# Chapter 13

# Repetition

Repetition is a major aspect of music. The theory of musicality as an aspect of speech perception forces us to ask why exact repetition (free and non-free) occurs in music, even though it does not occur in normal speech.

Human perception of speech melody is intrinsically time translation invariant, and this creates problems when repetition or near-repetition occurs *within* a single speech melody. The solution is to maintain a **repetition count**, i.e. to distinguish the first occurrence from a second occurrence of a melodic fragment within a melody.

A secondary question is: When does the brain know *not* to keep count any more? A suggestion is that those features that normally come at the end of a melody may serve the function of resetting the repetition count to zero.

## 13.1 Repetition as a Super-Stimulus

Recall the relationship between aspects of music and the perception of speech:

- An aspect of speech is perceived by a cortical map.
- A corresponding aspect of music is a super-stimulus for that particular cortical map.

As a result, we sometimes see features of music that appear not to exist at all in speech, for example the occurrence of musical scales. Even when we can recognise the similarity between a musical aspect and a speech aspect, such as the rhythms of speech and the rhythms of music, the musical version may have regularities not apparent in the speech version.

Repetition is an aspect of music where there is a high degree of regularity, with apparently no analogue in speech. Musical phrases are sometimes exactly repeated within a tune without any variation at all, and are often repeated an exact number of times—usually twice, sometimes more.

This kind of exact repetition does not normally occur in speech (although there is one major exception—see Section 13.7 which discusses **reduplication**), and speech would generally sound strange or contrived if it did occur. So what's going on?

I defined **free** and **non-free** repetition in Chapter 4. Free repetition is where the major components of music are repeated freely, such as choruses and verses of a song.

In some cases a tune exists in a cyclic time frame, in the sense that the end of one repetition blends directly into the beginning of the next repetition, and the musicality of each portion of the tune depends on what precedes it and what follows it, so the performer has no choice but to perform endless repetitions of the tune. In other cases the tune comes to a stop before starting again each time, but is still repeated an indefinite number of times as part of a single performance. Many recorded performances of popular songs come to an end by "fading out", suggesting that those producing the songs could not find a satisfactory way to end them. The only thing that prevents a song containing freely repeated components from repeating them forever is that the audience will get bored if the song goes on for too long.

Non-free repetition is perhaps of more interest. Components repeated non-freely are components *within* a major component of a song. They can range from single notes, to portions of a bar, to as much as a quarter of the song, e.g. a tune might take the form $[AABB]$.[1] The non-free aspect is that the repetitions occur a fixed number of times. Thus each occurrence of the repeated phrase has assigned to it some count of its location within the repetition. (With free repetition there is no sense of keeping count, unless perhaps we keep count consciously: freely repeated verses and choruses just go on and on.)

This suggests that somewhere there is a cortical map that keeps this count of the number of repetitions that have occurred for a phrase. In as much as repetitiveness is a perceived quality that we can be consciously aware of, there probably exists some cortical map representing it, since most perceptual qualities have corresponding cortical maps that process and represent them.

There is at least one difficulty with this theory of a cortical map that encodes a repetition count: it has to deal with **nested repetition**. This is where a non-freely repeated component itself contains non-freely repeated

[1] Here $A$ and $B$ etc. are used to refer to particular phrases (not notes), so that, for example, $[AA]$ refers to some phrase $A$ being repeated twice.

components. For example, a tune might have structure $[ACBBACBB]$, where the component $B$ is repeated twice inside the repeated component $[ACBB]$. Such nested repetition does occur within popular and traditional music—a good example is "Funiculi Funicula" (Denza & Turco, 1880). The representation of nested repetition in a cortical map would appear to require a separate dimension of count for each nested level of repetition.

Possibly even more common than exact non-free repetition is **partial** non-free repetition. A melody may have distinct components $[ABCD]$, but reduced to the rhythm only it may read $[A'A'C'D']$, where the first two phrases have different melody, but the same rhythm. There are many common variations on this partiality:

- The sequence of notes may be the same, but the rhythm may be different.
- The rhythm and the up and down contour may be the same, but the melody is translated up the scale, so that the exact intervals between corresponding pairs of notes in the partially repeated phrases are different.

Another variation is repetition of the beginning of a phrase, which may be either exact or partial, but then a variation occurs at the end of the phrase. This type of variation is often associated with a sense of progression.

Whereas non-free repetition requires the inclusion of a repetition count in the perceptual state, there is no such requirement for partial repetition, since the aspects of the music not being repeated provide the information that distinguishes the first repetition from the second.

## 13.2 Reasons for Perception of Repetition

If we were writing a computer program to perceive melodies as sequences of notes, there are two ways that we might deal with repeated sequences:

1. The program could ignore repetition. The melody is just treated as a sequence of notes, all of which are individually recorded and processed by the program. If repeated sequences happen to occur, the program doesn't care; it just processes each repetition in due course.
2. Or, the program could be written so as to recognise repeated sequences. The program would have to include some definition of what was a significant repetition, for example some minimum length of a repeated sub-sequence. When a repeated sequence was recognised, instead of processing it all over again, the program could just record that the repetition had occurred, and it would record which previous sub-sequence

of the melody it was that was repeated. Some file compression algorithms work this way.[2]

As I have stated previously, the human brain does not always solve a problem the same way that we might program an electronic computer to solve that problem. A computer programmer writing a program to process sequences of values in melodies would probably have the program write each sequence of note values that it was processing to a corresponding series of numerically indexed locations in memory. The index values would form an implicit global frame of reference against which the note values were indexed. The index values could be used to identify the location of a sequence of values that had already occurred and which was being repeated. It is, however, very unlikely that the brain uses any type of numerical indexing system to store the data it processes.

## 13.3 Perceptual State Machines

A **state machine** is a system that has a set $S$ of possible states, a set $E$ of possible events, and a **transition function** $F$ which maps each pairing of input state $s_{in}$ and event $e$ to an output state $s_{out}$. We can use this concept to model how the brain processes sequences of values such as notes in a melody: the state corresponds to the state of activity in a cortical map, and the events correspond to the information coming into the cortical map about each musical note. For a given initial state $s_0$, we can model the perception of a melody as the updating of the state by the sequence of events representing the notes of the melody.

If the transition function is such that the state machine's current state has no dependence on more than the previous $N$ values, then the state machine will automatically recognise repetitions of $N$ or more notes, in the sense that it will always be in the same state at the end of two identical sub-sequences of $N$ or more notes.

A state machine whose state depends on only a limited set of previous values is **forgetful**, because all past history eventually gets "forgotten" by the state machine.

A state machine with this forgetfulness property recognises a repeated sequence, but the disadvantage is that the machine is then completely incapable of knowing how many times the sequence has been repeated, since it is always in the same state at the end of a sequence, whether it be the first repetition or the second or the hundredth.

Forgetfulness is related to time translation invariance of perception. We wish the response to a sufficiently extended sequence of values to be the same

[2] See *A Universal Algorithm for Sequential Data Compression* Jacob Ziv and Abraham Lempel (IEEE Transactions on Information Theory 1977)

whenever that sequence occurs; thus the state of a system responding to the sequence must not maintain any state information too persistently.

### 13.3.1 A Neuronal State Machine

The following is a simple model of a neuronal state machine that responds to the occurrence of a sequence of information values representing the sequence of notes in a musical melody:

- Each note value is represented by a group of sub-values, where each sub-value relates to one aspect of the music. These sub-values are derived from the symmetry-invariant encodings of the notes in the melody.
- Simplifying slightly, each sub-value is represented by the activation of a neuron in a corresponding cortical map (simplified in that we are ignoring population encoding).
- There is a **state cortical map** within which individual neurons encode for individual states.
- The neuron for the current state and the neurons for the sub-values of the current value in the sequence activate the neuron in the state map representing the next state.

## 13.4 The Flow Model

Suppose that there are $n$ sub-values per note value. Each value is therefore a point in an $n$-dimensional space. Suppose, hypothetically, that each value corresponding to a note in the melody is unique within that melody. And further suppose that each point in the $n$-dimensional space is represented by a neuron. Then in order to represent the sequential progress of the melody, all we need is a connection from each neuron to the next neuron in the $n$-dimensional space, such that each neuron activates the neuron representing the next step in the melody.

We can imagine these connections between neurons as being like a sequence of arrows in the $n$-dimensional space. In effect the $n$-dimensional space is the same as the space of possible states. We start at the state representing the beginning of the melody, we follow the arrows, and eventually we reach the state representing the end of the melody.

This model is somewhat idealised. There are at least two major objections to it as a realistic model of how the brain represents and processes information about melodic sequences:

- Cortical maps are not $n$-dimensional; in general they are no more than 2-dimensional with regard to representing numerical values.

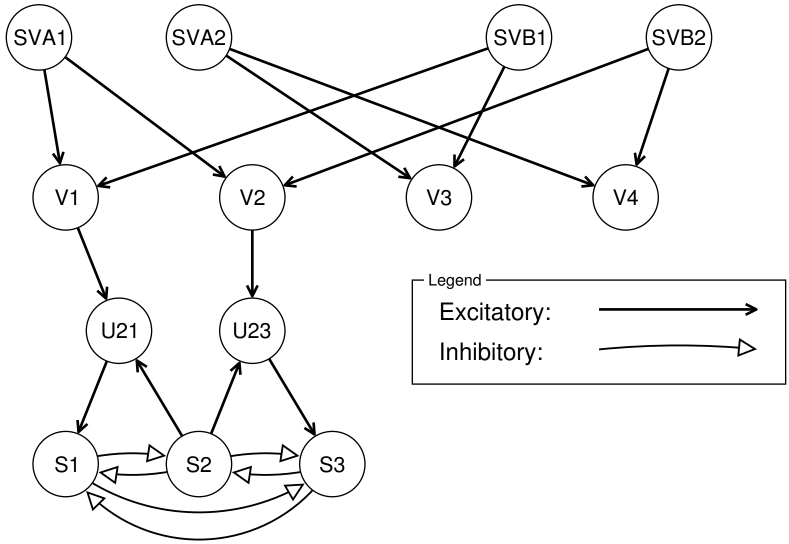


**Figure 13.1.** A neuronal state machine representing melody perception. Neurons SVA1 and SVA2 represent possible values of sub-value A; neurons SVB1 and SVB2 represent possible values of sub-value B. The different combinations of sub-values form the full values as represented by neurons V1, V2, V3 and V4. Neurons S1, S2 and S3 represent three possible states of the state machine. Update neuron U21 represents the rule that value V1 (i.e. sub-values A1 and B1 combined) should cause state S2 to transition to state S1. Similarly neuron U23 represents the rule that value V2 (sub-values A1 and B2 combined) causes state S2 to transition to state S3. Mutual inhibition between the state neurons ensures that only one state is active at a time, and that the transition to the new state deactivates the old state.

- If the state in the $n$-dimensional map ever repeats, then our model of the melodic sequence will get stuck—it will be forced to go around forever in the same loop. We can explain this better by considering what happens when we reach a particular state for the second time: the flow of state changes is entirely determined by the current state and the arrow from that state to the next, so the next state after visiting a state a second time has to be the same next state that happened when the state was visited the first time.

We will find a pragmatic way to solve the 2-dimensional limitation, but first I will look at the "stuck-in-a-loop" problem.

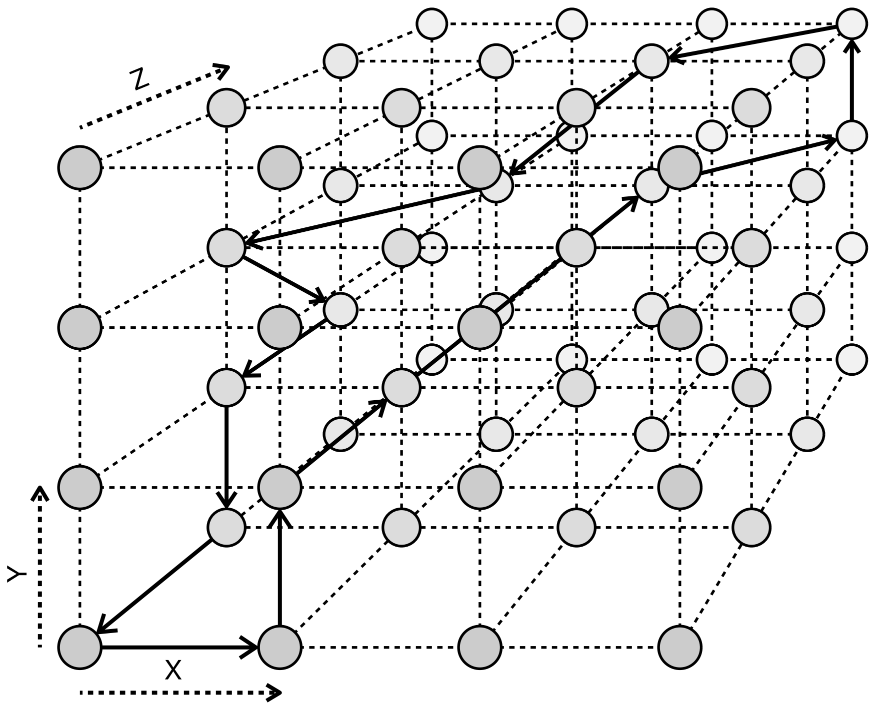


**Figure 13.2.** N-dimensional state. A hypothetical N-dimensional cortical map (in this case N=3, as that is the most that can practically be shown on a diagram). The arrows show a closed-loop path through the map. Comparing to Figure 13.1, the neurons shown simultaneously represent the "V", "U" and "S" neurons, and the arrows represent the connection from each "S" neuron via a "U" neuron to the next "S" neuron.

### 13.4.1 Breaking Out of the Loop

When we get to a state in the $n$-dimensional space for the second time, what we really want to know is that we have already been there before, and we want to provide this information about having been there before as *an extra dimension* of our state. The connections between neurons will then be able to take into account this extra dimension of information, to in effect say "we are now here for the second time, so don't follow the arrow we took the first time, instead follow this other arrow".

### 13.4.2 Almost Exact Repetitions

Musical repetitions are often exact repetitions. The repetition count (of how many times we have been here before) will be 0 the first time around, 1 the

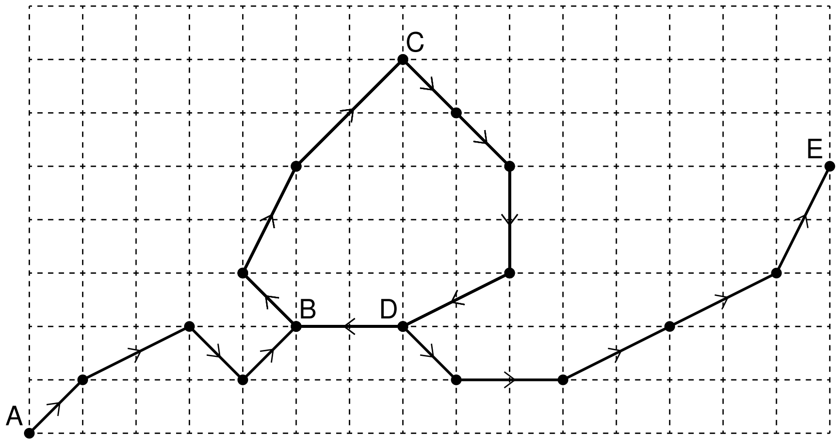


**Figure 13.3.** A path representing flow in a state space (here assumed to be just 2-dimensional). The flow starts at point $A$, goes to point $B$, goes around the loop $B$ to $C$ to $D$ three times, and then exits the loop at $D$ and finishes at $E$.

second time round, 2 the third time, and so on. These exact repetitions do not normally occur in speech,[3] yet if we suppose that there exist special mechanisms for the perception of repetition in melody, then those mechanisms must presumably exist for the purpose of perceiving speech melody.

Although speech melodies do not contain exact repetitions, it is entirely possible that they can contain repetitions that are close enough to being exact to cause a partial occurrence of the problems caused by exact repetition, and for which the addition of repetition count as an extra dimension of information is required. Thus the system not only represents "we have been here before", it also supports "we have been close enough to this spot before that it might cause confusion". This could be regarded as a fuzzy numerical attribute, i.e. 0 represents "we have not been here before", 1 represents "we have been here before", and values between 0 and 1 represent "we have been close to this point before".

### 13.4.3 Faking $n$ Dimensions in 2-Dimensional Maps

The other problem with the flow theory is that cortical maps are only 2-dimensional (with a very thin 3rd dimension that would not be able to represent a continuous numerical attribute), whereas the flow is in an $n$-dimensional space of perceived values.

[3]Except for the very short exact repetitions caused by **reduplication**, as discussed in more detail in Section 13.7.

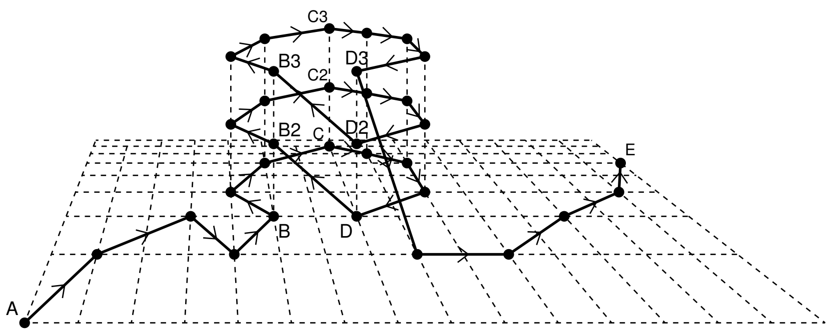


**Figure 13.4.** As in Figure 13.3, but now "Have we been here before?" is represented as an additional dimension. The flow goes from $A$ to $B$ to $C$ to $D$. When it comes around to $B$ a second time the repetition is represented by being lifted up to $B2$. From there it loops around $C2$ to $D2$ to $B3$ to $C3$ to $D3$. When it exits the loop at $D3$ the flow returns to "ground level" because the state is no longer a repetition of a previous state.

Given $n$ dimensions, we could imagine all possible **coordinate projections**[4] to 2-dimensional subspaces representing all possible pairs of the $n$ dimensions. There will be $n(n-1)/2$ such subspaces. For each of these subspaces, and for each melody, we can define a flow of motion in a hypothetical cortical map that represents the subspace. In each subspace there will be a path corresponding to the flow of the melody in that subspace. There will be many more collisions in the 2D subspaces than there are in the original $n$-dimensional space.

What happens if we leave the direction of flow undefined at the points where these collisions occur? Even if a collision occurs in one 2D subspace, there will not necessarily be a collision in all the other subspaces. It is likely that the flow will be defined in a sufficient number of other subspaces that the direction of flow can be fully reconstructed in the $n$-dimensional space.

The one occasion where a collision will occur in all the 2D subspaces is when there is an exact repetition, and this is precisely where some means of introducing repetition count as an extra dimension is needed in order to maintain enough state information to record the full path of the melody.

[4]A **coordinate projection** from an $n$-dimensional space to an $m$-dimensional space is a mapping defined by a subset of the numbers 1 to $n$, of size $m$, such that coordinate positions in the subset are retained, and coordinate positions not in the subset are not retained. For example, for $n = 5$ and $m = 2$, the subset $\{2, 4\}$ defines a projection that maps the point $(32, 67, -9, 21, 8)$ to $(67, 21)$.

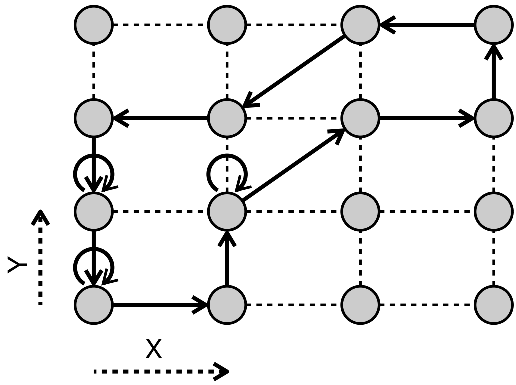


**Figure 13.5.** Projection of Figure 13.2 onto X and Y dimensions. In the original 3 dimensions the state loop shown formed a simple loop that defined a unique transition from each state in the path to the next one. This is no longer true in the projection because in some cases different points in the original path project down to the same points in the projection (because they only differ in their Z coordinate).

## 13.5 Non-Free Repetition: Summary

The theory of musical repetition given here is not as fully developed as the components of my theory relating to other aspects of music. A full understanding of repetition and the mechanisms of recording sequential information in cortical maps is one of the missing pieces of the puzzle that is required to properly complete the theory (see Chapter 15 for a fuller discussion of the incompleteness of the super-stimulus theory).

Within the framework of the super-stimulus theory, the following is a summary of my current understanding of repetition as it occurs in music:

- Exact non-free repetition is a common and well-defined aspect of music.
- Musical repetition must be a super-stimulus for an aspect of speech melody perception.
- Special perception of repetition is required in a system that is designed to work when perceiving information that is not repetitive.
- The characteristics of a perceptual system that enable it to automatically recognise repetitive sub-sequences also cause that system to fail when it is perceiving sequences that contain non-freely repeating sub-sequences.

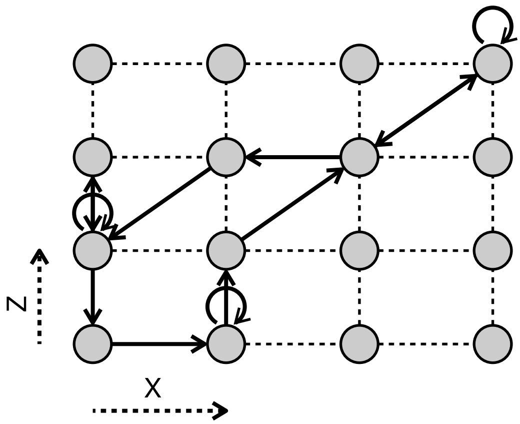


**Figure 13.6.** Projection of Figure 13.2 onto X and Z dimensions.

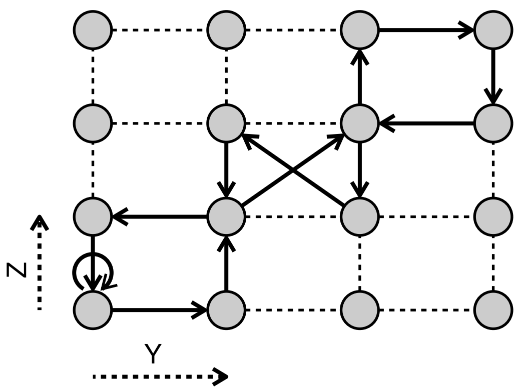


**Figure 13.7.** Projection of Figure 13.2 onto Y and Z dimensions.

## 13.6 Free Repetition and Home Chords

The theory of home chords given in this book explains how the home chord is determined, and what perceptual purpose is served by the cortical map that makes this determination. The theory also explains why a home chord occurs at the start of a melody: the diatonic scale causes more than one stable state to be possible in the home chord cortical map, and the map does not enter one of these states until the corresponding chord actually occurs.

This leaves unexplained the second major fact about home chords and notes, which is their occurrence at the *end* of a melody. Associated with the occurrence of the home note and chord at the end of the tune is the tendency of the home chord to be preceded by a **dominant 7th** chord. Referring to examples on the white notes scale with home chord C major or A minor, the chord that precedes the home chord is the dominant 7th chord with root note a perfect fifth higher than the root note of the home chord, i.e. G7 = GBDF precedes C major, and E7 = EG♯BD precedes A minor.

Another common feature of the home note/home chord combination is the length of the last note. Very commonly the final home note is a single note that starts simultaneously with the final occurrence of the home chord at the start of the final bar, and continues for all or most of that bar.

Taken together, these features of a final note/chord combination that define the end of a tune are called a **cadence**.

I have not been able to discover any convincing explanation of why this combination of chords and a long final home note wants to occur at the end of a tune. But I can make one pertinent observation:

> A tune cannot freely repeat, until it has first *ended*.

So it may be that the purpose of a home chord, optionally preceded by a dominant 7th, is to *end* a tune by resetting the state of some or all of the cortical maps involved in perception of music/speech (in particular resetting any repetition counts), so that the tune can then be freely repeated.

We have seen that the perception of non-free repetition requires a keeping of the repetition count. If we consider a cortical map that is responding to a repeated sequence, especially one that repeats from the beginning of the tune so far, we might ask how the cortical map knows if it is meant to be a non-free repetition or a free repetition. In other words, is it meant to be keeping count or not? Keeping count for a free repetition would introduce a spurious dimension of perception into the perception of the music. Failing to keep count of a non-free repetition would remove a dimension of perception that was required.

It is possible that the default is to assume all repetition is non-free, and that the effect of a cadence is to reset the repetition count (of everything) back to zero, so that any following repetitions are perceived as free repetitions. A cadence might perhaps represent a prototypical sentence ending, prototypical in the sense that the human brain is partly predetermined to end a sentence with an intonation resembling a cadence, even though specific languages may adopt alternative intonations for the ends of sentences (in effect overriding the predetermined default).

In a musical cadence, the state of the harmonic cortical map matches the state of the home chord cortical map. It is possible that the same match occurs in a speech "cadence", even though, in the case of speech, neither map would be in a state corresponding to a musical chord.

## 13.7 Reduplication

There is one major exception to my earlier assertion that exact repetition does not occur in natural language. This is the phenomenon of **reduplication**. Reduplication is where all or part of a word is duplicated within the word to make a new word.

One family of languages where reduplicated words are common is the family of Polynesian languages. For example, in New Zealand Maori, "toru" means "three" and "torutoru" means "few".[5] In Hawaiian (another Polynesian language), "wiki" means "hurry" and "wikiwiki" means "quick".[6]

Reduplication is conspicuous by its absence in English and most other Indo-European languages. Reduplicated words sound strange to the English ear, and one could suppose that we positively avoid constructing them. Perhaps we are always in such a hurry to say what we want to say that saying something twice seems like a waste of time. The nearest we get to using reduplication is the use of phrases like "itsy-bitsy", "hodge-podge", "lovey-dovey", "shilly-shally" and "hoity-toity", all of which are highly colloquial and informal in their usage.

Exact non-free repetition occurs in music on a much larger time frame than lexical reduplication. But we cannot rule out the possibility that the same cortical map is responding to both forms of repetition. After all, we are supposing that music is a super-stimulus for speech perception, so the long non-free exact repetitions in music may be a super-stimulus for perception of the short non-free exact repetitions caused by reduplication.

[5] *The Reed Reference Grammar of Maori* Winifred Bauer (Reed Books 1997)
[6] *Hawaiian Dictionary* Pukui and Elbert (University Press of Hawaii 1977)