

ジュウシマツの歌文法研究

ジュウシマツの歌（さえずり）は、とても興味深い性質を秘めています。例えば、育った環境によって歌の構造が異なったり、自分の歌を維持する為に個別に練習したり…。

ジュウシマツの歌の構造は複雑で個体ごとに異なり、それはどんな歌を聴いて育ったのかに大きく依存します。また、脳構造やその働き、発生システムなど人間との類似点が多数存在します。そこで、人間の音声言語獲得の解明にむけた良モデルになるのではないか？と期待されています。

本研究室では理化学研究所と共同で、ジュウシマツの歌の構造を理解すべく日々研究を進めています。以下、研究概要を手順に従って解説していきます。



白色：ジュウシマツ
黒色：その祖先種コシジロキンパラ

1. ソナグラムと音素の切り分け

ジュウシマツの歌を録音し、ソナグラムで音の**パワースペクトル**つまり音の強弱を表示させると右図のようになります。塊ごとに分かれてますね？この1つ1つを**音素**と言います。同じ形の音素には同記号を割り当てていきますが、この作業はとても大変です。そこで、この音素の記号割り当てを自動化しよう!! というのが、この過程での目的になります。

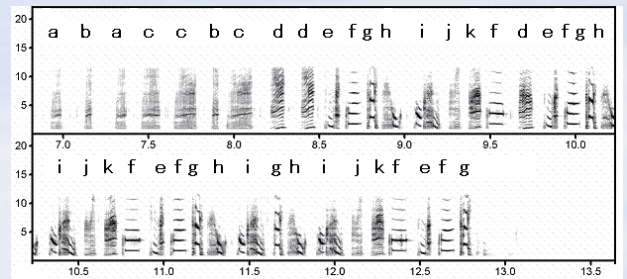
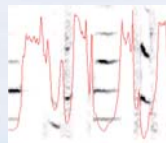


図1：ジュウシマツ (Hikari) の歌のソナグラム表示

WienerEntropy の出力



- (1) WienerEntropy という関数を用いて各音素に切り分けます。
- (2) 各音素を定量化することによって特徴量を取り出し、それらと比較します。
- (3) コホーネンネットを用いて類似性の高い音素を集め、それらには同記号を割り当てます。

2. チャンクの選定による構造の解析

音のデータを記号のテキストデータにすることができました（図1）。次に、それを用いて歌の構造を解析することを考えます。ジュウシマツの歌は、図1の ghi のような特定の音素のパターン、**チャンク**で構成されています。これは、人間の言語でいう単語に相当するものであり、まずどの音素の組がチャンクであるかを判断する必要があります。これを達成するために、**N グラムモデル**を用います。N グラムモデルとは、あるN個の単語（音素）の後にはどの単語（音素）が、どのくらいの確率で現れるのかを予測するものです。（例えば、音素 a の後には、30%の確率で音素 b が、70%の確率で音素 c が現れる、というように構造を解析することができます。）

これを用いて図1のテキストデータを解析すると、図2のようになります。この解析結果から、どれがチャンクと言えるのかを判断し、より分かりやすい形で表すことを目指しています。

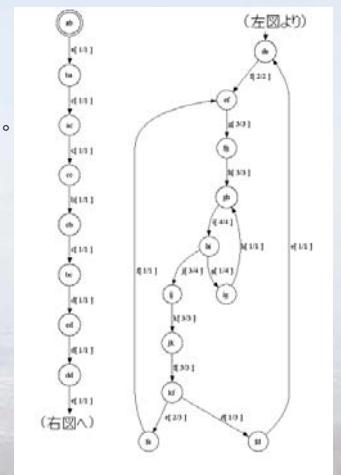
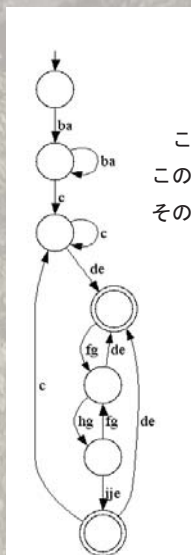


図2：図1のNグラムによる解析

3. 歌文法の抽出

この過程では、ジュウシマツの歌を言語理論に基づいて解析するという研究を行っています。

現在、**k 可逆言語**という枠組みを用いた解析で、興味深い結果が得られています。次にその一例を示します。



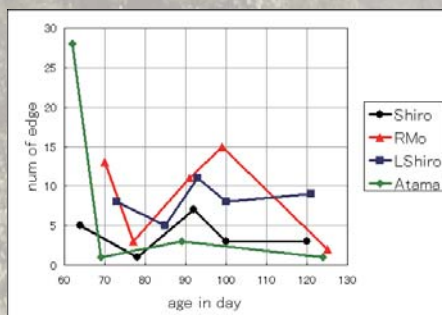
例：（歌データの一部）

↓チャンク化

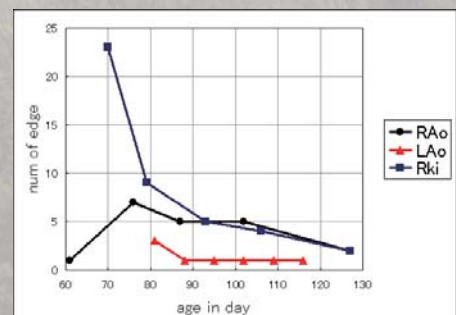
babababacdefghgfghgfghgijedefghgfghgije... ba ba ba ba c de fg hg fg hg fg hg ije de fg hg fg hg ije...

これに k 可逆言語の学習アルゴリズムを適用することで左図のように歌を表現することができます。

このように歌文法を可視化することで文法の持つ複雑さ等を評価することができます。これは7羽のジュウシマツの歌を数日おきに録音し、その文法の複雑さ（遷移の数）の変化をプロットしたものです。文法の変化のタイプが大きく二種類に分かれていることがわかります。



タイプ1：90日齢前後で遷移数が増加



タイプ2：成鳥になるにつれ遷移数が減少

図3：k 可逆言語の

学習アルゴリズムを適用