Salience as a fundamental cue to musical pattern recognition: a Parallel Distribuited Processing approach

Alessandro D'Ausilio°, Frederic G. Piat#, Antonio Mastacchi°, Marta Olivetti Belardinelli°
° Dept. of Psychology, University of Rome "La Sapienza"
#Pierre et Marie lab for Artificial Intelligence University of Paris VI

The aim of the present study is to test if a self-organized Neural network (Piat, 2000) is able to classify four classes of stimuli, previously widely tested by our team, and to replicate human results. The variables considered here are salience and tonality. We consider the first as a fundamental dimension to make the recognition of musical fragments possible, on the other hand we understand, both from our simulation and the results on human subjects, that tonal cues do not have such importance. The theorical frame of this work is Tulving's model of memory, by the assumption that in subjects the salience carries out Remember answer and tonality Know answer respectively as an expression of the episodic and semantic store.

A great part of the distributed model, designed for musical stimuli recognition, seems to be unrealistic, for two reasons: first, because they are based on a tonal *a priori* knowledge; the second reason is because these models do not include a self-organizing learning phase, as happens in human learning. Therefore in this work, we try to verify if a neural network is able to classify 4 groups of musical stimuli as human subjects are able to. We used Piat's ARTIST model, that is not designed with an a priori knowledge except for the categorical/perceptive discrimination of pitches.
Piat's model is made up of two layers of neurons, the input field (F1) with 72 nodes (12 semitones for 6 octaves), the second field (F2) for categorization. The connections between layers are both bottom-up and top-down. Learning occurs through the modification of weights, that progressively tune the F2 layer to be most responsive to a certain input pattern. The synaptic weights between F1 and F2, store the long-term memory of the model. Each note activates its corresponding neural node input proportionally to its loudness, and according to a temporal exponential decay. Subsequently, a similarity measure between the input pattern and the F2 category is computed; then, with a "winner takes all" strategy, the most activated is chosen, and the others set to zero. At the end of the cycle the winner node propagates back to F1 (top-down). Upon learning, a parameter (vigilance) is used as a changeable threshold, whether or not to admit an input into a category or to form a new one.
It is clear, in its simplicity, that the model has no *apriori* tonal knowledge, but it is extracted from the input; the discriminative learning ability shown by the net, is not subject to any law except for the recursive activation between F1 and F2 and vice versa. We adopted Piat's model because our team's previous research has clearly demonstrated that a subject's musical experience is based on superficial characteristics, and not on a deep tonal analysis of the musical pieces.
Such research has highlighted that the salience, operatingly defined as redundancy of interval or rhythmic parameters, is one crucial dimension for the recovery of melodies from memory.
This result has been obtained by means of four classes of musical stimuli constructed for this purpose, intercrossing two variables: the tonality and the salience, both on two levels: presence and absence. These two variables give place to four classes of stimuli, respectively: Tonal-Salient, nonTonal-Salient, Tonal-nonSalient, nonTonal-nonSalient.
To these specification two series of 48 stimuli have been created, with an average length of 7 seconds, equally distributed between the 4 classes, the first composed by Fabio Cifariello Ciardi and the second by Massimo Caltagirone. As we used these stimuli in repeated experimental research, we were able to demonstrate that, according to Tulving's model, the correct recognition is due to the episodic memory (Remember answers), however the semantic memory produces more generic familiarity impressions (Know answers).
Our hypothesis is that the first modality would have to be based on a superficial analysis of the stimulus characteristics, tied to a single interaction experience with it. The second one, instead, with the listening repetition and a consequent generalization and extraction of rules, could take

advantage of more complex and hierarchical elaborations, which need a high level of musical competence, able to codify all the characteristics of tonal relations in the piece, in issue, in real time. The innovation of the experimental data achieved with these stimuli on human subjects is that the Remember answers are tied to the characteristics of salience of the stimuli, and the Know answers are connected to the categorization based on tonality.
The input of the 2 series into the net produced results consistent with our hypothesis, that the variable salience would provide a better system of classification by the net. In fact, the presence of this variable in association with the tonality produces the best classification and with a low degree of dispersion around the prototypes of its class. The ability of the net is a little bit weaker in classifing stimuli S-NT. On the contrary the tonality alone, as we hypothesised, does not allow the same degree of fine discrimination. That is in agreement with the results obtained with human subjects: the single presence of the tonality produces confusion in recovery from memory. The absence of tonality and salience, as it was simple to imagine, allows the net to produce as many classes as stimuli.
The conclusion is that the tonality, even if it is difficult to find an exhausting definition, is surely an important dimension in the musical acculturation process. But it is necessary to take into consideration other more superficial variables. Here the salience has been considered as one of the most important variables in naives subject to musical recognition, in this paper our hypothesis is strengthened by the results collected with a self-organizing network.

References

Bharucha, J. J., (1987). *Music cognition and perveptual facilitation:A connectionist framework*, Music Perception, 5(1), 1-30.

Gjerdingen, R. O., (1990). *Categorization of musical pattern by a self-organizing neuronlike networks*, Music perception, 7, 339-370.

Piat, F. G., (2000). *ARTIST: A connectinist model of music acculturation*, Proceedings paper of ICMPC 2000.

Olivetti Belardinelli, M., (2000). *Structuring factors in musical processing*, Proceedings of the 6th international Conference on Music Perception, Keele University, August 5-10.

Olivetti Belardinelli, Rossi-Arnaud, C., Pitti, G.,Vecchio, S. M., (2000). *Looking for the anchor points for musical memory*, Proceedings of the 6th International Conference on Music Perception, Keele University, August 5-10.

Olivetti Belardinelli, M., Rossi-Arnaud, C., (1999). *Recollection and familiarity in recognition memory for musical themes*. XI Conference of European Society for Cognitive Psychology. Proceedings (A. Vandierendonck, M. Brysbaert, K. Van Der Goten, Eds.), Academia Press, p. 193.

Tulving, E., (1985). How many memory system are there?, American Psychologist, 40(4), 385-398.