical cadence are psychologically relevant if they mirror the structure of attributes encountered in a given repertory that listeners are likely to learn and remember, where category membership is determined not by essential features, but by family resemblance.

## Keywords

Musical Cognition, Form, Musical Modelling, Mathematics and Formalisation, Psychology of Music.

## REFERENCES

Caplin, William E., 1998. *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. Oxford/New York: Oxford University Press.

- , 2004. ‘The Classical Cadence: Conceptions and Misconceptions’, *Journal of the American Musicological Society* 57/1: 51–118.
- Müllensiefen, Daniel, and Pendzich, Marc, 2009. ‘Court Decisions on Music Plagiarism and the Predictive Value of Similarity Algorithms’, *Musicae Scientiae* 13/1 suppl.: 257–95.
- Martin, Nathan John, and Pedneault-Deslauriers, Julie, 2015. ‘The Mozartean Half Cadence’, in Markus Neuwirth and Pieter Bergé (eds.), *What is a Cadence? Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire*. Leuven: Leuven University Press, 185–213.
- Posner, Michael I., 1986. ‘Empirical Studies of Prototypes’, in Colette Craig (ed.), *Noun Classes and Categorization*. Amsterdam: John Benjamin, 53–61.
- Rosch, Eleanor H., 1973. ‘Natural Categories’, *Cognitive Psychology* 4, 328–50.
- , 1978. ‘Principles of Categorization’, in Eleanor Rosch and Barbara B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale (NJ): Erlbaum, 28–48.
- Rosch, Eleanor H., and Mervis, Carolyn B., 1975. ‘Family Resemblances: Studies in the Internal Structure of Categories’, *Cognitive Psychology* 7, 573–605.
- Saitou, Naruya, and Nei, Masatoshi, 1987. ‘The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees’, *Molecular Biology and Evolution* 4/4: 406–25.
- Tversky, Amos, 1977. ‘Features of Similarity’, *Psychological Review* 84/4: 327–52.
- Tversky, Amos, and Hutchinson, J. Wesley, 1984. ‘Nearest Neighbor Analysis of Psychological Spaces’, *Psychological Review* 93/1: 3–22.